

序

材料是人类技术进步的标志。人们习惯用石器时代、青铜时代和铁器时代来表达人类文明史的不同阶段，而我们也亲身经历了从钢铁时代向以硅芯片为代表的电子材料时代的过渡。当今，材料被广泛地应用于各式各样的常规及高精尖技术：诸如机械、运输、建筑、能源、医疗、通信、计算等，发挥了十分关键的作用。而材料本身也品种繁多，五花八门：诸如金属与合金、半导体、电介质、陶瓷、玻璃、塑料、水泥乃至层出不穷的复合材料。材料科学与工程所探讨的是材料的制备、结构、性能与功效之间相互关系，涉及的领域十分宽广，内容丰富多彩，但也头绪纷纭，令人有应接不暇之感。

和人类使用材料的漫长历史相比，科学家研究材料科学的历史比较短暂。它始于 19 世纪中叶，当时关注的问题是钢铁热处理过程的显微组织的变化，研究的工具限于金相显微镜。随后是相平衡的热力学与统计力学的建立，为探讨材料的相平衡与相变提供了理论基础。20 世纪 20 年代，原子结构的阐明，量子力学的建立又为理解材料的根本结构提供了新的理论依据。另一方面 X 射线衍射、电子衍射与电子显微术，乃至多种电子隧道探针技术，又为探测材料的微观结构提供了新的手段。固体物理学（后延拓为凝聚态物理学）与物理化学为全面理解材料结构与性能关系提供了理论基础。合成化学的发展又为合成层出不穷的新材料提供了手段。到 20 世纪 50 年代，金属学（包括化学冶金与物理冶金）业已初具规模。随后半导体技术的发展又对材料的纯度和完整性提出了苛刻的要求，促使材料的研究向纵深发展。另一方面，金属学的理论与实验手段也向陶瓷学的领域延伸。到 1960 年左右，“材料科学”的名称问世，迄今已有 40 余年的历史。随后非晶材料科学、液晶材料科学和高分子材料科学又相继问世，丰富了材料科学的内涵。在 20 世纪末，软物质科学蓬勃发展，将液晶、高分子、胶体等领域的研究贯通起来，和传统硬物质科学形成鲜明的对照，软硬兼顾，相得益彰，从而将无机材料科学和有机材料科学辩证地融合起来。

材料科学的基本原理根植于凝聚态物理学、物理化学与合成化学。但由于各种材料的分支学科的学术背景不尽相同，诸分支学科的汇总融合也有一个历史过程，不能一蹴而就。传统的材料科学教科书也反映了这一历史情

况，流于分门罗列、兼容并蓄，而未能熔铸成一整体。读过之后，难免使人有支离破碎，杂乱无章之感。因此，迄今为止，在全世界范围之内，尚欠缺一本能够概括材料科学的众多侧面，而又井然有序、条理清楚的入门教科书。值此世纪之交，学科的交叉融合正在加速进行，这就为将材料科学融贯成为一门成熟的学科提供了必要的基础。这就是我们撰写这本书的时代背景和面临的挑战。

本书力图超越传统的处理方法，引入连贯统一的理论框架来概括不同类型材料丰富多彩的特征，希望能够做到既充分反映了材料的多样性，又体现了科学在概念上的统一性。为了本书主体鲜明，纲举目张，就不得不删除枝节，突出主干。因而本书的重点就放在材料本身的规律性上，对表征材料的实验技术和制备材料的具体工艺只是略有涉及，未作详论。我们并不低估这些问题的重要性，而是认为这部分内容放在其他课程中学习为宜。另外，本书也和固体物理学做了适当的分工，以免相互重复。本书既可以作为高等院校讲授材料科学的入门教科书，也可供从事这方面的研究工作者参考，还可以视为在一定程度上体现了我们对于材料科学整体化观点的一本著作。由于我们才疏学浅，编写过程十分仓促，尚希各位专家学者给予批评指正。

本书的编撰得到国内材料科学界许多专家学者的参与和支持。其中包括中科院金属所的叶恒强院士，中南工业大学金展鹏教授及刘立斌博士，清华大学熊家炯教授及朱嘉麟教授，南京工业大学曾燕伟教授；南京大学的翟宏如教授，李齐教授，袁仁宽教授，丁世英教授，王牧教授，王慧田教授，张荣教授，马余强教授，陈延峰教授，韩民副教授，吴迪博士，濮林博士，杨斌博士等。特别值得注意的是，本书既凝聚了不少老中年专家的心血、学识和智慧，也发挥了许多杰出青年科学家（其中有4位长江计划特聘教授）的探索精神与创新思维。我们特别感谢北京科技大学柯俊院士对本书的编写给予的关怀和支持，他在百忙之中还对书稿中涉及材料发展史方面的内容予以审阅；我们也感谢中国科技大学侯建国教授提供的图片。我们还感谢参加2000年夏在南京召开的《材料科学导论》研讨会专家们对编写本书所提出的宝贵意见和热情鼓励。国内材料科学界还有许多科学家对本书的编写提出了建议及提供图片资料，使我们获益匪浅。

国家教育部面向21世纪教学内容和体系改革项目，南京大学教务处，南京大学材料科学与工程系和南京大学物理系对作为本书重要基础的前期教学研究曾给予长期支持。南京大学材料科学与工程系，南京大学物理系，南京大学固体微结构实验室的同事们对本书提出了许多有益的建议。南京大学相关科研组的研究生在书稿的资料收集、文字加工、绘图等方面做了

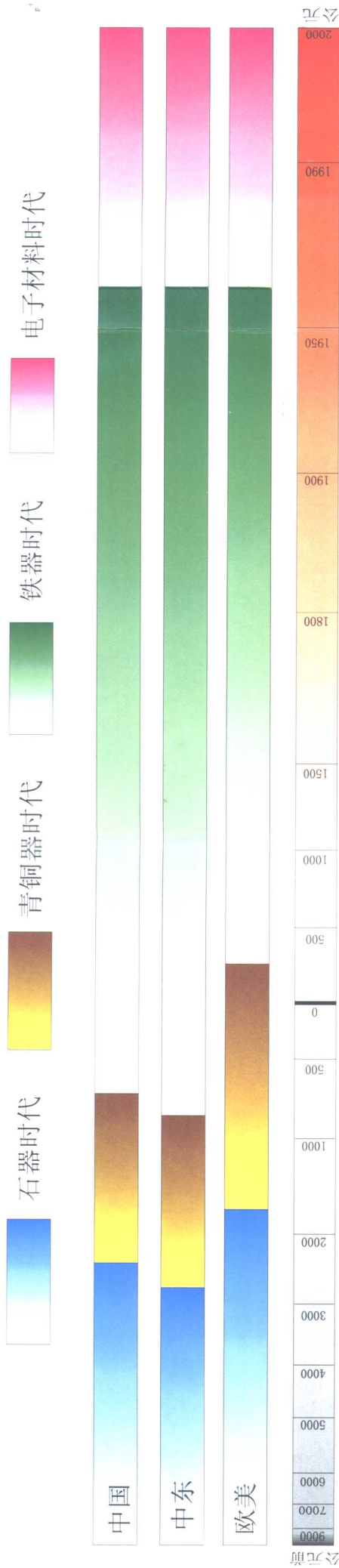
大量工作：特别是殷春荣、周岚、杨颖三位同学更为此付出了辛勤的劳动；吕鹏工程师也参加了绘图等工作；杨颖同学还绘制了彩图。编者对这些单位和个人无私帮助与深情厚谊表示衷心的感谢。

南京大学百年校庆在即，正是南京大学良好的学术环境为本书完成提供了重要条件。我们谨以此书作为微薄的礼物献给南京大学百年校庆。

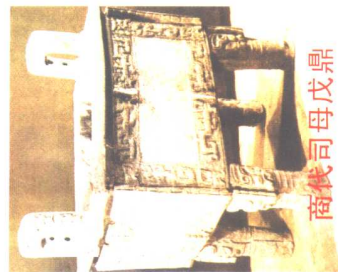
冯 端 师昌绪 刘治国

2002 年元月

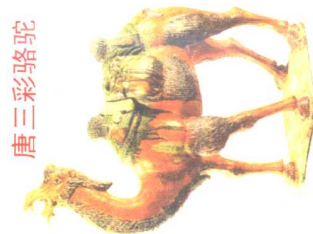
材料 与 文明



石铲



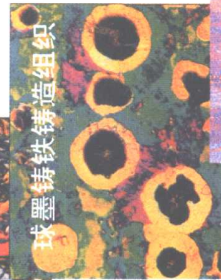
商代司母戊鼎



唐三彩骆驼



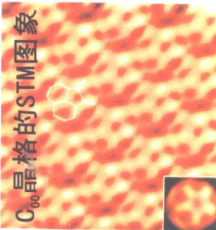
钢中的马氏体及其精细结构



球墨铸铁铸造组织



最早的晶体管



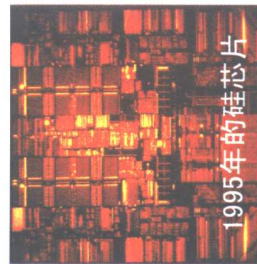
C₆₀晶格的STM图像



人面鱼纹陶盆



沧州铁狮子



1995年的硅芯片

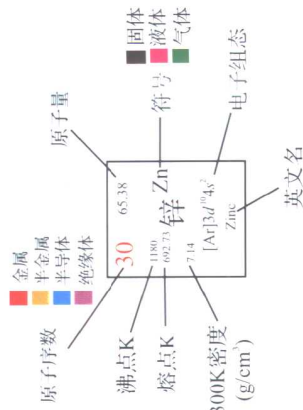


高温合金柱晶组织

元素周期表

GROUP
IA

1 20.268 1.0079 1.008 0.0899 Hydrogen H 1s ¹	2 4.00260 4.215 0.95 0.1787 Helium He 1s ²	3 6.941 6.941 1.01218 1.012 0.53 Lithium Li 1s ² 2s ¹	4 9.01218 9.012 1.363 1.156 0.97 Beryllium Be 1s ² 2s ²	5 10.81 10.81 12.011 12.01 2.34 Boron B 1s ² 2s ² 2p ¹	6 12.011 12.01 14.0067 14.01 2.62 Carbon C 1s ² 2s ² 2p ²	7 14.0067 14.01 15.9994 16.00 2.70 Nitrogen N 1s ² 2s ² 2p ³	8 15.9994 16.00 17.003 17.00 2.82 Oxygen O 1s ² 2s ² 2p ⁴	9 18.998403 18.998 20.179 20.18 2.34 Fluorine F 1s ² 2s ² 2p ⁵	10 20.179 20.18 21.017 21.02 2.96 Neon Ne 1s ² 2s ² 2p ⁶	11 22.98977 22.99 24.305 24.31 0.97 Sodium Na [Ne]3s ¹	12 24.305 24.31 26.98154 26.98 0.97 Magnesium Mg [Ne]3s ²	13 26.98154 26.98 28.0855 28.09 2.70 Aluminum Al [Ne]3s ² 3p ¹	14 28.0855 28.09 30.97376 30.97 2.43 Silicon Si [Ne]3s ² 3p ²	15 30.97376 30.97 32.06 32.07 2.33 Phosphorus P [Ne]3s ² 3p ³	16 32.06 32.07 33.9545 33.96 2.07 Sulfur S [Ne]3s ² 3p ⁴	17 35.453 35.46 36.96 36.97 3.17 Chlorine Cl [Ne]3s ² 3p ⁵	18 39.948 39.95 40.08 40.09 1.784 Argon Ar [Ne]3s ² 3p ⁶	19 39.0983 39.10 40.78 40.79 0.86 Potassium K [Ar]4s ¹	20 40.08 40.09 42.01 42.02 1.55 Calcium Ca [Ar]4s ²	21 44.9559 44.96 46.08 46.09 3.0 Scandium Sc [Ar]3d ¹ 4s ²	22 47.90 47.91 48.96 48.97 4.5 Titanium Ti [Ar]3d ² 4s ²	23 50.9415 50.94 51.996 52.00 5.8 Vanadium V [Ar]3d ³ 4s ²	24 51.996 52.00 54.938 54.94 7.19 Chromium Cr [Ar]3d ⁵ 4s ¹	25 54.938 54.94 55.847 55.85 7.43 Manganese Mn [Ar]3d ⁵ 4s ²	26 55.847 55.85 58.9332 58.93 7.86 Iron Fe [Ar]3d ⁶ 4s ²	27 58.9332 58.93 60.00 60.01 8.90 Cobalt Co [Ar]3d ⁷ 4s ²	28 58.93 58.94 63.546 63.55 8.90 Nickel Ni [Ar]3d ⁸ 4s ²	29 63.546 63.55 65.38 65.39 8.96 Copper Cu [Ar]3d ¹⁰ 4s ¹	30 65.38 65.39 69.72 69.73 9.14 Zinc Zn [Ar]3d ¹⁰ 4s ²	31 69.72 69.73 72.59 72.60 5.91 Gallium Ga [Ar]3d ¹⁰ 4s ¹ 4p ¹	32 72.59 72.60 74.9216 74.92 5.32 Germanium Ge [Ar]3d ¹⁰ 4s ² 4p ²	33 74.9216 74.92 76.90 76.91 5.72 Arsenic As [Ar]3d ¹⁰ 4s ² 4p ³	34 78.96 78.97 81.07 81.08 4.04 Selenium Se [Ar]3d ¹⁰ 4s ² 4p ⁴	35 79.904 79.91 83.80 83.81 3.12 Bromine Br [Ar]3d ¹⁰ 4s ² 4p ⁵	36 83.80 83.81 85.47 85.48 3.74 Krypton Kr [Ar]3d ¹⁰ 4s ² 4p ⁶	37 85.4678 85.47 87.62 87.63 1.53 Rubidium Rb [Kr]5s ¹	38 87.62 87.63 89.904 90.00 2.6 Strontium Sr [Kr]5s ²	39 89.904 90.00 91.22 91.23 4.5 Yttrium Y [Kr]4d ¹ 5s ²	40 91.22 91.23 92.9064 92.91 8.55 Zirconium Zr [Kr]4d ² 5s ²	41 92.9064 92.91 94.90 94.91 10.2 Niobium Nb [Kr]4d ⁴ 5s ¹	42 94.90 94.91 96.91 96.92 10.2 Molybdenum Mo [Kr]4d ⁵ 5s ¹	43 96.91 96.92 98.91 98.92 11.5 Technetium Tc [Kr]4d ⁵ 5s ²	44 98.91 98.92 101.07 101.08 12.2 Ruthenium Ru [Kr]4d ⁷ 5s ¹	45 101.07 101.08 102.9055 102.91 12.4 Rhodium Rh [Kr]4d ⁸ 5s ¹	46 102.9055 102.91 104.91 104.92 12.0 Palladium Pd [Kr]4d ¹⁰	47 104.91 104.92 106.91 106.92 10.5 Silver Ag [Kr]4d ¹⁰ 5s ¹	48 106.91 106.92 108.91 108.92 10.5 Cadmium Cd [Kr]4d ¹⁰ 5s ²	49 108.91 108.92 111.90 111.91 7.31 Indium In [Kr]4d ¹⁰ 5s ² 5p ¹	50 111.90 111.91 114.82 114.83 7.30 Tin Sn [Kr]4d ¹⁰ 5s ² 5p ²	51 114.82 114.83 116.91 116.92 6.8 Antimony Sb [Kr]4d ¹⁰ 5s ² 5p ³	52 116.91 116.92 118.91 118.92 6.24 Tellurium Te [Kr]4d ¹⁰ 5s ² 5p ⁴	53 126.9045 126.91 127.60 127.61 4.92 Iodine I [Kr]4d ¹⁰ 5s ² 5p ⁵	54 131.30 131.31 132.91 132.92 5.89 Xenon Xe [Kr]4d ¹⁰ 5s ² 5p ⁶	55 132.9054 132.91 137.33 137.34 3.5 Cesium Cs [Xe]6s ¹	56 137.33 137.34 138.85 138.86 3.5 Barium Ba [Xe]6s ²	57 138.9055 138.91 140.91 140.92 6.7 Lanthanum La [Xe]5d ¹ 6s ²	58 140.91 140.92 141.91 141.92 6.7 Cerium Ce [Xe]4f ¹ 5d ¹ 6s ²	59 140.9077 140.91 141.91 141.92 6.7 Praseodymium Pr [Xe]4f ³ 6s ²	60 141.91 141.92 143.91 143.92 7.0 Neodymium Nd [Xe]4f ⁴ 6s ²	61 143.91 143.92 144.24 144.25 6.475 Promethium Pm [Xe]4f ⁵ 6s ²	62 144.24 144.25 146.91 146.92 7.54 Samarium Sm [Xe]4f ⁶ 6s ²	63 150.4 150.41 151.96 151.97 5.36 Europium Eu [Xe]4f ⁷ 6s ²	64 151.96 151.97 153.91 153.92 7.89 Gadolinium Gd [Xe]4f ⁷ 5d ¹ 6s ²	65 158.9254 158.93 160.93 160.94 8.27 Terbium Tb [Xe]4f ⁹ 6s ²	66 160.93 160.94 162.50 162.51 8.54 Dysprosium Dy [Xe]4f ¹⁰ 6s ²	67 162.50 162.51 164.9304 164.93 8.80 Holmium Ho [Xe]4f ¹¹ 6s ²	68 164.9304 164.93 167.26 167.27 9.05 Erbium Er [Xe]4f ¹² 6s ²	69 167.26 167.27 168.9342 168.94 9.33 Thulium Tm [Xe]4f ¹³ 6s ²	70 173.04 173.05 175.04 175.05 9.84 Ytterbium Yb [Xe]4f ¹⁴ 6s ²	71 174.967 174.97 176.93 176.94 9.84 Lutetium Lu [Xe]4f ¹⁴ 5d ¹ 6s ²	72 176.93 176.94 178.93 178.94 9.84 Hafnium Hf [Xe]4f ¹⁴ 5d ² 6s ²	73 178.93 178.94 180.93 180.94 9.84 Tantalum Ta [Xe]4f ¹⁴ 5d ³ 6s ²	74 180.93 180.94 182.93 182.94 10.3 Tungsten W [Xe]4f ¹⁴ 5d ⁴ 6s ²	75 183.85 183.86 186.207 186.21 19.3 Rhenium Re [Xe]4f ¹⁴ 5d ⁵ 6s ²	76 186.207 186.21 188.91 188.92 22.4 Osmium Os [Xe]4f ¹⁴ 5d ⁶ 6s ²	77 192.22 192.23 194.22 194.23 22.5 Iridium Ir [Xe]4f ¹⁴ 5d ⁷ 6s ²	78 194.22 194.23 196.9665 196.97 21.4 Platinum Pt [Xe]4f ¹⁴ 5d ⁹ 6s ¹	79 196.9665 196.97 200.59 200.60 19.3 Gold Au [Xe]4f ¹⁴ 5d ¹⁰ 6s ¹	80 200.59 200.60 204.37 204.38 19.3 Mercury Hg [Xe]4f ¹⁴ 5d ¹⁰ 6s ²	81 204.37 204.38 207.2 207.21 11.4 Thallium Tl [Xe]4f ¹⁴ 5d ¹⁰ 6s ² 6p ¹	82 207.2 207.21 208.9804 208.98 9.4 Lead Pb [Xe]4f ¹⁴ 5d ¹⁰ 6s ² 6p ²	83 208.9804 208.98 210.5 210.6 9.8 Bismuth Bi [Xe]4f ¹⁴ 5d ¹⁰ 6s ² 6p ³	84 210.5 210.6 222.0 222.1 9.4 Polonium Po [Xe]4f ¹⁴ 5d ¹⁰ 6s ² 6p ⁴	85 210.5 210.6 223.0 223.1 9.4 Astatine At [Xe]4f ¹⁴ 5d ¹⁰ 6s ² 6p ⁵	86 222.0 222.1 226.0254 226.03 9.91 Radon Rn [Xe]4f ¹⁴ 5d ¹⁰ 6s ² 6p ⁶	87 223.0 223.01 226.0254 226.03 9.91 Francium Fr [Rn]7s ¹	88 226.0254 226.03 227.0278 227.03 10.07 Radium Ra [Rn]7s ²	89 227.0278 227.03 232.0381 232.04 10.07 Actinium Ac [Rn]7s ² 5f ¹	90 232.0381 232.04 238.0289 238.03 11.7 Thorium Th [Rn]7s ² 5f ²	91 232.0359 232.04 238.0289 238.03 15.4 Protactinium Pa [Rn]5f ² 6d ¹ 7s ²	92 238.0289 238.03 244.0642 244.07 15.4 Uranium U [Rn]5f ³ 6d ¹ 7s ²	93 238.0289 238.03 247.0 247.1 19.8 Neptunium Np [Rn]5f ⁴ 6d ¹ 7s ²	94 244.0642 244.07 250.1077 250.11 19.8 Plutonium Pu [Rn]5f ⁶ 7s ²	95 244.0642 244.07 252.0832 252.09 13.6 Americium Am [Rn]5f ⁷ 7s ²	96 252.0832 252.09 257.10 257.11 13.51 Curium Cm [Rn]5f ⁷ 6d ¹ 7s ²	97 257.10 257.11 261.10 261.11 13.51 Berkelium Bk [Rn]5f ⁹ 7s ²	98 257.10 257.11 262.10 262.11 13.51 Californium Cf [Rn]5f ¹⁰ 7s ²	99 261.10 261.11 267.10 267.11 13.51 Einsteinium Es [Rn]5f ¹¹ 7s ²	100 267.10 267.11 271.10 271.11 13.51 Fermium Fm [Rn]5f ¹² 7s ²	101 267.10 267.11 272.10 272.11 13.51 Mendelevium Md [Rn]5f ¹³ 7s ²	102 271.10 271.11 277.10 277.11 13.51 Nobelium No [Rn]5f ¹⁴ 7s ²	103 277.10 277.11 283.10 283.11 13.51 Lawrencium Lr [Rn]5f ¹⁴ 7s ² 5f ¹
--	--	---	---	---	--	---	--	---	---	---	--	--	---	---	--	--	--	---	--	--	--	--	---	--	--	---	--	---	--	---	---	---	--	--	---	---	--	---	--	--	---	---	--	--	---	--	---	--	---	---	---	---	---	--	--	---	--	--	---	--	---	--	---	--	--	---	--	---	---	---	---	--	---	--	---	---	--	---	--	--	---	---	--	--	--	--	--	--	--	---	---	--	--	--	--	---	--	--	---	---	--	--



注: 物态与导电态均指室温常压下的物相

目 录

综览.....	1
---------	---

第 1 篇 材料的结构

第 1 章 结构问题的不同侧面	17
1.1 键合	17
1.1.1 原子结构和周期表	17
1.1.2 离子键	20
1.1.3 共价键	23
1.1.4 分子	25
1.1.5 金属键	27
1.1.6 弱键	28
1.2 对称性	29
1.2.1 对称操作	29
1.2.2 对称与不对称	30
1.2.3 对称元素的组合	31
1.3 对称破缺——无序与有序	34
1.3.1 概念的引入	34
1.3.2 从概念到现实——有序-无序转变的实例	35
1.3.3 能与熵的角逐——有序-无序转变的物理根源	38
1.3.4 实际结构——有序相中的缺陷	40
1.4 结构与信息	41
1.4.1 信息	41
1.4.2 遗传信息的载体——DNA 分子	42
1.4.3 蛋白质分子	44
参考文献	46
第 2 章 晶态	47
2.1 周期结构和点阵	47
2.1.1 点阵	48
2.1.2 点阵与点群的类型	49
2.1.3 点阵的几何关系	53

2.1.4 对称性对物性常量的制约	54
2.2 空间群	55
2.2.1 晶体的微观对称性	55
2.2.2 空间群	56
2.2.3 空间群的国际符号	57
2.2.4 晶体结构的描述	58
2.3 若干晶体结构的实例	60
2.3.1 堆积结构	60
2.3.2 键联结构	62
2.3.3 钙钛矿结构及其家族	65
2.4 超出空间群的结构	67
2.4.1 色群和磁结构	67
2.4.2 无公度调制结构	67
2.4.3 准晶和 Penrose 拼砌	68
参考文献	70
第3章 非晶态与液晶态	71
3.1 导向非晶态	71
3.1.1 熔化	71
3.1.2 玻璃化转变	73
3.1.3 位置无序的统计描述	75
3.2 无机玻璃	78
3.2.1 无规密堆模型	78
3.2.2 无规网络模型	79
3.2.3 硫系玻璃	81
3.3 聚合物	82
3.3.1 聚合物的结构和构型	82
3.3.2 无规行走的线团模型	84
3.3.3 溶胀的线团	85
3.3.4 交联	87
3.4 液晶态	88
3.4.1 概述	88
3.4.2 向列相和胆甾相	91
3.4.3 近晶相和柱状相	91
3.4.4 溶致液晶 (自组装膜的有序结构)	92
3.4.5 聚合物的有序结构	94

参考文献	96
第4章 点缺陷	97
4.1 空位与填隙原子	97
4.1.1 热平衡态的点缺陷浓度	97
4.1.2 点缺陷的形成能	98
4.1.3 点缺陷的迁移激活能	99
4.2 点缺陷的产生	100
4.3 离子晶体中的点缺陷与色心	103
4.3.1 离子晶体中的点缺陷	103
4.3.2 离子导电和超离子导电现象	104
4.3.3 色心	105
4.4 杂质原子	108
参考文献	110
第5章 位错与向错	111
5.1 晶体中位错的几何特征	112
5.2 位错的弹性性质	116
5.3 位错核心结构	118
5.4 位错与其他缺陷之间的交互作用	121
5.4.1 位错间的交互作用	122
5.4.2 位错与溶质原子间的交互作用	123
5.4.3 位错与自由表面的交互作用	124
5.5 位错的产生和增殖	126
5.5.1 位错的产生	126
5.5.2 位错的增殖	128
5.6 向错	131
参考文献	134
第6章 表面与界面	135
6.1 晶体的表面	135
6.1.1 晶体外形与表面微观形貌	135
6.1.2 表面弛豫与重构	138
6.1.3 技术晶体的表面	141
6.1.4 纳米粒子的表面	142
6.2 平移界面	143
6.2.1 堆垛层错	143
6.2.2 反相畴界	147

6.2.3 晶体学切变面	150
6.3 孪晶界面	152
6.3.1 反映孪晶界面	152
6.3.2 旋转孪晶界面	155
6.3.3 铁电畴界	155
6.4 晶界	158
6.4.1 小角度晶界	158
6.4.2 大角度晶界	162
6.5 相界	164
参考文献	166
第7章 多层次与非均质材料	168
7.1 非均质材料	168
7.1.1 概述	168
7.1.2 非均质材料的微结构特征	169
7.1.3 有效媒质近似——二相合金微结构与物性 关联问题处理实例	173
7.1.4 非均质介观结构实例	174
7.2 结构的几何相变：逾渗	183
7.2.1 逾渗的概念	183
7.2.2 逾渗的若干实例	185
7.3 无规结构：分形几何	186
7.3.1 支离破碎世界的描述	186
7.3.2 物质结构中的分形	188
参考文献	190

第2篇 材料的物性

第8章 理解物性的基本概念	191
8.1 量子特征	191
8.1.1 波粒二象性	191
8.1.2 经典统计与量子统计	193
8.2 固态电子理论	198
8.2.1 金属自由电子理论	198
8.2.2 能带概念的引入	199
8.2.3 布里渊区与能态密度	202
8.3 晶格振动理论	205

8.3.1 简谐近似	205
8.3.2 爱因斯坦模型与德拜模型	206
8.3.3 格波与晶格振动模	208
8.4 相互作用的电子体系	210
8.4.1 准粒子体系	210
8.4.2 强磁性与超导电性	211
8.4.3 强关联电子体系	214
参考文献	216
第9章 输运性质	217
9.1 金属与合金的输运性质	217
9.1.1 经典电导理论	218
9.1.2 量子电导理论简介	219
9.1.3 电导的实验研究	221
9.1.4 热电效应	223
9.2 半导体的输运性质	224
9.2.1 半导体的能带	225
9.2.2 载流子和费米能级	226
9.2.3 迁移率和散射	230
9.2.4 霍尔效应	231
9.2.5 p-n 结	233
9.2.6 金属与半导体接触	235
9.2.7 光电导与光伏效应 (太阳能电池)	238
9.2.8 晶体管	239
9.2.9 量子半导体器件	242
9.3 聚合物、非晶态材料等的输运性质	244
9.3.1 导电聚合物	244
9.3.2 非晶态半导体	248
9.3.3 离子晶体	251
参考文献	253
第10章 磁学性质	254
10.1 材料磁性来源与分类	255
10.1.1 原子磁矩	255
10.1.2 材料磁性概述和分类	258
10.2 磁有序与交换作用	264
10.2.1 铁磁有序	264

10.2.2	反铁磁和亚铁磁性	268
10.2.3	交换作用	272
10.2.4	强磁材料的本征性能	277
10.3	强磁材料的磁畴与磁化	280
10.3.1	影响磁畴结构的能量	280
10.3.2	强磁体中的磁畴结构	283
10.3.3	磁化曲线与磁滞回线	285
10.4	磁化动力学与铁磁共振	290
10.4.1	磁化进动及转动	290
10.4.2	张量磁化率和铁磁共振	291
10.4.3	自旋波共振	292
10.4.4	磁化与反磁化的动态过程	293
10.5	强磁材料及应用	295
10.5.1	强磁材料按组成与结构的分类	295
10.5.2	强磁材料的应用	296
10.6	巨磁电阻效应和磁电子学	303
10.6.1	磁电阻和巨磁电阻	303
10.6.2	自旋相关导电	304
10.6.3	磁电子学和巨磁电阻的应用	308
	参考文献	310
第 11 章	铁电性质	311
11.1	电介质的极化与铁电性	311
11.1.1	电介质的极化	311
11.1.2	自发极化与铁电性	314
11.2	压电性与热释电性	323
11.2.1	压电性与压电材料	323
11.2.2	热释电性与热释电材料	330
	参考文献	332
第 12 章	超导电性质	333
12.1	引言	333
12.2	超导电性的基本特征	334
12.2.1	零电阻效应	334
12.2.2	迈斯纳效应	336
12.2.3	超导体临界参数	337
12.3	超导体的热力学性质	338

12.3.1	自由能	338
12.3.2	熵	340
12.3.3	潜热	340
12.3.4	比热容	340
12.3.5	超导能隙	341
12.4	迈斯纳态的电磁性质	341
12.4.1	二流体模型	341
12.4.2	London 理论	342
12.4.3	G-L 理论	344
12.4.4	相干长度	345
12.4.5	表面能	346
12.4.6	磁通量子化	347
12.5	第二类超导体	348
12.6	超导隧道效应	350
12.7	超导电性的微观图像	352
12.7.1	实验和理论的启示	353
12.7.2	电子-声子机制	355
12.7.3	Cooper 对	355
12.7.4	BCS 理论图像	356
12.8	超导材料	357
12.8.1	强磁场超导材料	357
12.8.2	超导电子材料	358
12.8.3	高温超导材料	359
	参考文献	362
第 13 章	光学性质	363
13.1	光波在线性介质中的传播	363
13.1.1	光波在均匀的各向同性电介质中的传播	364
13.1.2	光波在导体中的传播	365
13.1.3	光波在各向异性介质中的传播	365
13.1.4	光波在平面介质薄膜波导中的传播	370
13.1.5	光波在光纤中的传播	372
13.1.6	光波在光子晶体中的传播	374
13.2	光发射	376
13.2.1	光的自发辐射和受激跃迁	377
13.2.2	激光产生的原理	379

13.2.3	典型的三能级和四能级激光系统·····	381
13.2.4	Nd:YVO ₄ 和半导体激光器 ·····	382
13.2.5	一种新的激光发射机理与材料——纳米硅发光·····	386
13.3	激光频率变换·····	387
13.3.1	非线性极化·····	387
13.3.2	块状晶体中的光频变换·····	389
13.3.3	光学超晶格中的光频变换·····	393
13.4	非线性折射率与非线性吸收·····	397
13.4.1	非线性折射率·····	397
13.4.2	非线性吸收·····	403
	参考文献·····	407
第 14 章	力学性质 ·····	409
14.1	弹性与广义弹性·····	410
14.1.1	弹性参量·····	410
14.1.2	常规弹性的物理本质·····	413
14.1.3	高弹性的物理本质·····	415
14.1.4	粘弹性·····	417
14.2	塑性形变与粘性流变·····	419
14.2.1	晶体的理论屈服强度·····	419
14.2.2	晶体塑性形变的几何学与结晶学关系·····	423
14.2.3	塑性流变——从低温到高温·····	426
14.2.4	高聚物粘性流变·····	430
14.3	断裂·····	437
14.3.1	理论断裂强度·····	437
14.3.2	格里菲斯裂纹理论·····	438
14.3.3	脆性与韧性·····	441
14.4	材料的强化和增韧·····	442
14.4.1	金属材料的强化·····	442
14.4.2	陶瓷材料的增韧·····	449
14.4.3	复合材料·····	453
	参考文献·····	455

第 3 篇 材料制备的科学基础

第 15 章	材料制备的概述 ·····	457
15.1	材料合成与加工的意义和内涵·····	457

15.2 基于液相-固相转变的材料制备	458
15.2.1 从熔体制备单晶材料	458
15.2.2 从熔体制备非晶材料	459
15.2.3 溶液法材料制备	460
15.2.4 溶胶-凝胶法	461
15.3 基于固相-固相转变的材料制备	463
15.3.1 固相反应法制备粉末	463
15.3.2 陶瓷成型和烧结	464
15.3.3 固相外延	465
15.3.4 高压制备	465
15.4 基于气相-固相转变的材料制备	466
15.4.1 真空蒸发镀膜	466
15.4.2 溅射和激光脉冲沉积	467
15.4.3 化学气相沉积	469
15.4.4 分子束外延	470
参考文献	475
第 16 章 相图原理	476
16.1 吉布斯相律	477
16.2 杠杆定律	478
16.3 单元体系的温度~压力图 (p - T 图)	479
16.4 二元相图	480
16.4.1 匀晶相图	480
16.4.2 共晶相图与包晶相图	482
16.4.3 具有中间化合物的相图	484
16.4.4 偏晶相图与综晶相图	486
16.5 三元相图	487
16.5.1 三元相图的表示法	487
16.5.2 三元相图的分类	493
参考文献	499
第 17 章 固体中的扩散、化学反应与烧结	502
17.1 固体中的扩散	502
17.1.1 扩散的基本特点与唯象理论	502
17.1.2 质点迁移的微观机构与扩散系数	509
17.1.3 扩散的热力学理论	514
17.1.4 短路扩散与非平衡态下点缺陷的扩散	517

17.1.5	带电质点的扩散行为·····	520
17.1.6	金属中的电迁移和热迁移·····	521
17.2	固相化学反应·····	522
17.2.1	固相反应动力学的基本特征·····	522
17.2.2	固相反应动力学方程·····	523
17.2.3	影响固相反应的因素·····	529
17.3	固态烧结·····	532
17.3.1	固态烧结的初期动力学·····	533
17.3.2	固态烧结中、后期动力学模型·····	536
17.3.3	烧结过程中的晶粒生长与二次再结晶·····	537
17.3.4	液相烧结和热压烧结·····	539
	参考文献·····	540
第 18 章	相变的基本原理 ·····	541
18.1	相变的基本结构特征·····	542
18.1.1	重构型相变和位移型相变·····	542
18.1.2	马氏体型相变·····	544
18.1.3	有序-无序相变 ·····	545
18.1.4	其他形式相变·····	546
18.2	相变热力学·····	548
18.2.1	重要的热力学函数·····	548
18.2.2	一级相变与高级相变·····	549
18.2.3	朗道相变理论简介·····	552
18.2.4	朗道理论的推广·····	554
18.3	固态相变动力学·····	555
18.3.1	新相胚核形成过程·····	555
18.3.2	胚核生长和粗化过程·····	562
18.3.3	相变动力学速率形式理论·····	568
18.3.4	成核-生长和失稳分解过程 ·····	575
18.3.5	失稳分解的动力学方程·····	579
18.4	结语·····	584
	参考文献·····	586
第 19 章	薄膜的外延生长 ·····	587
19.1	外延的基本物理过程·····	587
19.1.1	基本概念·····	587
19.1.2	表面成核·····	589

19.1.3	表面动力学	592
19.1.4	外延界面的结构	593
19.2	同质外延	595
19.2.1	同质外延的生长模式	595
19.2.2	杂质分凝与控制	596
19.2.3	Si 和 GaAs 的同质外延	596
19.3	异质外延	597
19.3.1	异质外延系统的结构与分类	597
19.3.2	应变层的稳定性	600
19.3.3	异质外延的生长模式	601
	参考文献	602
第 20 章	生长界面的稳定性——枝晶生长和分形生长	603
20.1	引言	603
20.2	生长界面的稳定条件	605
20.2.1	温度梯度对界面稳定性的影响	605
20.2.2	浓度梯度对界面稳定性的影响	606
20.2.3	界面能对界面稳定性的影响	608
20.2.4	平界面的失稳条件	608
20.3	枝晶生长	611
20.3.1	控制枝晶生长的几个物理因素	612
20.3.2	枝晶主干的生长——针状晶体的生长	614
20.4	分形的概念和分形聚集	618
20.4.1	分形形态的表征	618
20.4.2	由扩散控制的分形聚集过程	620
20.4.3	界面能对聚集形态的影响	622
	参考文献	623
第 21 章	软物质的自组织	624
21.1	引言	624
21.1.1	什么是软物质	624
21.1.2	软物质怎样表现	627
21.2	自组织的原理和控制	628
21.2.1	自组织趋向有序途径	628
21.2.2	软物质自组织的形式	630
21.2.3	控制和设计自组织形貌的途径	639
21.3	软物质的复相分离	641

21.3.1 软物质复相分离的热力学·····	641
21.3.2 软物质复相分离的动力学·····	648
21.4 结语：自组装的基本问题和研究前沿·····	650
参考文献·····	653

第4篇 展 望

第22章 材料表征及其进展 ·····	654
22.1 性能检测·····	654
22.2 显微组织分析原理·····	655
22.2.1 电子与样品相互作用·····	655
22.2.2 X射线与物质的相互作用·····	657
22.2.3 离子束分析·····	658
22.2.4 中子衍射与小角度中子散射·····	659
22.3 显微结构表征·····	660
22.3.1 形貌观察·····	660
22.3.2 结构测定·····	660
22.3.3 化学组分分析·····	662
22.4 从材料出发的综合分析·····	664
22.5 材料性能与组织表征技术发展展望·····	665
22.5.1 分析技术的交叉与综合·····	665
22.5.2 分析结果的定量化与可视化·····	666
22.5.3 低维材料制备与分析测试结合·····	667
参考文献·····	668
第23章 材料设计及其进展 ·····	669
23.1 什么是材料设计·····	669
23.2 材料设计的发展概况·····	670
23.2.1 前期研究的回顾·····	670
23.2.2 当前面临的挑战·····	671
23.3 材料设计的范围与层次·····	672
23.4 材料设计的主要途径·····	674
23.4.1 材料知识库和数据库技术·····	674
23.4.2 材料设计专家系统·····	676
23.4.3 材料设计中的计算机模拟·····	677
23.4.4 基于第一性原理的计算·····	679
23.5 基于第一性原理的主要计算方法·····	680

23.5.1	密度泛函理论	681
23.5.2	准粒子方程, GW 近似	682
23.5.3	Car-Parrinello 方法	684
23.6	国外研究动态与展望	685
23.6.1	1995 年美国 NRC 的专门报告	685
23.6.2	美国能源部的 SSI 计划	690
23.6.3	美国西北大学 SRG 的工作	693
23.7	我国材料设计研究的进展概况	694
	参考文献	695
第 24 章	计算相图及其进展	697
24.1	相图计算的基本原理——Gibbs 法则	699
24.2	纯物质的自由能——点阵稳定性常数	700
24.2.1	温度关系模型	700
24.2.2	压力关系模型	701
24.2.3	纯组元相图的计算	701
24.3	热力学模型	702
24.3.1	二元化学计量比相的热力学模型	702
24.3.2	替换溶液模型	702
24.3.3	准化学模型	705
24.3.4	缔合物模型	705
24.3.5	团簇变分法 (CVM 法)	707
24.3.6	亚点阵模型	707
24.4	相图优化和计算过程	716
24.5	材料优化设计实例——双相不锈钢的成分设计	716
	参考文献	721
第 25 章	材料前沿的若干问题	723
25.1	科学、技术与材料	723
25.1.1	技术对材料研制的促进	723
25.1.2	当代技术的挑战与机遇	724
25.2	新材料发展的主要动向	727
25.2.1	开拓超微结构的领域	727
25.2.2	探测电子关联体系中的宝藏	732
25.2.3	开拓有机材料的新领域	735
25.3	结语	740
	参考文献	741

综 览

撰稿人 师昌绪

材料是人类用于制造物品、器件、构件、机器或其他产品的那些物质。材料是物质，但不是所有物质都可以称为材料。如燃料和化学原料、工业化学品、食物和药物，一般都不算是材料。但是这个定义并不那么严格，如炸药、固体火箭推进剂，一般称之为“含能材料”，因为它属于火炮或火箭的组成部分。

材料是人类赖以生存和发展的物质基础。20 世纪 70 年代人们把信息、材料和能源誉为当代文明的三大支柱。80 年代以高技术群为代表的新技术革命，又把新材料、信息技术和生物技术并列为新技术革命的重要标志。这主要是因为材料与国民经济建设、国防建设和人民生活密切相关。

材料除了具有重要性和普遍性以外，还具有多样性。由于多种多样，分类方法也就没有一个统一标准。从物理化学属性来分，可分为金属材料、无机非金属材料、有机高分子材料和不同类型材料所组成的复合材料。从用途来分，又分为电子材料、航空航天材料、核材料、建筑材料、能源材料、生物材料等。更常见的两种分类方法则是结构材料与功能材料；传统材料与新型材料。结构材料是以力学性能为基础，以制造受力构件所用材料，当然，结构材料对物理或化学性能也有一定要求，如光泽、热导率、抗辐照、抗腐蚀、抗氧化等。功能材料则主要是利用物质的独特物理、化学性质或生物功能等而形成的一类材料。一种材料往往既是结构材料又是功能材料，如铁、铜、铝等。传统材料是指那些已经成熟且在工业中已批量生产并大量应用的材料，如钢铁、水泥、塑料等。这类材料由于其量大、产值高、涉及面广，又是很多支柱产业的基础，所以又称为基础材料。新型材料（先进材料）是指那些正在发展，且具有优异性能和应用前景的一类材料。新型材料与传统材料之间并没有明显的界限，传统材料通过采用新技术，提高技术含量，提高性能，大幅度增加附加值而成为新型材料；新材料在经过长期生产与应用之后也就成为传统材料。传统材料是发展新材料和高技术的基础，而新型材料又往往能推动传统材料的进一步发展。

材料是人类社会进步的里程碑

人类发展的历史证明，材料是社会进步的物质基础，是人类进步程度的

主要标志，所以人类社会的进步以材料作为里程碑。纵观人类发现材料和利用材料的历史，每一种重要材料的发现和广泛利用，都会把人类支配和改造自然的能力提高到一个新水平，给社会生产力和人类生活水平带来巨大的变化，把人类的物质文明和精神文明向前推进一步。

早在一百万年以前，人类开始用石头做工具，使人类进入旧石器时代。大约一万年以前，人类知道对石头进行加工，使之成为精致的器皿或工具，从而使人类进入新石器时代。在新石器时代，人类开始用皮毛遮身。8000年前中国就开始用蚕丝做衣服，4500年前印度人开始种植棉花，这些都标志着人类使用材料促进文明进步。在新石器时代，人类已知道使用自然铜和天然金，但毕竟数量太少，分散细小，没有对人类社会产生重要影响。

大约在8000~9000年前，人类还处于新石器时代，已发明了用粘土成型，再火烧固化而成为陶器。陶器的出现，不但用于器皿，而且成为装饰品，是对精神文明的一大促进，历史上虽无陶器时代的名称，但其对人类文明的贡献是不可估量的。在烧制陶器过程中，偶然发现金属铜和锡，当然那时还不明白这是铜、锡的氧化物在高温下被炭还原的产物，进而又生产出色泽鲜艳、又能浇铸成型的青铜，从而使人类进入青铜时代。这是人类较大量利用金属的开始，也是人类文明发展的重要里程碑。世界各地开始青铜时代的时间各不相同，希腊在公元前3000年前，埃及在公元前2500年前，巴比伦在公元前19世纪中叶，印度大约在公元前3000年已广泛使用青铜器。中国的青铜器在公元前2700年已经使用了，至今约5000年的历史，到商周（公元前17世纪到公元前3世纪）进入了鼎盛时期，如河南安阳出土的达875kg的鼎、湖北随县的编钟、西安青铜车马都充分反映了当时中国冶金技术水平和制造工艺的高超。

公元前13~14世纪前，人类已开始用铁，3000年前铁工具比青铜工具更为普遍，人类开始进入了铁器时代。中国最早出土的人工冶铁制品约在公元前9世纪。到春秋（公元前770~476年）末期，生铁技术有较大突破，遥遥领先于世界其他地区，如用生铁退火而制成韧性铸铁及以生铁炼钢技术的发明，促进了中国生产力的大发展，对战国和秦汉农业、水利和军事的发展起到很大作用。早在公元2世纪中国的钢和丝绸已驰名罗马帝国。生铁技术在公元前5世纪即春秋末叶已经在黄河长江流域传播。这些技术于公元6~7世纪传入朝鲜半岛、日本和北欧，推动了整个世界文明的进步。

随着世界文明的进步，18世纪发明了蒸汽机，19世纪发明了电动机，对金属材料提出了更高要求，同时对钢铁冶金技术产生了更大的推动作用。1854年和1864年先后发明了转炉和平炉炼钢，使世界钢产量有一个飞速发展。如1850年世界钢产量为6万吨，1890年达2800万吨，大大促进了机

械制造、铁道交通及纺织工业的发展。随之电炉冶炼开始，不同类型的特殊钢相继问世，如 1887 年高锰钢、1900 年 18-4-1 (W18Cr4V) 高速钢、1903 年硅钢及 1910 年奥氏体镍铬 (Cr18Ni8) 不锈钢，把人类带进了现代物质文明。在此前后，铜、铝也得到大量应用，而后镁、钛和很多稀有金属都相继出现，从而金属材料在整个 20 世纪占据了结构材料的主导地位。

随着有机化学的发展，19 世纪末叶西方科学家仿制中国丝绸发明了人造丝，这是人类改造自然材料的又一里程碑。20 世纪初，人工合成有机高分子材料相继问世，如 1909 年的酚醛树脂 (电木)，1920 年的聚苯乙烯，1931 年的聚氯乙烯及 1941 年的尼龙等，以其性能优异，及资源丰富、建设投资少、收效快而得到迅速发展。目前世界三大有机合成材料 (树脂、纤维和橡胶) 年产量逾亿吨。而且有机材料的性能不断提高，附加值大幅度增加，特别是特种聚合物正向功能材料各个领域进军，显示其巨大的潜力。

陶瓷本来用作建筑材料、容器或装饰品等。由于其资源丰富、密度小、高模量、高硬度、耐腐蚀、膨胀系数小、耐高温、耐磨等特点，到了 20 世纪中叶，通过合成及其他制备方法，做出各种类型的先进陶瓷 (如 Si_3N_4 、 SiC 、 ZrO_2 等)，成为近几十年来材料中非常活跃的研究领域，有人甚至认为“新陶器时代”即将到来。但由于其脆性问题难以解决，且价格过高，作为结构材料没有得到如钢铁或高分子材料一样的广泛应用。

复合材料是 20 世纪后期发展的另一类材料。众所周知，天然材料很多是复合材料，如木材、皮革、竹子等。事实上，人类很早就制造复合材料，如泥巴中混入碎麻或麦秆用以建造房屋，钢筋水泥是脆性材料和韧性材料的复合。近几十年来，利用树脂的易成型和金属韧性好，无机非金属的高模量、高强度、耐高温，做成了树脂基复合材料或金属基复合材料，前者已得到广泛应用，后者以其制作困难、价格高而受到一定限制。为了改善陶瓷的性能，也制成陶瓷基复合材料。碳是使用温度最高的材料 (可达 2500°C)，为了克服热震性能差，并提高其力学性能而制出的碳-碳复合材料已广泛用于军工，并扩展到民用。

以上仅以结构材料为主线，概述了材料的发展历史。可以看出，自 19 世纪中叶现代炼钢技术出现以后，金属材料的重要性急剧增加，一直到 20 世纪中叶，人工合成有机材料、陶瓷材料及先进复合材料迅速发展，金属材料的重要性逐渐下降，但一直到 21 世纪上半叶，金属材料仍将占重要位置。

功能材料自古就受到重视，早在战国 (公元前 3 世纪) 已利用天然磁铁矿来制造司南，到宋代用钢针磁化制出了罗盘，为航海的发展提供了关键技术。功能材料是信息技术及自动化的基础，特别是半导体材料出现以后，加速了现代文明的发展，1947 年发明了第一只具有放大作用的晶体管，10 余

年后又研制成功集成电路，使以硅材料为主体的计算机的功能不断提高，体积不断缩小，价格不断下降，加之高性能的磁性材料不断涌现，激光材料与光导纤维的问世，使人类社会进入了信息时代。因为硅是微电子技术的关键材料，所以有人称之为“硅片为代表的电子材料时代”，再一次说明材料对人类文明起了关键的作用。

材料科学的形成与内涵

“材料”是早已存在的名词，但“材料科学”的提出只是 20 世纪 60 年代初的事。1957 年前苏联人造卫星首先上天，美国朝野上下为之震惊，认为自己落后的主要原因之一是先进材料落后，于是在一些大学相继成立了十余个材料研究中心。采用先进的科学理论与实验方法对材料进行深入的研究，取得重要成果。从此，“材料科学”这个名词便开始流行。

“材料科学”的形成实际是科学技术发展的结果。

首先，固体物理、无机化学、有机化学、物理化学等学科的发展，对物质结构和物性的深入研究，推动了对材料本质的了解；同时，冶金学、金属学、陶瓷学、高分子科学等的发展也使对材料本身的研究大大加强，从而对材料的制备、结构与性能，以及它们之间的相互关系的研究也愈来愈深入，为材料科学的形成打下了比较坚实的基础。

其次，在材料科学这个名词出现以前，金属材料、高分子材料与陶瓷材料都已自成体系，目前复合材料也获得广泛应用，其研究也逐步深入。但它们之间存在着颇多相似之处，对不同类型材料的研究可以相互借鉴，从而促进学科的发展。如马氏体相变本来是金属学家提出来的，而且广泛地被用来作为钢热处理的理论基础；但在氧化锆陶瓷中也发现了马氏体相变现象，并用来作为陶瓷增韧的一种有效手段。又如材料制备方法中的溶胶-凝胶法，是利用金属有机化合物的分解而得到纳米级高纯氧化物粒子，成为改进陶瓷性能的有效途径。虽然不同类型的材料各有其专用测试设备与生产装置，但各类材料的研究检测设备与生产手段有颇多共同之处。例如显微镜、电子显微镜、表面测试及物性与力学性能测试设备等。在材料生产中，许多加工装置的原理也有颇多相通之处，可以相互借鉴，从而加速材料的发展。

第三，许多不同类型的材料可以相互替代和补充，能更充分发挥各种材料的优越性，达到物尽其用的目的。但长期以来，金属、高分子及无机非金属材料自成体系，缺乏沟通。由于互不了解，不利于发展创新，对复合材料的发展也极不利。

尽管从材料发展需要和共性来看，有必要形成一门材料科学，但是由于各类材料的学科基础不同，还存在不小的分歧，特别是无机材料与有机材料

之间分歧较大；但由于软物质科学的崛起，还可以软硬兼顾，互相借鉴。

材料科学所包括的内容往往被理解为研究材料的组织、结构与性质的关系，探索自然规律。这属于基础研究。实际上，材料是面向实际、为经济建设服务的，是一门应用科学，研究与发展材料的目的在于应用，而材料又必须通过合理的工艺流程才能制备出具有实用价值的材料来，通过批量生产才能成为工程材料。所以，在“材料科学”这个名词出现后不久，就提出了“材料科学与工程”。工程是指研究材料在制备过程中的工艺和工程技术问题。许多大学的冶金系、材料系也就此改变了名称，多数改为“材料科学与工程系”，偏重基础方面的就称“材料科学系”，偏重工艺方面的称“材料工程系”。也有不肯放弃“冶金”而称为“冶金与材料科学系”的，如英国的剑桥大学。同时，有关材料科学或材料科学与工程方面的杂志和书籍也应运而生。第一部《材料科学与工程百科全书》由美国麻省理工学院的科学家主编，由英国 Pergamon 自 1986 年陆续出版。它对材料科学与工程下的定义为：材料科学与工程就是研究有关材料组成、结构、制备工艺流程与材料性能和用途的关系的知识的产生及其运用。换言之，材料科学与工程是研究材料组成、结构、生产过程、材料性能与使用效能以及它们之间的关系。因而把组成与结构 (composition-structure)、合成与生产过程 (synthesis-processing)、性质 (properties) 及使用效能 (performance) 称之为材料科学与工程的四个基本要素 (basic elements)。把四要素连结在一起，便形成一个四面体 (tetrahedron)，如图 1(a)。

考虑在四要素中的组成/结构并非同义词，即相同成分或组成通过不同的合成或加工方法，可以得出不同结构，从而材料的性质或使用效能都不会相同。因此，我国有人提出一个五个基本要素的模型，即成分 (composition)、合成/加工 (synthesis/processing)、结构 (structure)、性质 (properties) 和使用效能 (performance)。如果把它们连接起来，则形成一个六面体 (hexahedron)，如图 1(b)。

材料科学与工程五要素模型的特点主要有两个：一是性质与使用效能有一个特殊的联系，材料的使用效能便是材料性质在使用条件下的表现。环境对材料性能的影响很大，如受力状态、气氛、介质与温度等。有些材料在一般环境下的性能很好，而在腐蚀介质下性能却下降显著；有的材料在光滑样品时表现很好，而在有缺口的情况下性能大为下降，特别有些高强度材料表现尤为突出，但凡有一个划痕，就会造成灾害性破坏。因此，环境因素的引入对工程材料来说十分重要。二是材料理论和材料设计或工艺设计有了一个适当位置，它处在六面体的中心。因为这五个要素中的每一个要素，或几个相关要素都有其理论，根据理论建立模型 (modelling)，通过模型可以进行

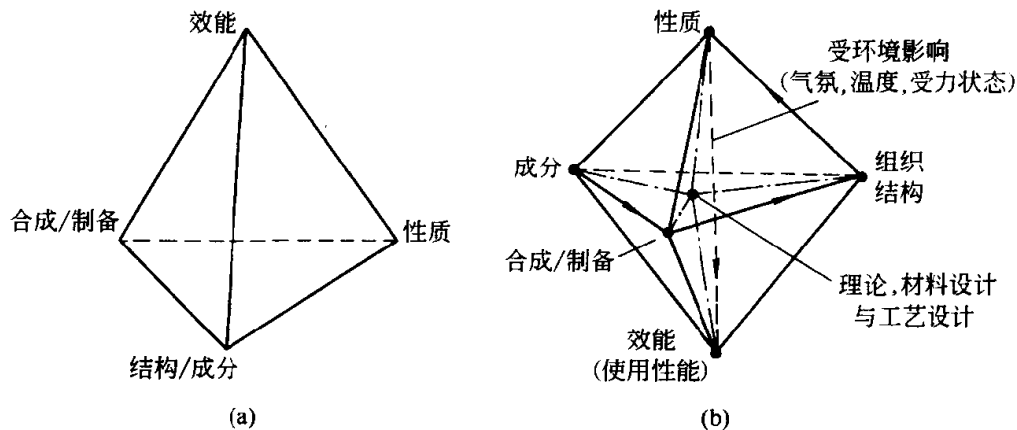


图 1 材料科学与工程要素图

(a) 材料科学与工程四要素；(b) 材料科学与工程五要素

材料设计或工艺设计，以达到提高性能及使用效能、节约资源、减少污染或降低成本的最佳状态。这是材料科学与工程最终努力的目标。有人设想提出性能指标或使用效能要求，可以通过材料配方，采用最佳工艺，就可制备出符合要求的材料或器件。应该说明，目前国际流行的仍是四要素模型，五要素模型在国际上也有人引用。

根据以上所述，材料科学有三个重要属性：一是多学科交叉，它是物理学、化学、冶金学、金属学、陶瓷、高分子化学及计算科学相互融合与交叉的结果，如生物医用材料要涉及医学、生物学及现代分子生物学等学科；二是一种与实际使用结合非常密切的科学，发展材料科学的目的在于开发新材料，提高材料的性能和质量，合理使用材料，同时降低材料成本和减少污染等；三是材料科学是一个正在发展中的科学，不像物理学、化学已经有一个很成熟的体系，材料科学将随各有关学科的发展而得到充实和完善。

先进材料是社会现代化的先导

材料既是人类社会进步的里程碑，又是社会现代化的物质基础与先导。特别是先进材料的研究、开发与应用反映着一个国家科学技术与工业水平。现仅就现代科学技术的发展与先进材料的关系举几个典型事例来说明。

电子技术的发展

从电子技术的发展可以看出材料所起的作用。1906 年发明了电子管，从而出现了无线电技术、电视机、电子计算机。1948 年发明了半导体晶体管，致使电子设备的小型化、轻量化、节能化及成本的降低、可靠性的提高与寿命的延长。1958 年出现了集成电路，使计算机及各种电工设备发生再一次飞跃。如以 1946 年电子管计算机与 1976 年微机的一些指标来对比，由

于集成电路的采用，使计算机体积缩小了 30 万倍，功耗降低了 5 万多倍，重量降低了 6 万倍，平均故障率也大为减少，而且价格大幅度下降。这样为计算机的普及创造了条件。

随着芯片集成度的不断提高，单元体积和价格不断下降，图 2 说明芯片发展历程及与硅晶片尺寸的关系。可以看出，40 年间（1958~1998），芯片集成度提高了一百万倍，每单元价格下降到一百万分之一，这是因为单元价格与特征尺寸（feature length）的平方成正比，与晶片直径的平方成反比。特征尺寸的缩小，一方面与制作技术有关，另一方面与相关材料的不断改进有直接关系。为了适应集成度的不断提高，特征尺寸不断缩小，晶片尺寸要不断增加，对硅单晶的质量要求也要不断提高。由于芯片密度的大幅度提高，用以制造计算机的性能大为提高，如 20 世纪末计算机的计算速度已逾每秒万亿次。信息存储是现代化的另一重要标志，其要求是容量大，密度高，易于快速随机存取，可擦除和反复使用，这就要求材料不断改进才能满足。迄今已出现磁存储、半导体存储和光存储等，如一张光盘可以存储 10 万幅图像或 50 万页文字信息量。计算机是工业自动化的关键，计算机控制的精度决定于传感器的敏感程度。因此，高精度、高灵敏度、性能稳定的各种类型的敏感材料便成为关键。

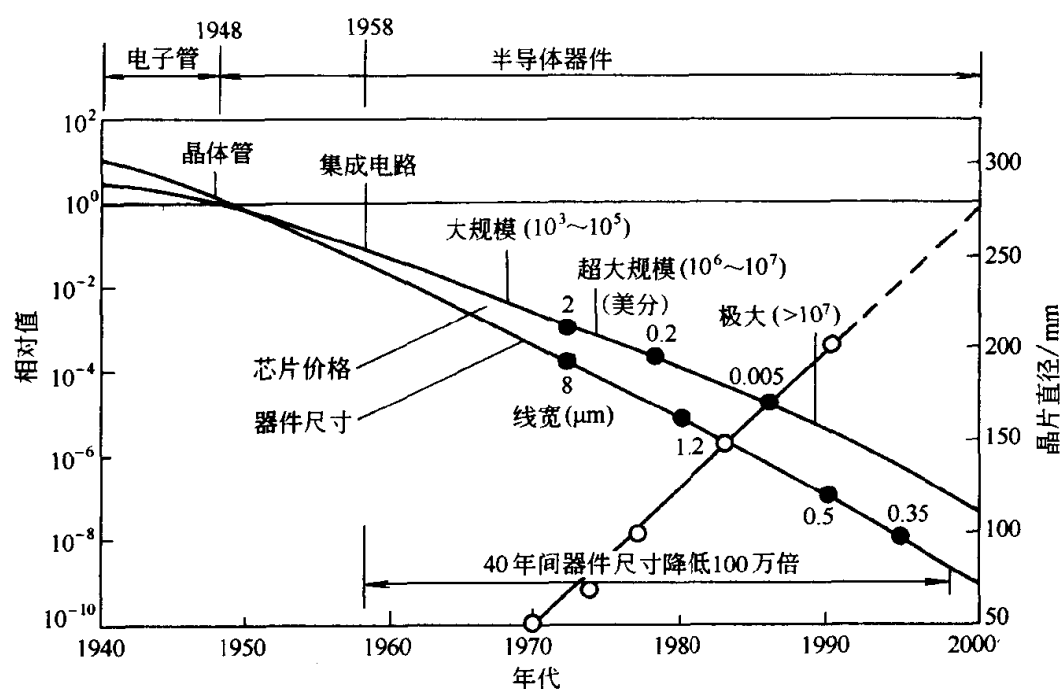


图 2 集成电路发展历程

光纤通信的诞生

1966 年高琨提出当光纤传输的损耗小于 20dB/km 时可实现光通信。

1970 年采用石英掺杂氧化锗等, 达到了这一指标, 而后光纤通信比集成电路发展更快, 代替了同轴电缆。因为其信息容量比同轴电缆大几个量级, 而且重量轻, 消耗原材料少, 保密性强, 又抗电磁干扰, 所以几十年内, 从连接各大陆的海底电缆到很多家庭, 都遍布了光缆, 成为现代信息高速公路与信息网络化的重要组成部分。光纤通信除了光导纤维以外, 还要有激光器和探测器, 这些都是化合物半导体。表 1 为光纤通信系统的发展历程, 除第四代尚在研制阶段外, 其余都达到应用状态。

表 1 光纤通信的发展历程

技术指标	第一代 (短波长)	第二代 (长波长)	第三代 (长波长)	第四代 (超长波长)
波长/ μm	0.83	1.3	1.55	2~5
光纤	熔石英	熔石英	熔石英 (+ 铒)	氟化物玻璃
激光器	GaAlAs	InGaAsP	InGaAsP	GaInAsSb
探测器	Si	Ge, InGaAsP	InGaAsP	InAsGaSb
损耗/ $\text{dB}\cdot\text{km}^{-1}$	2~3	0.5~1	<0.1	0.1~0.001
传输距离/km	约 10	约 30	约 400	>500

除光纤通信外, 通信设备小型化是“漫游电脑”时代的必要条件, 因此, 促进了低电压、低功耗、高速集成电路的发展。

航空航天技术的进步

21 世纪全球经济的一体化, 一靠信息网络的大发展, 形成全球性的电子商业; 二是运输工具的高效、远程、大容量。前者是电子及光电技术发展的结果, 后者是航空技术与水上、地面运输工具不断改进的结果, 这些都与材料有着密切联系。以飞机而论, 由于战争的需要, 20 世纪 40 年代出现了喷气技术, 这种技术的实现是以高温材料及高性能结构材料为依托的。高温合金的发展, 使涡轮温度不断提高; 高比强度、高比刚度结构材料的不断进步, 使今天大型客机的安全性及有效载荷大为提高, 续航时间不断延长, 飞机发动机寿命显著延长, 油耗不断下降。有人对歼击机的发展有一个估计, 飞机性能的改善有 2/3 依靠于材料, 航空发动机性能的提高在很大程度上也有赖于材料的改进。由此说明材料对飞机的重要性。地面运输工具如汽车和火车, 水上运输工具如轮船的发展莫不与材料有密切关系。

航天与卫星是科学技术现代化的重要标志之一。对材料的要求除了高比强度和高比刚度之外, 还需要耐超高温、抗辐射、耐粒子云以及原子氧侵蚀等的材料, 对比强度和比刚度的要求更为苛刻, 因为航天飞行器每减重 1kg, 可使运载火箭减轻几百千克。这就说明材料在航天技术的重要性。

从以上几个方面,说明材料,特别是先进材料与现代文明的关系十分密切,而且先进材料是先进技术的先导。没有半导体和其他功能材料,就不会有今天的信息社会,没有高温和超高温材料及高比强度、高比刚度材料,就不会有今天的航空航天技术,从而全球经济的一体化、人类社会的频繁往来就会发生很大困难。

传统材料在国民经济中的地位与可持续发展

谈到传统材料往往和“夕阳工业”联系在一起。实际上传统材料与国民经济支柱产业是密不可分的。钢铁曾经是衡量一个国家实力的重要标志,今天在一些工业发达国家也仍把钢铁视为支柱产业,因为钢有不可代替的优良性能,其价格又十分低廉,表2为一些结构材料的价格与其比强度的关系。钢的价格与比强度的比值仅高于水泥和木材,而比任何其他材料都低。同时根据多年统计结果,很多机械产品和材料的价格都在上涨,而钢的价格由于其工艺的不断改进、降低能耗和提高劳动生产率而长期未上涨。建筑行业在任何一个国家都视为支柱产业,建筑材料是建筑行业的重要组成部分。它涉及人民生活的提高,量大面广。水泥、玻璃、木材以及砖、瓦、石材都属于传统产业。汽车工业在很多国家也视为支柱产业,但是所用材料钢、铁、铝、塑料及玻璃等占了汽车用结构材料的90%~95%。人工合成纤维、树脂、塑料、橡胶在国民经济中占有非常重要位置,而且逐年增加,这都属于传统产业。此外,机械制造、造船、机车等莫不是以钢铁及其他传统材料为基础的,所以说传统材料是国民经济的基础,不能稍有忽视。

表2 一些结构材料的价格与其比强度的关系

材 料	钢	铝	钛	水泥	Al ₂ O ₃	碳纤维	木材	聚丙烯
比强度	5.2	11.1	13.3	0.8	97.4	160.9	12.5	3.0
价格/比强度(相对值)	1	3.0	16.7	6.4	4.0	5.2	0.8	3.8

另外,传统材料有以下几个特点。

① 传统材料量大面广。传统材料只要有一点改进,收益就十分可观。据估计,美国道路与桥梁的使用寿命增加1%,其收益可达300亿美元。世界易拉罐每年消耗量以600亿计,其造价每个省1分,则有6亿元的收益。所以对传统材料性能的改进和生产成本的下降必须给以足够的重视。

② 传统材料是矿产资源消耗大户,根据专家估计,全世界铁矿资源最多能维持200~300年,铝矿近百年,而铜、铅、锌矿只能够几十年。因此,除了提高矿产的收得率外,要开拓新的矿源,探索某些关键材料的代用品。如铝和镁有大致相似的性质,而镁在地壳表层储量仅次于铁、铝,居第三

位,但在海水中储量巨大 (2.1×10^{15} 吨),可以说取之不尽,用之不竭,所以开展以镁代铝的研究应给予重视。

③ 传统材料在制备过程中污染严重。根据美国统计,含毒污水的排放,近 80% 来自化工、钢铁与有色金属冶炼,因而有机材料和金属是重要污染源。水泥生产除了高能耗外,也是一个重要污染源。1997 年我国生产水泥 5 亿吨,占世界产量近 40%,产生 CO_2 近 4 亿吨, SO_2 49 万吨, NO_x 100 万吨,粉尘 1100 多万吨。这些废气加剧了温室效应,促进了酸雨。而炼 1t 钢所产生的 CO_2 至少 2 倍于水泥。所以从大气污染看,材料的生产对环境的污染是十分严重的。

资源的大量消耗和环境污染都不利于材料的可持续发展。为此,必须注意以下几点。

① 开发节约资源、低污染的生产流程。当前正在开发一种所谓“零排放”流程,即所输入的原料及能源全部生成产品而没有废物排出,这当然是一种理想状态,但是至少应尽量开发充分利用资源而少污染的生产流程。

② 发展环境友好材料,或称环境材料 (ecomaterials)。就是指与环境相适应的材料,即节约资源、少污染、易回收或可降解的材料。

③ 开发高性能、长寿命材料是节约资源、减少污染最有效的途径。因此,应把提高传统材料性能作为主要奋斗目标。如钢的强度大幅度提高;开发既抗腐蚀又可大幅度提高强度的水泥,以减少水泥用量,从而减少污染和节约资源。

④ 用新技术改造传统材料生产流程。一方面提高劳动生产率,改善产品质量,降低成本,另一方面使传统材料升级换代,扩大材料的用途,以增加竞争能力。

总之,传统材料仍有很多值得探索的科学技术问题。即使在工业发达国家,也都花很大力气在这方面开展研究与开发工作,否则便无竞争能力。根据我国 1998 年在《国家重点基础研究发展规划》中的新一代钢铁材料重大基础研究的结果表明,利用钢的晶粒细化、成分的纯净化和均匀化,在不加任何元素的情况下,可使碳钢的强度从 200MPa 提高到 400MPa,使低合金钢的强度从 400MPa 提高到 800MPa,而且生产实践证明,在不明显改变现有生产条件就可实现工业生产。

材料科学技术发展的重点

材料科学技术为开发先进材料、改进现有材料和合理使用材料服务。根据国际动向并结合国内具体情况,材料科学技术发展的重点有以下几个方面。

材料制备工艺与技术的开发

任何一种新材料从发现到应用于实际，必须经过适宜的制备工艺才能成为工程材料。高温超导自 1986 年发现以后到 20 世纪末，已有 15 年的历史，但仍不能普遍应用于电力，主要是因为没有找到价廉而稳定的生产线材的工艺。 C_{60} 也是如此，尽管在发现之初认为它的用途十分广泛，但到 20 世纪末仍处于科研阶段。传统材料也需要不断改进生产工艺或流程，以提高产品质量、降低成本和减少污染，从而提高竞争能力。

分子束外延技术的出现，可以控制薄膜的生长精确到几个原子的厚度，从而实现了“原子工程”或“能带工程”，为原子、分子设计提供了有效手段。快冷技术（即每秒冷却速度达 $10^4 \sim 10^8 K$ ）的采用，为金属材料的发展开辟了一条新途径。首先是金属玻璃的形成，提高了金属强度、耐磨耐腐蚀性能和磁学性能。其次通过快冷可得到超细晶粒，成为改进性能的有效方法。第三是通过快冷发现了准晶，由此改变了晶体学的传统观念。所以材料制备方法的研究与开发成为材料科学技术的重点。

材料制备工艺的重点：一是工艺流程的智能化；二是实现原子或分子加工，使材料或器件依照人们的意志达到微型化、多功能化和智能化。

材料的现代制备工艺或技术往往与极端条件密切相联，如利用空间失重条件进行晶体生长，可得出无偏析或低偏析的材料；强磁场、强冲击波、超高压、超高真空以及强制快冷等都可能成为改进材料性能的有效手段。

材料的应用研究与开发

材料的广泛应用是材料科学技术发展的主要动力，实验研究出来的具有优异性能的材料不等于具有实用价值，必须通过大量应用研究，才能发挥其应有的作用，本书既不涉及具体材料，更不会谈到材料的应用，因此这里对材料的应用问题稍加讨论，如此对材料有一个比较全面的了解。

材料的应用要考虑以下几个因素：一是材料的使用性能（performance）；二是使用寿命（durability）及可靠性（reliability）；三是环境适应性（environmental compliance），包括生产过程与使用期间；四是价格（cost）。当然，不同材料及使用的对象不同，考虑的重点就不一样，有些量大面广的材料，价格低廉是主要的，因而生产要低成本，检验不十分复杂，如建材与包装材料；相反，有些关键技术所用关键材料，如航空航天及医用生物材料，一旦发生意外，则损失严重，因而必须高质量、安全可靠，加强检验，否则后果不堪设想，所以有时检验费用比材料本身花费还高。以航空发动机所用高温合金为例，作为涡轮叶片及涡轮盘材料，一旦在飞行过程中出现断裂，很可能造成机毁人亡，因而在要求长寿命（几万小时）的同时，对可靠性的要求特别严格。为了保证材料的质量，采用三次熔炼、真空感应炉熔炼，以保证

严格控制成分（去气、去有害杂质）；再用电渣重熔，以去除非金属夹杂物；最后真空自耗电弧重熔，可以得到无宏观缺陷的合金锭，如此保证材料质量的均一性和完整性，再经锻造，或重熔铸造加工成零件，最后经过高灵敏度的检验合格后，再装机使用。对医用生物材料来说，质量保证更为严格，因为一旦因质量事故而产生不良后果，则后患无穷。

材料是否有竞争能力，除了质量和质量稳定性以外，还有材料的生产成本和价格。价格的影响因素很多，其中产量是决定因素之一。因为高产量不但可以实现自动化，而且质量稳定，成品率大幅度提高，成本明显下降。因此扩大应用范围是促进生产量必由之路。

材料应用研究又是机械部件与电子元件失效分析的基础。失效分析的准确性与时效性代表一个国家的科学技术水平，因而通过材料的应用研究来培养一批既有实践经验又有理论基础的工程技术人才是一个重要途径。通过材料应用研究也可以发现材料中规律性的东西，进而指导材料的改进和发展。

开发先进材料，发展高技术产业

当今是高技术主宰着社会，一方面高技术促进社会的发展，又是国防安全的保证，另一方面高技术是传统产业改造和发展支柱产业不可或缺的组成部分，而先进材料又是高技术的先导和基础，所以开发先进材料必须受到高度重视。根据目前情况，先进材料的重点有以下几个方面。

(1) 信息功能材料将得到更高的重视 人类已进入信息时代，为了实现装置的小型化、低功耗、多功能化和智能化，信息功能材料将受到更高的重视。所谓信息功能材料，就是指信息的产生、获取、存储、传输、转换、处理、显示所需材料。主要用于计算机、通讯和控制，成为 3C (Computer, Communication 及 Control) 所需材料。其特点是要求高，发展快，种类繁多，这里不可能一一列举。但总的来说，信息技术 (IT, Information Technology) 以集成电路为基础 (细胞)，计算机为中枢 (大脑)，网络为媒介 (神经)。

现仅以集成电路所需材料为例说明。图 2 为集成电路发展的历程。根据 Moore's 定理，到 2010 年集成电路的特征尺寸为 $0.07\mu\text{m}$ 。由于制作困难，再加发热过高以及量子效应等可能达到极限，这时硅晶片的直径将在 450mm 以上，对硅单晶质量要求进一步提高，如表 3。因此硅单晶的生长一直是研究的主题，因为晶体直径愈大，杂质的偏析愈严重，愈容易出现缺陷。在硅集成电路达到极限以后，下一代集成电路将如何发展，是目前正在探索中的问题：量子点、光集成，还是采用别的途径？

计算机作为控制装置的组成部分，敏感元件传感器是关键，因此敏感材料的灵敏度和稳定性决定计算机可能控制的精度。敏感材料是利用材料对物

表 3 1998 年公布的硅片表面颗粒度指标

参 数	年 份					
	1997	1999	2002	2005	2008	2011
IC 工艺/nm	250	180	130	100	70	50
颗粒大小/nm	125	90	65	50	35	25
颗粒密度/cm ⁻²	≤0.14	≤0.088	≤0.055	≤0.038	≤0.023	≤0.014

理、化学和生物的变化而产生不同反应发展出来的，种类繁多，发展很快。表 4 列举了几种无机非金属敏感材料及其原理，由此说明对物质的基础研究至关重要。

表 4 无机非金属敏感材料

探测性能	原 理	材 料	探测性能	原 理	材 料
氧含量	体离子导电	Zr _{1-x} CaO _{2-y}	电压	晶界面隧道	ZnO·Bi ₂ O ₃
湿度	表面体离子导电	MgCr ₂ O ₄ ·TiO ₂	PTC 热敏电阻	晶界面相变	Ba _{1-x} Ce _x TiO ₃
压力	压电	PbZr _{1-x} TiO ₃	化学	表面电子导电	ZrO·CuO
酸度	表面化学反应	IrO _{2-x}	光学	光电阻	CdS

(2) 先进结构材料的研究与开发是永恒的主题 结构材料一般数量大，资源与能源消耗高，污染严重，对可持续发展有决定性作用。结构材料性能的提高，无疑会减缓上述压力，因性能的提高，使用寿命的延长，可减少材料的用量。特别对运载机械来说，工作温度愈高，效率也愈高，所以，高温材料的研究与开发是当务之急。高比强度、高比刚度，对提高机械性能十分重要；耐磨、耐蚀、抗疲劳、抗老化是延长使用寿命的关键，因此，在材料的发展中，必须给予高度重视。从材料类型来说，金属材料在 21 世纪上半叶仍占主要地位。先进陶瓷材料要进一步提高韧性，降低成本。有机材料以其再生资源的优势和优异性能，而且可实现分子设计而进一步得到发展。先进复合材料类型很多，有些已得到广泛应用，如碳纤维增强树脂基复合材料；有的因价格高，制作难度大而需进一步发展新工艺，如金属基复合材料。有些超高温材料，如碳/碳复合材料，必须解决抗氧化问题才能得到更大范围的应用。

(3) 能源材料的开发有广阔的前景 能源的利用是人类进步的标志之一。目前化石燃料一是储量有限，二是污染严重，必须大力开发无污染、可再生能源。太阳能虽然密度低，又受气候影响，但辐射于地表的能量，一万倍于人类开发的能源。因此，开发光-电转换效率高而又价廉长寿命的材料是当务之急。海水中氢的同位素氘可谓取之不尽，用之不竭，认为是人类的最终能源。科学家已采用多种方法（如等离子、激光）点火，实现可控热核聚变，用聚变能发电，预测 21 世纪内可实现商业化，其中抗辐射、耐高温、

耐氢脆材料是关键之一，必须得到解决。除开发新能源以外，节能也十分重要，如超导的利用。低温超导以金属（NbTi 合金、Nb₃Sn 等）为主，但必须液氮冷却（4K），成本太高而难以普及。20 世纪 80 年代发现可在液氮温度下出现超导现象的氧化物超导体，称为高温超导，但制成符合工业应用的线材和块体还存在差距，不能普遍应用。利用超导输电，可以减少线路损失；超导储电，效率可提高；超导电机及超导电池功率大、体积小、损耗低，十分诱人。目前正在探索室温超导，一旦有所突破，前途将更是无量。

（4）有机高分子材料将有更大发展 有机高分子材料以其可再生，资源丰富，又有优异性能，特别可实现分子设计，将会更大发展。高分子聚合物不但是重要的结构材料，而且正在发展成为重要的功能材料，从半导体到超导体都有发展前途，作为电导体，其电导率可与铜相比，高分子光学材料种类繁多，正在研究开发。高分子材料也具有铁磁性，这些都是 21 世纪要开展研究与开发的重要领域。但是高分子的某些缺点也是必须下大力气才能得到解决的，如稳定性、抗老化性及阻燃性。

（5）生物材料将受到更大的重视 随着生物技术的发展及人类寿命的延长与生活质量的提高，生物医用材料已成为人们非常关注的领域，人的器官更换、药物缓释及组织工程的发展将逐步深入。生物模拟是另一正在兴起的学科，使材料的功能进一步提高，并能达到自恢复、自修复或智能化。生物材料的更长远目标是使生物技术原理用于工业化生产，改变高温、高压及耗能高的生产方式，催化剂已迈出第一步。光合作用使水和 CO₂ 合成碳水化合物，人类正在寻求利用生物技术，通过像催化剂的效应一样以工厂方式合成粮食，如此除了可以解决世界粮食问题，CO₂ 过剩问题也可得到解决。当然，这是很长的历史过程。

（6）纳米材料及制备技术的研究与开发迫在眉睫 当物质到纳米尺度时，由于其尺寸效应、晶界效应和量子效应等，材料显示出奇特的物理、化学性能，或其生物功能有明显改变，利用这一效应可大幅度提高结构材料的强度，改善其脆性。对功能材料来说，应用方面更广，高效率的催化作用，以及多变的光学性质、电学性质、磁学性质和热学性质等，将使目前的功能材料性能明显提高。这些改变将影响到工业和国防的各个方面，所以有人认为纳米技术将如信息技术或生物技术一样，将导致下一代工业革命。应该指出，纳米科学技术仍处于基础研究阶段，特别是纳米电子学、纳米医学所需材料，尚处于探索阶段，但是纳米技术用于结构材料的改性及用作某些功能材料已显示出明显的优越性，有些已进入产业化。

材料设计

材料通过设计而得到所需要的性能是材料科学技术工作者的奋斗目标。

材料的研制从以经验为主过渡到以科学设计为主，从而进入“设计材料”(materials by design)时代。材料设计概念的风行始于20世纪80年代初，随着有关学科的发展，设计材料愈来愈受到重视。材料设计可以分为几个层次，最基础的是量子设计(quantum design)，这是由于电子运动而引起的多种现象，如光学，这是功能材料的基础。其次为原子设计(atomic design)，这是纳米技术的基础。原子排列是决定材料力学性质与化学性质的基础，原子操纵可能是下一代芯片的一种技术。第三个层次是微观设计(micro-design)，即微米级结构的设计，金属的相变、晶界的控制都属于这一范畴。第四个层次是宏观设计(macro-design)，以毫米到厘米为对象，像金属在凝固过程的结构与偏析便属于此。

材料设计的最终目标是根据最终需求，设计出合理成分，制订最佳生产流程，而后生产出符合要求的材料。应该指出，材料设计十分复杂，如模型的建立往往是基于平衡态，而实际材料多处于非平衡态，如凝固过程的偏析和相变等。材料的力学性质往往对结构十分敏感，因此，结构的任何微小变化，性能都会发生明显变化，何况有些性质，如脆性、裂纹的萌生与扩展等还很不清楚，因此，要想得到确切真实的结果决非易事。另外，相图也是材料设计不可或缺的组成部分。

可以肯定，随着科学技术的进步，各学科的发展与对材料的了解日益深入，计算机日益巨型化和智能化，一个更加科学化发展新材料的时期将会到来。要做到这一点，最关键的是建立完善的数据库与知识库，发展符合实际的物理模型，这就要求应用数学家、物理学家、化学家、材料科学家与工程技术等方面人员的密切合作。

科学仪器与检测装置

科学技术的发展很大程度上依赖于新科学仪器的不断发明和性能的不断提高，以物理学诺贝尔奖金获得者为例，大约有半数得益于新仪器的发明或新测试手段的采用。工业产品质量的改进往往取决于检测装置精度的提高。所以工业发达国家对科学仪器与检测装置的开发与制造都给以足够的重视。以电子光学仪器为例，1863年金属材料研究首先应用光学显微镜。几十年后出现了电子显微镜、扫描电镜(SEM)、高分辨电镜(HREM)，其分辨率达0.2nm，足以观察到原子。后又有扫描透射电镜(STEM)，不但可以观察到原子，而且可以分析微小区域的化学组成和结构。20世纪80年代初出现的扫描隧道显微镜(STM)，可以在非真空条件下观察到原子，并用来进行原子加工，一种更为引人注意的显微技术称之为扫描探针显微镜(scanning probe microscope)，简易价廉，它基于压电原理，可以在不同结构中观察到单个原子或分子。原子力显微镜(AFM)可以观察金属脆性或韧

性断裂过程。红外原子力显微镜 (infrared AFM) 可以观察高分子聚合物的苯环。这些分析装置与高性能计算机相结合, 可以得出结构细节, 使结构分析从以经验为主, 费时而不准确, 过渡到直接、快速地解决材料中的问题。

除显微结构以外, 材料的晶体结构分析开始于劳厄 (Laue) 等, 于 1912 年发现 X 射线通过晶体产生衍射花样, 随后又推广到电子衍射与中子衍射, 成为探明材料原子结构的重要手段, 而后又出现了基于不同原理的多种谱仪, 成为分析表征材料结构与性能的有力装置。材料科学工作者不仅利用现有先进设备研究和开发材料, 同时也应致力于发明与创造新的检验与测试装置, 因为任何已商业化的设备至少落后了 5 年, 突破性的工作很难在这些已商品化的设备上做出来。

检测是控制工艺流程及产品质量的主要手段, 如无损检测装置不但可以检查宏观缺陷, 也可监控裂纹的萌生与扩展, 提高无损检测装置的精度十分重要, 如先进陶瓷的临界裂纹, 只有几十微米, 没有精密的无损探伤仪, 陶瓷的质量控制便谈不上, 当前检测装置除了扩大品种, 提高灵敏度和稳定性以外, 主要发展在线、动态以及适应各种恶劣环境的检测装置。

总之, 材料的重要性任何时候都不会下降, 但是随着科学技术的进步, 材料与器件或构件的结合愈来愈密切, 全新的人造材料或人工合成材料愈来愈占更主要的位置, 具有创造性的新材料将层出不穷。但是, 在很多情况下, 新材料的出现往往难以预料, 如高温超导和 C_{60} 等, 因此, 材料科学工作者必须富有探索精神, 同时也要为他们建立一个宽松环境和氛围, 在打好本学科基础上, 重视交叉学科, 广泛接触实际, 抓住任何一个“灵感”火花, 才会做出创新成果。

第 1 篇 材料的结构

本篇的主题为材料的结构，这是材料科学的核心问题之一。在进入正式论述之前，用一章篇幅鸟瞰一下材料结构所涉及的不同学科的侧面。下面，用两章来论述材料的理想结构，从晶态出发，延伸扩大到晶态之外。再用三章来论述材料的实际结构，探讨不同维度的晶体缺陷问题和不同物态间的界面问题。最终，以对多层次和非均质结构的讨论，来扩大视野，进一步认识材料结构中所蕴含的复杂性。

第 1 章 结构问题的不同侧面

撰稿人 冯 端 俞圣宪

我们可以从不同的学科侧面来观察材料结构的问题。首先，从化学的侧面来理解原子是如何键合成分子或晶态与非晶态的物质；其次，从晶体学的侧面来分析对称性在结构方面所起的作用；再从统计热力学的侧面来认识对称破缺所扮演的重要角色；最终，从信息论的论点来阐述结构与信息的关系，强调了结构中的信息在生物学中的重要意义。

1.1 键合

1.1.1 原子结构和周期表

材料的最基本的结构单元是原子（包括带电的原子，即离子），原子的基本性质都包含在元素周期表内。元素周期表首先由门捷列夫等化学家 19 世纪从经验规律中概括出来的，到 20 世纪量子力学建立后方才奠定了其理论基础。这里，扼要地复述一下为理解材料结构所需的一些主要结果。

氢原子只有一个电子，处于球形对称的库仑场中，因而可用薛定谔方程确切地求出其能级结构。这一结果对我们理解周期表给予了富有启发性的主导线索。当然，周期表中的其他元素的原子含有 Z （ Z 为原子序数）个电子，其中心为带有 $+Z$ 电子电荷的原子核。这类多电子体系，由于电子不仅受到核的库仑场的作用，还受到其他电子的库仑相互作用，这使得电子能级的计算成为复杂的量子力学多体问题。由于很难得到确切解，通常乞助于单电子近似来求解。

根据对氢原子中电子态的量子力学计算结果,我们将多电子原子中的电子态也用四个量子数来描述,前三个量子数 n (主量子数), l (角量子数), m (磁量子数)规定了电子的空间(或轨道)运动,而 m_s (自旋量子数)则规定了电子的自旋($+\frac{1}{2}$ 或 $-\frac{1}{2}$)。

一个自由、孤立的原子,其电子的实际分布区域是从靠近原子核起一直散布到无限远。由于电子的能级是确定的,根据不确定原理,其位置是在空间中散布的。由薛定谔方程的解,可得到电子的径向分布函数。根据电子的径向分布函数,可以确定电子出现大几率的区域范围,所以通常外层电子径向分布函数的极大值被视为原子的轨道半径 r_0 。比如说氢原子,其径向分布函数的极大值就作为它的玻尔半径。大约有 50% 的电子落在玻尔半径(0.0529nm)内,而在 0.25~0.3nm,则电子出现的几率为 99% (如图 1-1)。对于孤立原子来说,电子的大致分布可以用玻尔半径来估算,但是其实际尺度要稍微大一点,仅能做到量级上的估计。

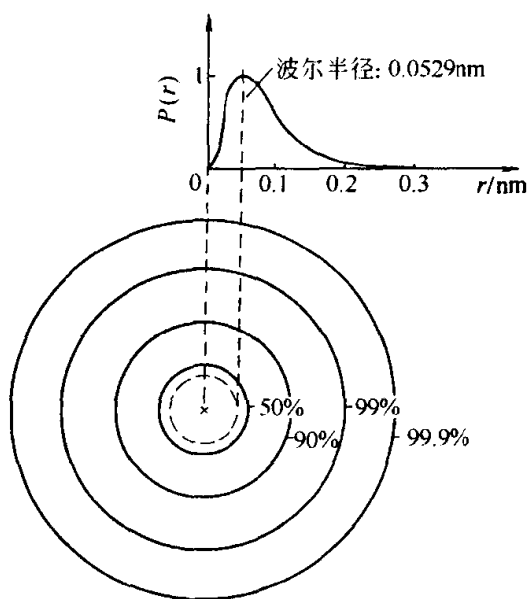
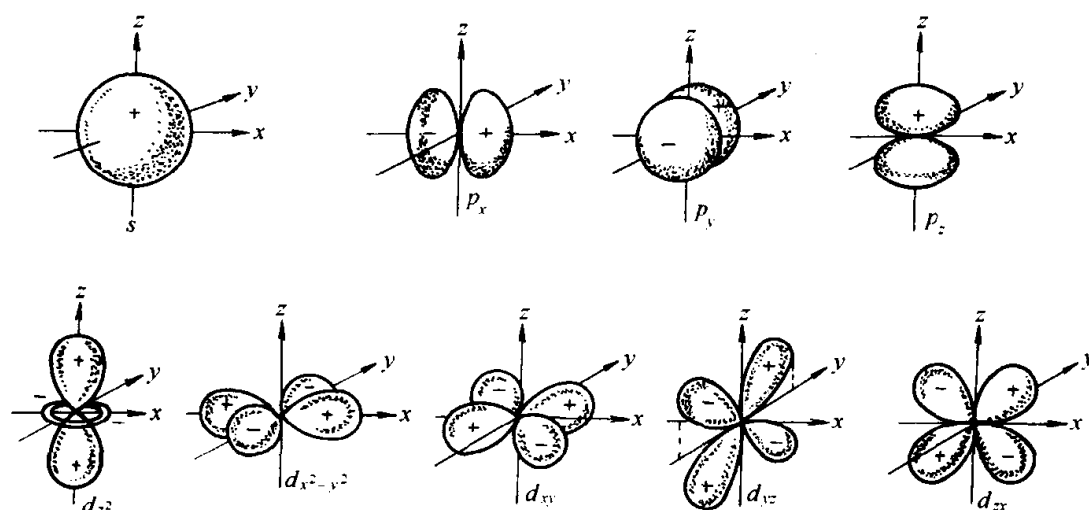
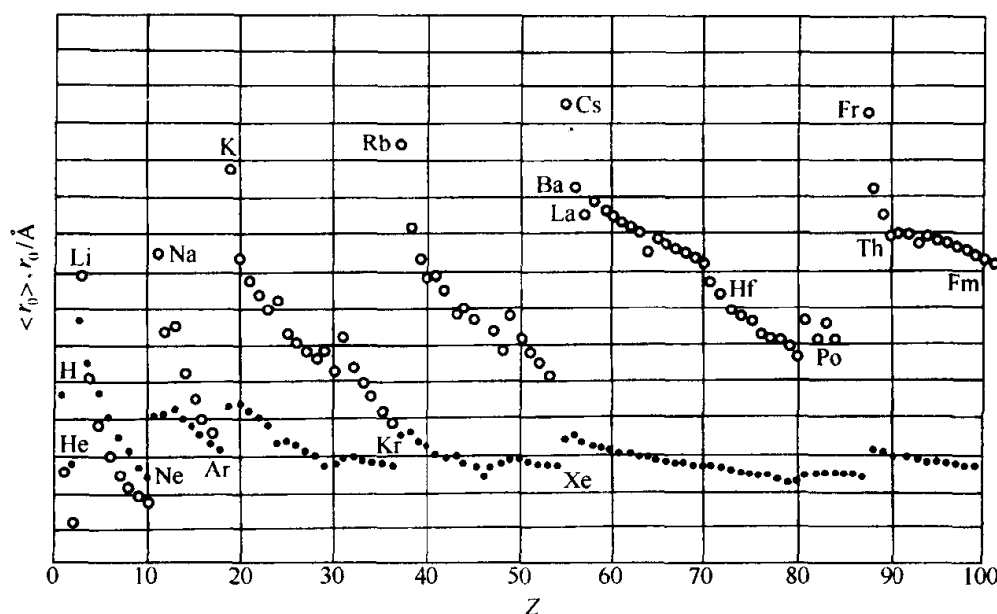


图 1-1 氢原子电子的径向分布

除了 s 轨道($l=0$)是球形对称分布外, p ($l=1$), d ($l=2$)轨道分布的形状均偏离球形对称。图 1-2 表示了计算出来的 s , p , d 波函数在空间分布的情况。根据壳层模型,可大致列出元素的电子组态和轨道半径 r_0 。图 1-3 显示了原子半径值与 Z 的关系。可看出,在同一周期内, r_0 随 Z 的增大而缩减,而随新的壳层的被填充又增大(其数值在 $0.8 \times 10^{-10} \sim 2.5 \times 10^{-10} \text{m}$ 之间)。

电子的具体布居遵循以下原则。

① 泡利不相容原理:同一电子态只有一个电子占据。

图 1-2 原子的 s , p , d 轨道的示意图图 1-3 中性原子的均方根半径 $\langle r_0 \rangle$ (黑点) 和外层轨道半径 r_0 (圆) 和 Z 的关系注: $1\text{Å} = 0.1\text{nm}$

② 能量最低原理: 电子优先地占据能量较低的原子轨道, 使整个原子体系能量最低。

③ 洪德定则: 在未填满的壳层中, 电子的自旋值应尽量地大。

主量子数 n 和角量子数 l 大体上决定了原子轨道能量 (即能级) 的高低。当 l 相同时, n 愈大, 能级 E 也愈高, 例如 $E_{1s} < E_{2s} < E_{3s}$; $E_{2p} < E_{3p} < E_{4p}$ 。当 n 相同时, l 愈大, 能级也愈高, 如 $E_{3s} < E_{3p} < E_{3d}$ 。如果

n 与 l 都不相同, 情况就比较复杂, 需要考虑电子间多体相互作用的一些效应, 诸如其他电子的存在会减弱原子核对外层电子的吸引力, 导致了原子的能级出现交错现象, 如 $E_{4s} < E_{3d}$, $E_{6s} < E_{4f}$, $E_{7s} < E_{5f}$ 。这类现象造成了过渡、稀土与钢系元素中电子结构的复杂性, 而洪德定则又使有些原子或离子具有磁性。图 1-4 为 Pauling 根据光谱实验数据以及理论结果所提出的多电子原子近似能级图。其中小圆圈代表原子轨道, 按能级高低顺序排列起来。电子从最低能级填起, 逐级填充, 形成了原子的壳层结构 (见表 1-1) 和周期表。能透彻理解并善于应用周期表是做一个材料科学家的必要条件。在本书的书后刊出元素周期表和与之相关的材料数据, 以备读者查阅。

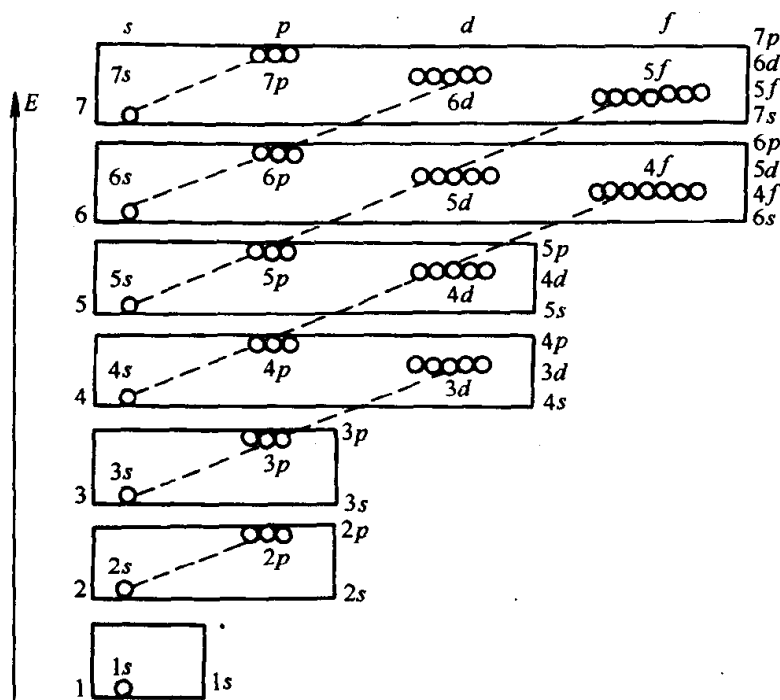


图 1-4 多电子原子中的能级图

对于多电子的原子, 它的化学键中真正起作用的是外壳层电子。对于大部分元素来说, 内壳层电子不起任何作用, 但是过渡、稀土与钢系元素中内壳层电子也可能起作用的, 特别是 d 或 f 壳层会参与一些效应, 而更深壳层的电子则可完全忽略。

1.1.2 离子键

正负离子间的静电相互作用是离子键的根源, 首先来看两个原子为什么会形成一对正负离子。

一般说来, 周期表开端的元素的最外壳层只有一个电子, 其电离能 I^+ 小, 容易电离而失去电子:

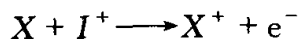


表 1-1 原子的壳层结构 (部分)

量 子 数				原子轨道		电子层可容纳电子总数	
n	l	m	m_s	名称	容纳电子数		
1 (K 层)	0	0	$\pm 1/2$	$1s$	2	$2 \times 1^2 = 2$	
2 (L 层)	0	0	$\pm 1/2$	$2s$	2	$2 \times 2^2 = 8$	
	1	+1	$\pm 1/2$	$2p$	6		
		0	$\pm 1/2$				
3 (M 层)		-1	$\pm 1/2$			$2 \times 3^2 = 18$	
	0	0	$\pm 1/2$	$3s$	2		
	1	+1	$\pm 1/2$	$3p$	6		
		0	$\pm 1/2$				
	2	-1	$\pm 1/2$	$3d$	10		
		-2	$\pm 1/2$				
		+2	$\pm 1/2$				
4 (N 层)		+1	$\pm 1/2$			$2 \times 4^2 = 32$	
	1	0	$\pm 1/2$	$4p$	6		
		-1	$\pm 1/2$				
	2	+2	$\pm 1/2$	$4d$	10		
		+1	$\pm 1/2$				
		0	$\pm 1/2$				
		-1	$\pm 1/2$				
		-2	$\pm 1/2$				
	3		+3	$\pm 1/2$	$4f$		14
			+2	$\pm 1/2$			
			+1	$\pm 1/2$			
		0	$\pm 1/2$				
-1		$\pm 1/2$					
-2		$\pm 1/2$					
-3		$\pm 1/2$					

图 1-5 表示电离能随元素 Z 的变化。图中可以看到, 电离能随 Z 呈周期振荡; 同一周期中, 碱金属元素的电离能 I^+ 为最小, 容易失电子形成正离子。

原子除了吸收电离能失去电子外, 当其吸收电子变成负离子时, 即放出亲和能 I^- :

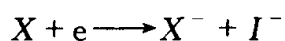


图 1-6 表示亲和能随 Z 的变化。同样的, 亲和能也随周期表的 Z 呈周期振荡; 同一周期中, 卤素元素的亲和能 I^- 最大, 容易得电子形成负离子。

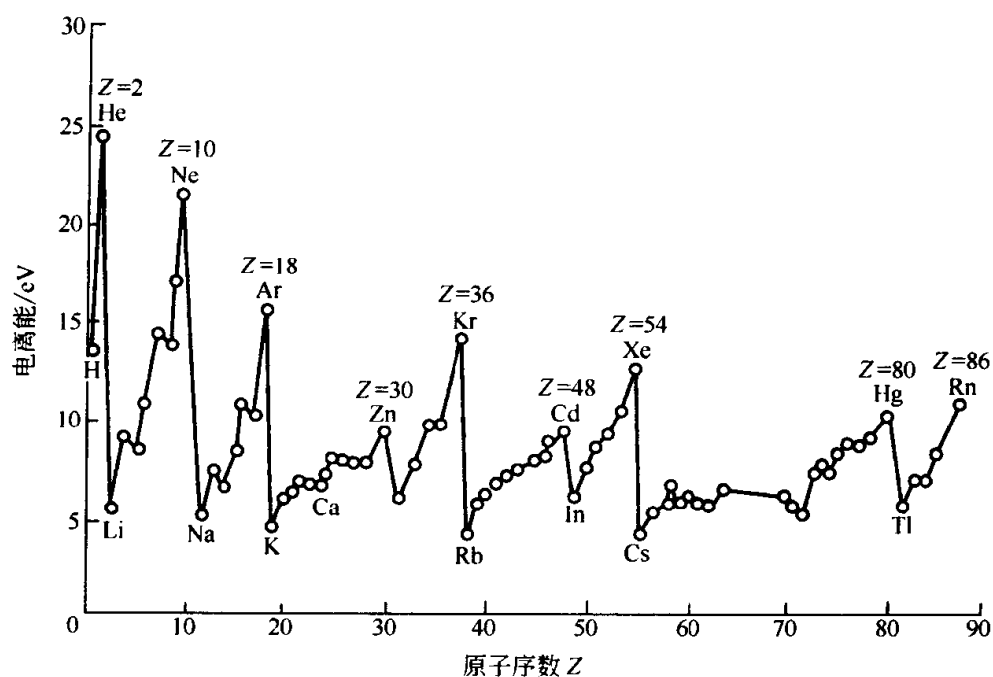


图 1-5 元素的电离势

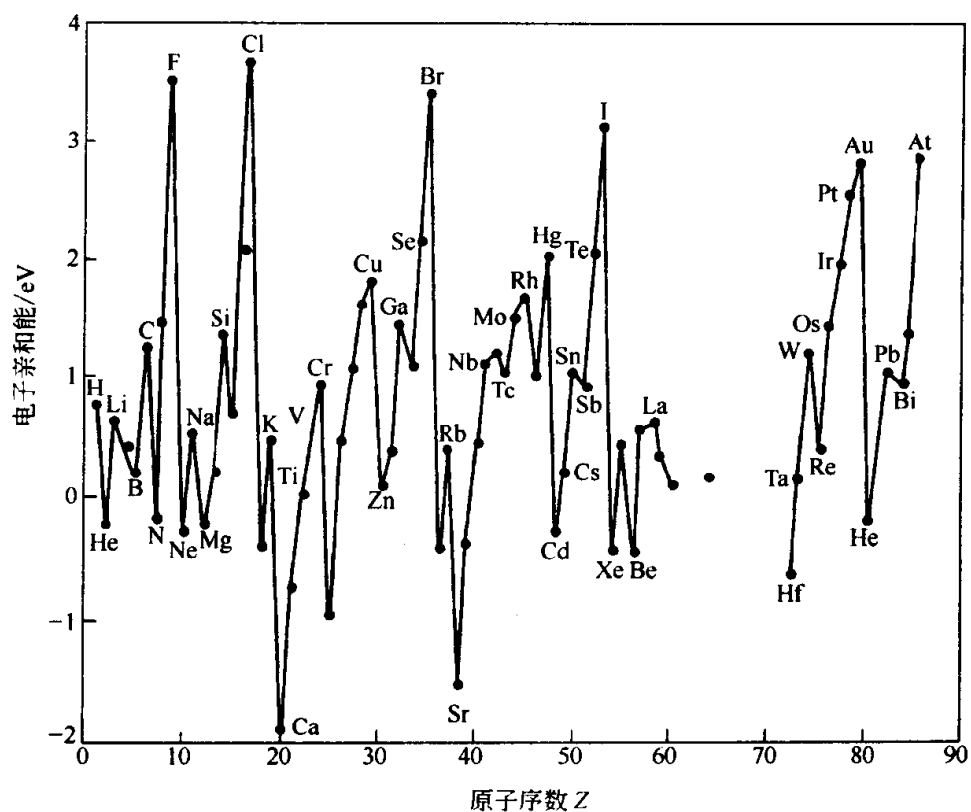
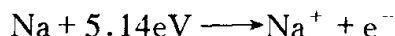


图 1-6 元素的电子亲和能

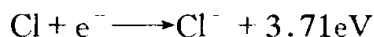
由于碱金属的低电离能 I^+ 和卤素原子的大的亲和能 I^- ，因此这两种元

素很容易形成离子键的分子。以 $\text{Na}^+ \text{Cl}^-$ 为例，可分为以下 3 个基本过程。

① Na 原子吸收能量，失去一个电子形成 Na^+ 。其电离能为 5.14eV 。



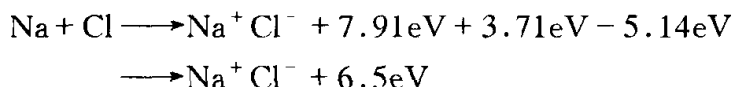
② Cl 原子得到一个电子形成 Cl^- ，释放能量。其亲和能为 3.71eV 。



③ Na^+ 和 Cl^- 通过静电相互作用，吸引而形成 $\text{Na}^+ \text{Cl}^-$ ，释放出 7.91eV 的能量。



因此，所得到的结果就是 Na 原子和 Cl 原子形成离子型的 NaCl 分子的同时释放 6.5eV 的能量：



NaCl 分子只存在于气体之中。固体 NaCl 中， Na^+ 为 Cl^- 所包围，反之亦然。这种形态就构成了我们通常所说的 NaCl 离子晶体。在 NaCl 离子晶体中并不存在独立的 NaCl 分子。也就是说，分子这种层次形式不存在于离子键形成的离子晶体之中。静电相互作用是构成离子键的主要能量来源。

1.1.3 共价键

共价键是化学键问题中最关键的一个问题。

离子键的形成需要电荷的转移，然而许多化学键的形成并不存在明显的电荷转移现象，而是同号电荷聚集在两个原子之间。也可以说是两个原子都贡献出电子，在这两个原子中间的区域有较高的电荷密度，抵消了部分原子核间的库仑斥力并且产生静电相互作用，因而有利于形成共价键。原子和原子之间有电子密度增大的区域的这个物理图像可以用来表征共价键。

共价键最基本的物理问题就是考虑其最简单的情况，即 H_2^+ 。这是一个单电子问题：一个电子在两个固定质子所产生的电场中运动，通过其薛定谔方程可以严格求出系列的能级解。

而一般的分子结构也涉及多电子体系，需要采用近似方法。其中最早使用的是由 Heitler 与 London 所提出的氢分子的价键理论：他们用量子力学方法来计算 H_2 分之能量 E 与两个 H 原子核距离 r 之间关系。如图 1-7 所示，如原子 A 和原子 B 的两个电子的自旋反向，则起成键轨道的作用[图 1-7 中的(b)]：即当 A、B 相接近时， H_2 分子能量降低，曲线的极小值对应于分子的平衡间距；如 A、B 电子的自旋平行，则起反键轨道的作用[图 1-7 中的(a)]。因而认为共价键的本质在于两个原子各有一个自旋相反的未成对的电子，由于原子轨道相重叠而构成价键轨道，导致体系的能量下降。

按照价键理论，共价键具有饱和性和方向性。各原子都有确定的不成对

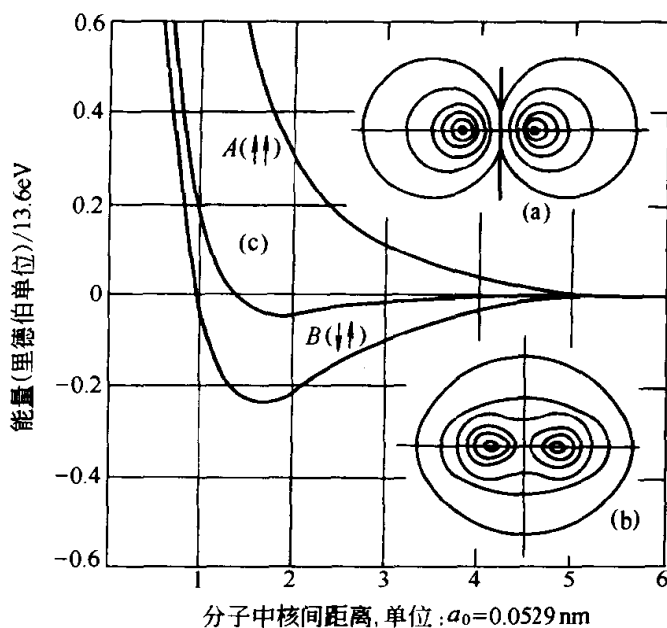


图 1-7 形成氢分子的能量曲线

(a) 两电子自旋同向形成反键态；(b) 两电子自旋反向形成键合态；(c) 经典曲线

的电子，对应于一定的共价键数。例如 H 和 Cl 为 1 价，O 与 S 为 2 价，N 为 3 价，C 为 4 价。另外电子轨道在空间有一定的取向，因而轨道的重叠也是有取向性的。例如 N 原子核外有 3 个未成对电子，其取向分别为 p_x ， p_y ， p_z ，因而对 N 分子 3 价键的方向也产生影响。

上述理论的基本概念为不同原子的价电子对占有了局域化的价键轨道，具有明确的直观图形，也和化学的经验规律吻合，但并不便于定量的理论计算。量子化学的主流采用另一种近似，即分子轨道法。这种近似的基本设想为电子系分子所公有。处于基态的分子中的电子占有了一系列最低能态的能级。对应于原子轨道的 s ， p ， d ， f 轨道，分子轨道亦有相应的 σ ， π ， δ ， ϕ 轨道。常见到的分子中含有 σ ， π 轨道。 σ 轨道的对称性有点类似于原子的 s 轨道，却远不及 s 轨道的球形对称，而呈圆柱形对称。 π 轨道则有点类似于 p 轨道。

参与共价键的主要是原子的外层电子（所谓的价电子），相邻原子壳层中的电子以共享的方式形成满壳层的稳定结构，产生强的共价键。周期表中共价键最明显的是 IV 族的 C、Si、Ge 元素。对称于 IV 组元素两边的 III 族和 V 族元素、II 族和 VI 族元素也是以共价键的方式组合形成稳定的化合物，成为半导体材料。

共价键是最主要的化学键，其键型（单键、双键或多键）、键长、键角等参量反映了它的特征。共价键是一种强键，如 C 键有高的结合能，这使

得有些有机材料由于有 C 元素的存在而相应地其强度也变高了。

共价键还可以与离子键相混合，形成部分的离子键。

1.1.4 分子

分子具有多种形式：从简单的氢分子 H_2 、水分子 H_2O 这种类型到相当复杂的有机分子，如苯环，乃至非常复杂的聚合物和生物分子，如胰岛素分子。

作为处于原子和大块物体之间的结构阶段，分子可以作为晶体的单元，这类晶体被称为分子晶体；也可以由少数或大量的原子或离子构成而单独存在。

晶体中是否有分子这个阶层要根据其键合的情况。一般而言，离子键、共价键和金属键的情况下，是不存在分子的。有些分子晶体是等同分子之间以范德瓦耳斯力键合形成。因此，分子之间的键合是弱键，而分子内部的键则以共价键为主。

近年来，分子的重要性在材料科学中日益突出，分子材料在材料领域中所占比例也越来越大，如高分子材料、液晶材料以及将来的生物材料都以分子为基础。

分子可以分为两大类：无机分子与有机分子。有机分子为含碳的化合物，而无机分子则是以碳以外元素为主的化合物。有机分子的品种数要比无机分子的品种多得多。无机分子中兼具离子键与共价键，而有机分子则比较单一，只有共价键。

分子具有特定的三维结构，呈现特征的形状。双原子分子的情况最为简单，两点确定一直线，因而其形状必然是线状的，原子以价键相连（可以是单键、双键或三键），如 H_2 ， N_2 ， CO 等。至于三原子分子，则有两种情况：其一是保持线状[如图 1-8(a)]，如 CO_2 ；也可以弯折[如图 1-8(b)]，如 H_2O 。更多原子所组成的分子，形态也就多样化了[如图 1-8(c)、(d)、(e)]。这就需要理论来理解其差异的根源。可以在价键理论的基础上提出更直观化的价电子壳层中电子对相互排斥这一基本概念：一个原子的外壳层中的诸电子对，由于斥力的作用，应相互回避，以保持最可能远的距离。在这里考虑的电子对既可以是与其他原子共享的价对 (valence pair)，也可能是非共享的孤对 (lone pair)，它们之间只有程度上的差异；而且不论是双键或三键，其主要效应是缩短键长，对分子总体形状影响不大，故不妨也当作一个电子对来处理。总的说来，2 个电子对分子的形状是线状的[如图 1-9(a)]，3 个电子对的形状是平面三瓣的(夹角为 120°)[如图 1-9(b)]，而 4 个电子对的形状是立体 4 瓣的，指向正四面体的 4 个顶角(各瓣之间的夹角为 109.5°)[如图 1-9(c)]。

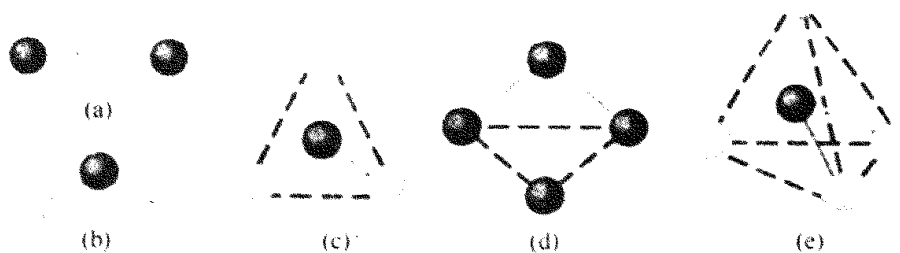


图 1-8 各种分子的形状

(a) 线状；(b) 弯折形；(c) 三角形；(d) 锥体；(e) 四面体

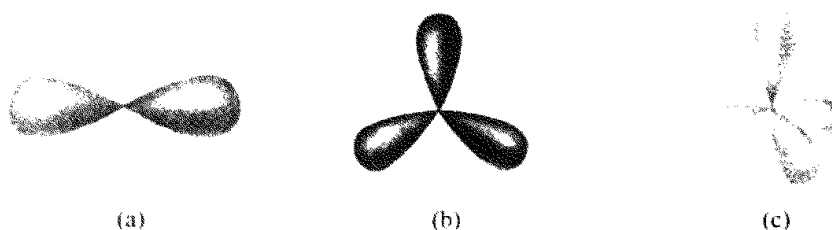


图 1-9 各种电子对的形状

上面有关价键具有方向性的论述尚有待于进一步从理论上来解释。Pauling 提出了杂化轨道的概念。以 CH_4 为例，原来 C 的 4 个价电子有 2 个是处于 $2s$ 轨道上，而另外 2 个处于能量略高一些的 $2p$ ($2p_x, 2p_y$ 或 $2p_z$) 轨道上，只能形成 2 个 p 轨道的键。如果将一个 $2s$ 轨道的电子激发到空的 $2p_z$ 轨道上，需要付出一定的能量，但可以形成 4 个键，对键能非常有利，成键能量不但能回收，而且绰绰有余。这样的 4 个键都是等同的，因而碳原子的键轨道即非纯的 s 态或 p 态，而是 3 个 p 轨道与一个 s 轨道形成的杂化轨道，用 sp^3 来表示，即每一轨道具有 $1/4s$ 态和 $3/4p$ 态[如图 1-10(a)]。一个 sp^3 轨道既非 s 态的球形，也非 p 态的哑铃形，而是两种形状的混杂。每一个 sp^3 轨道上含有一个电子，因而可以通过一个 C 原子的 4 个 sp^3 与 4 个 H 原子的 s 轨道相重叠而形成 CH_4 的 4 个 C—H 键，这些键都是 σ 键。

碳原子还有另一种杂化轨道，是一个 s 轨道与两个 p 轨道的杂化，称为 sp^2 轨道[如图 1-10(b)]。每一轨道具有 $1/3$ 个 s 态和 $2/3$ 个 p 态。这三个 sp^2 轨道处于同一平面上，夹角为 120° ，而剩下的 p 轨道则与此平面成 90° 。相邻原子的 sp^2 轨道相重叠形成 σ 键，而 p 轨道的侧面重叠则形成了 π 键(如图 1-11)。一个碳的双键是由一个 σ 键加一个 π 键。相对 σ 键而言， π 键较弱，因而 π 键只出现于 C 的双键或三键结构之中。

苯 (C_6H_6) 中的 6 个碳原子构成六角环的结构，如果用传统的单双键来表示，则具有等效的两种方式，而实际的结构则相当于两者平均值，即面

包圈式 π 电子云, 含有 6 个电子 (见图 1-12)。以苯环为基础可以衍生出许多具有苯环结构的分子, 如萘($C_{10}H_8$), 蒽($C_{15}H_{12}$), 并四苯($C_{20}H_{16}$), 并五苯($C_{25}H_{20}$)等, 其中一些是科学家注意研究的有机半导体材料。

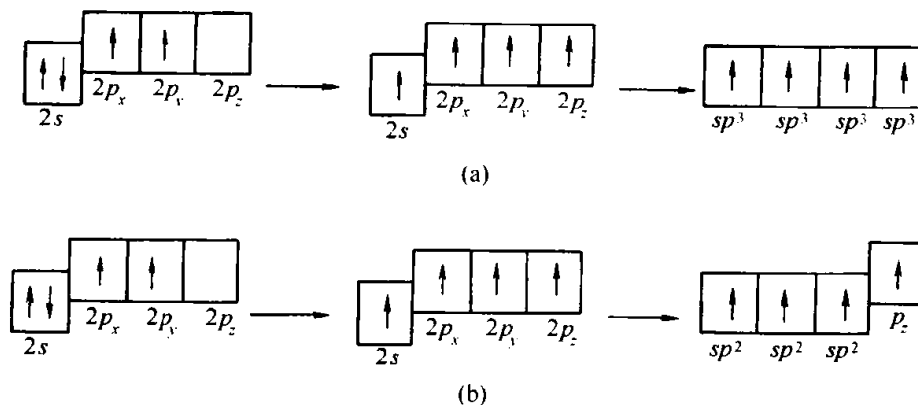


图 1-10 sp^3 与 sp^2 杂化轨道的形成

(a) sp^3 杂化轨道的形成; (b) sp^2 杂化轨道的形成

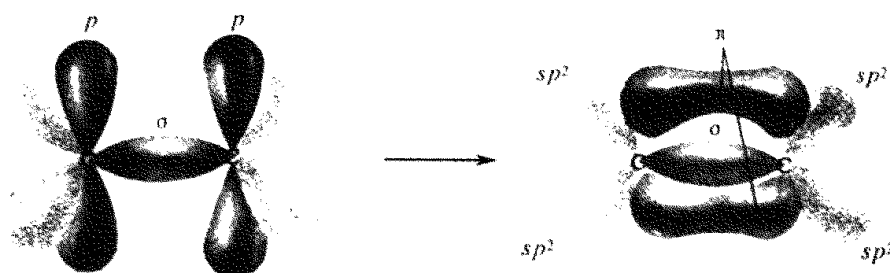


图 1-11 π 键的形成

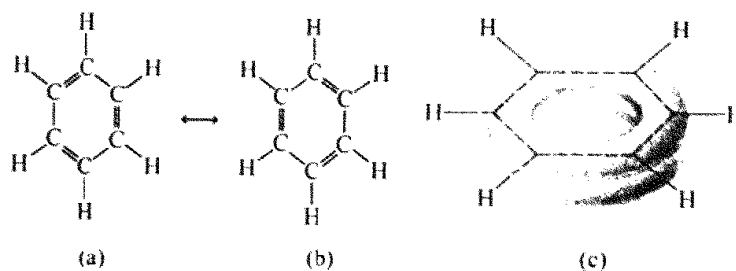


图 1-12 苯分子

(a)、(b) 等效的分子式; (c) π 电子云

聚合物和生物聚合物是更加复杂的有机分子, 我们将在有关章节中进行讨论。

1.1.5 金属键

金属键在本质上和共价键有类似的地方, 但是其外层电子比共价键更公有化, 电子自由游移于正离子之间, 遍及整个晶体, 构成近自由电子, 这就

像是正离子浸在近自由电子的海洋之中。金属键和共价键最明显的区别就是金属键缺乏方向性和饱和性。

金属键的外层电子的波函数互相重叠，这样就使孤立的能级逐步变成一个共有电子的能带。按照泡利不相容原理，能带从底部填起，一直填到费米能级为止。

粗略地来估计，金属的键合能主要有两项贡献（如图 1-13 所示）：其一

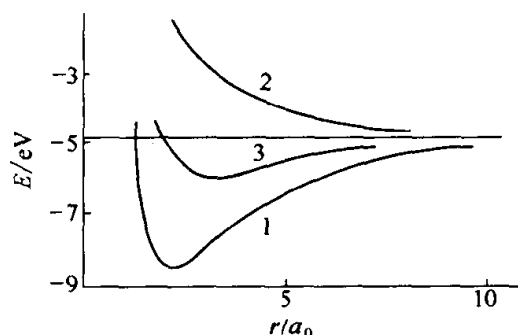


图 1-13 金属 Na 的点阵能量曲线

1— Na^+ 和自由电子的相互作用能；

2—电子的动能；3—总能

$a_0 = 0.0529\text{nm}$

是正离子与周围电子气的库仑吸引力；而另一是自由气的动能，将随浓度的增加而上升。

当然，要具体计算出某种金属的键合能是相当复杂的问题，涉及电子的能带结构，以及电子之多体系相互作用所起的修正。但基本上有规可循，可以进行从头算起的计算。

金属键有不同的类型。对于碱金属来说，金属性最明显，其离子相对而言较小；而过渡金属的离子则较大。

以电子密度的变化，可以将金属分为两种类型：丰满(full)型和开放(open)型。所谓丰满型金属是指金属离子之间的间距较小，而开放型的金属离子间的间距则较大。碱金属属于开放型金属，而过渡金属则接近于丰满型金属。显然，两者在性质上是有差异的。

以上所介绍的三种化学键：离子键、共价键以及金属键均属于强键。这三种强键的结合能的差别可参考相关资料。

1.1.6 弱键

一般而言，分子内起键合作用的是强键（主要是共价键），而分子之间的相互作用就弱得多了，可统称为弱键。有时键字还需打上引号，表明并非正规的键。

如果分子是非极性的，即不具有电偶极矩，那么是范德瓦尔斯力在起作用。不妨以最简单的单原子分子（如惰性元素分子）为例作一说明。这类原子的电子壳层正好填满。电子分布具有球对称性，因而并不存在电偶极矩。但由于电子在做不断的运动，导致偏离球对称分布的许多瞬间。这样一来，这些瞬间的电偶极子之间的相互作用构成了范德瓦尔斯吸引力的物理基础。如果两原子非常接近时，电子壳层间存在库仑力与泡利不相容原理导致的斥力又要占上风。因而可以用经验势函数来描述：

$$\Phi(r) = -\frac{A}{r^6} + \frac{B}{r^{12}}$$

被称为 Lennard-Jones 势。我们不妨比较一下两个 H 原子间和两个 He 原子间的最小势能，前者为 -4.6eV ，而后者仅为 -0.0008eV ($1\text{eV} = 1.602 \times 10^{-19}\text{J}$)，这表明范德瓦尔斯相互作用的确要比共价键要弱得多了。

至于极性分子之间，除了范德瓦尔斯力相互作用外，还要加上电偶极子之间的相互作用。特别值得注意的是分子之间可能存在氢键。考虑 H 原子与某些处于周期表右侧的非金属元素，如 F、O、N 等原子构成的分子，这样一来，构成共价键的电子对就会偏离 H 而更接近于另外的原子，其结果使分子具有强极性，H 带正电，而另外的原子带负电。以 HF 与 H_2O 分子为例（见图 1-14），可以看出，一个分子中 H 与另一分子中的 F 或 O 原子形成了以虚线表示的氢键。可以说氢键是弱键中的强者，但还要比共价键小一个数量级。单个氢键也许不是太起眼，而大群氢键的集体效应就会令人刮目相看了。氢键会导致分子液体中分子之间的缔合。水的某些性能异常（如沸点高达 373K ，熔点以上的密度异常）是和氢键有关，而氢键在生物分子结构中起重要作用，例如，DNA 与蛋白质的螺旋结构，就是靠一系列的氢键连结起来的。

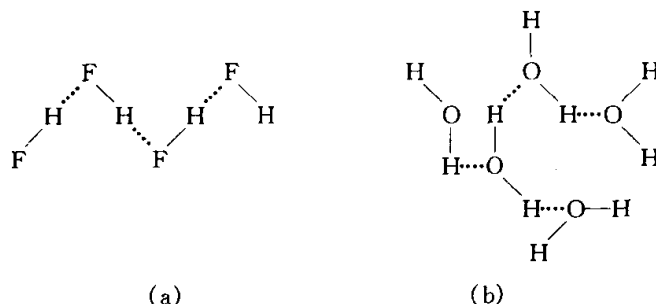


图 1-14 分子之间的氢键

(a) 氢氟酸 HF; (b) 水 H_2O

1.2 对称性

1.2.1 对称操作

人们对对称性的认识由来已久。我国西汉初期的古籍《韩诗外传》中就有耐人寻味的一段话：“凡草木花五出，独雪花六出。”似乎意识到雪花与草木花在对称性上的差异。如图 1-15 所示，雪花具有六重对称性，而草木花具有五重对称性。

在自然科学中首先重视对称性研究的是晶体学。到 19 世纪末，有关晶体对称性的理论业已完全确立。然后有关对称性的理论还延伸到物理性质与物理规律等方面去，成为当今自然科学的重要问题。

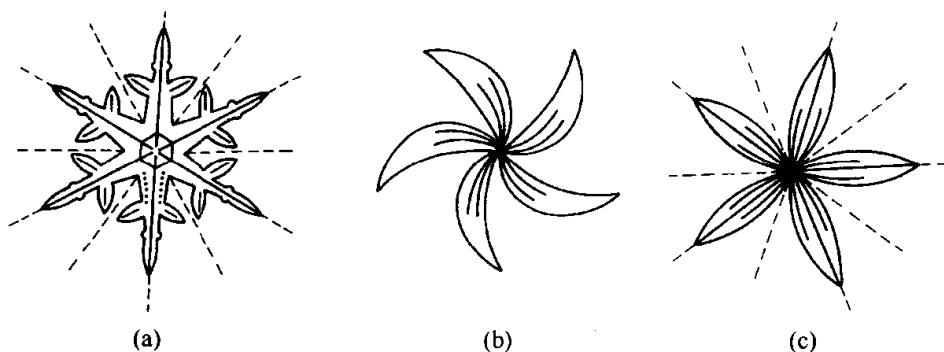


图 1-15 雪花与草木花

(a) 雪花; (b) 草木花; (c) 另一种草木花

那么, 什么是对称性? 这里需要做出更确切的解说: 如果将雪花沿其中轴旋转 $2\pi/6$ (即一圈的 $1/6$, 60°), 仍然和原图完全吻合; 对草木花情况相似, 只是将旋转角改为 $2\pi/5$ (即 72°)。从这里我们可以理解, 所谓对称性, 就在于再变换之中保持不变。在上述例子中, 变换就在于将图形沿中轴旋转一定角度, 而旋转后的图形与原图完全重合, 体现了其不变性。显然, 变化可以是多种多样的, 涉及数学上或物理上的坐标变换, 用科学术语来说, 就是对称操作。

除了上面提到的旋转对称操作 (2、3、4、5、6 重旋转轴, 相应于 $2\pi/2$ 、 $2\pi/3$ 、 $2\pi/4$ 、 $2\pi/5$ 、 $2\pi/6$ 的转角) 外, 还有反映对称操作 (mirror plane) 及反演对称操作 (inversion)。反映对称操作就是大家所熟悉的镜像反映, 它导致左右对称性。反映是相对于一个面 (镜面) 而言, 反演却是对一个点 (对称中心) 而言, 导致正向与反向对称性。这些对称操作可以组合构成一个新的对称操作。

1.2.2 对称与不对称

镜像反映, 例如, 对 $z=0$ 面镜像对称操作表明了物体的 $p(x, y, z)$ 变换为 $p'(x, y, z')$ 具有等同性。显然, 这种对称操作都不是能通过物体在空间中的现实运动来实现的。具有镜像对称的物体被称为对映体, 或称具有手性 (chirality) 的物体。左与右是相互对峙, 例如人的左手与右手, 左旋螺丝钉和右旋螺丝钉, 两者虽有相似性, 却在现实世界里无法实现它们之间的相互转变 (图 1-16)。这种左旋和右旋的手征性在生物学上十分重要, 很多生物分子就是左右不对称的, 而其旋光性也是不一样的, 如有些醌分子就有左旋右旋之分, 而 DNA 只有右旋的。

反演对称操作体现了正向与反向的等同性。反演的结果是将任意点的坐标 $p(x, y, z)$ 变为 $p'(\bar{x}, \bar{y}, \bar{z})$, (这里 \bar{x} 即 $-x$, 其余类推)。这一变换也可以分两步来实现, 即沿 z 轴旋转 180° , 使 $p(x, y, z) \rightarrow p'(\bar{x}, \bar{y}, z)$, 再对

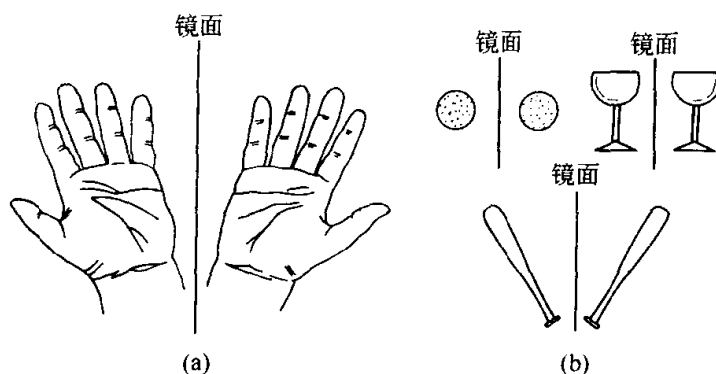


图 1-16 镜像反映

(a) 左右手; (b) 非手性物体

$z=0$ 处的镜面作反映, 令 $p'(\bar{x}, \bar{y}, z) \rightarrow p''(\bar{x}, \bar{y}, \bar{z})$ (见图 1-17)。这样, 空间反演之中就包含了镜像反映, 因而也是不能通过物体的直接运动所能实现的变换。

凡涉及镜像反映、空间反演及旋反轴的对称操作都是不能通过物体直接运动来实现的, 因而被称之为第二类对称操作, 以区别于可以通过物体直接运动来实现的第一类对称操作。

反演对称破缺时, 对称性被破坏, 具有不对称性。有偶极矩的分子必然不具有反演对称性。因此无论是对称或对称破缺 (欠缺了某一对称操作) 都具有重要意义。

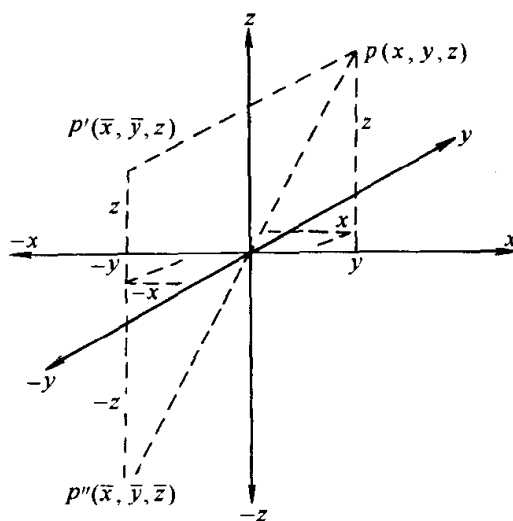


图 1-17 反演 = 180° 旋转 + 反映

1.2.3 对称元素的组合

对称元素的组合是受很大的制约的。下面介绍两个规律。

(1) 万花筒定理 和 n 重旋转轴共存的反映面不只有一个, 而是有 n 个。因为只要有一个镜面, 就可以通过 n 重旋转操作, 衍生出 n 个镜面; 而且由于镜面对穿中心, 相邻镜面的夹角就不是 $2\pi/n$, 而是 π/n 。反过来, 如果存在一对夹角为 π/n 的镜面 m 与 m' , 通过 m 与 m' 的相继反映, 就等效于作了 $2\pi/n$ 角或其整倍数的旋转, 这一关系清楚地显示于图 1-18 之中。

(2) 欧拉定理 两个相交的旋转轴必然引起第三个旋转轴。如果令两个相交的旋转轴为 n_1 及 n_2 , 被引发的第三个旋转轴为 n_3 , 转角分别为 $\alpha_1 =$

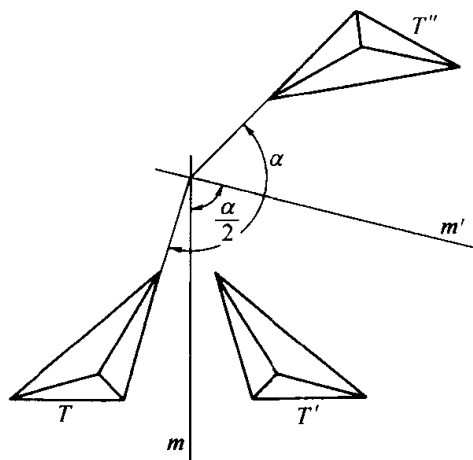


图 1-18 万花筒定理图

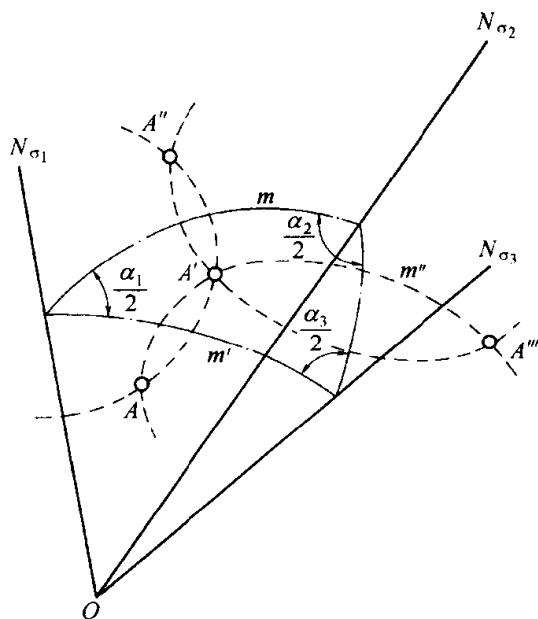


图 1-19 欧拉定理的证明

$2\pi/n_1$ 及 $\alpha_2=2\pi/n_2$, 被引发的第三个旋转轴为 n_3 , 转角为 $\alpha_3=2\pi/n_3$ (参看图 1-19)。这三个旋转轴 n_1, n_2, n_3 的交点与球面构成了一球面三角形, 由于球面三角形的三内角之和都大于 π , 因而存在一个不等式关系:

$$\frac{\pi}{n_1} + \frac{\pi}{n_2} + \frac{\pi}{n_3} > \pi$$

即:

$$\frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3} > 1 \quad (1-1)$$

有了上面两个规律, 就能得出对称元素的可能组合。对于某个特定物体来说, 其所有的对称操作正好构成一个闭合的群。群内的对称元素的操作不超出群的范围, 具有闭合性。群的元素可以是有限的, 从而形成一个离散群; 也可以是无限的, 构成一个连续群。

接下来考虑对称轴的组合问题: 由不等式 (1-1), 可以解决这个问题。如果 $n_1=2, n_2=2, n_3$ 就可以是任意的正整数, 没有丝毫限制, 即 $n_3=2, 3, 4, 5, 6, 7, 8, \dots$ 。具体的几何图像可以设想如下: 可以存在单一的 n 重轴 ($n \geq 2$), 和它正交的还有一组二重轴, 由于要求和 n 重轴相融洽, 相邻二重轴的夹角为 π/n 。由于这组二重轴都与 n 重轴正交, 二重轴的作用只能使 n 重轴反向, 不会偏离原来轴线的方向。这样的对称元素的组合, 可以构成封闭的体系, 满足群的要求。在这样的对称群中还可以插入镜面、对称中心, 或令旋反轴来取代旋转轴, 从而获得对称性更高的空间点群。

如果 n_1 、 n_2 、 n_3 中有两个高次轴 ($n > 2$)，则前面的不等式将对于可能的 n_1 、 n_2 、 n_3 值给予严格的制约。能够满足这一不等式的三个正整数的组合只有 233、432、235 三种。凡对称轴以 233 方式组合的，被称为四面体群；以 432 方式组合的，被称为八面体群；以 235 方式组合的，被称为二十面体群。

上面是以几何的观点来考虑对称元素的组合方式，而实际上，远古希腊的时候，人们就已发现了所有的正多面体——柏拉图体（如图 1-20）。所谓正多面体，顾名思义，是指那种多面体的每个面都是相同的，由边数为 p 的正多边形所构成；每个顶点也是等同的，都与 q 个正多边形相接，可用符号 $\{p, q\}$ 表示。正多面体总共只有五种：四面体（tetrahedron）， $\{3, 3\}$ ；六面体（hexahedron），称立方体（cube）， $\{4, 3\}$ ；八面体（octahedron）， $\{3, 4\}$ ；十二面体（dodecahedron）， $\{5, 3\}$ ；二十面体（icosahedron）， $\{3, 5\}$ 。这五种正

多面体中，四面体属于四面体群，立方体和八面体属于八面体群，而十二面体和二十面体则属于二十面体群。正多面体对结构的研究有很重要的帮助，

尤其是二十面体群具有五次对称性的特征很有意义：有些生物分子就是具有五次对称性，而近年来发现的具有截角二十面体形的碳分子 C_{60} 分子是特别值得注意的。

这里对 C_{60} 稍作介绍。首先我们必须对截角二十面体这个概念有个感性认识：将正二十面体的顶角全部削去即得到截角二十面体。截角二十面体更贴近球面，但不再是正多面体了。由于截去了 12 个顶角，出现了 12 个正五边形，

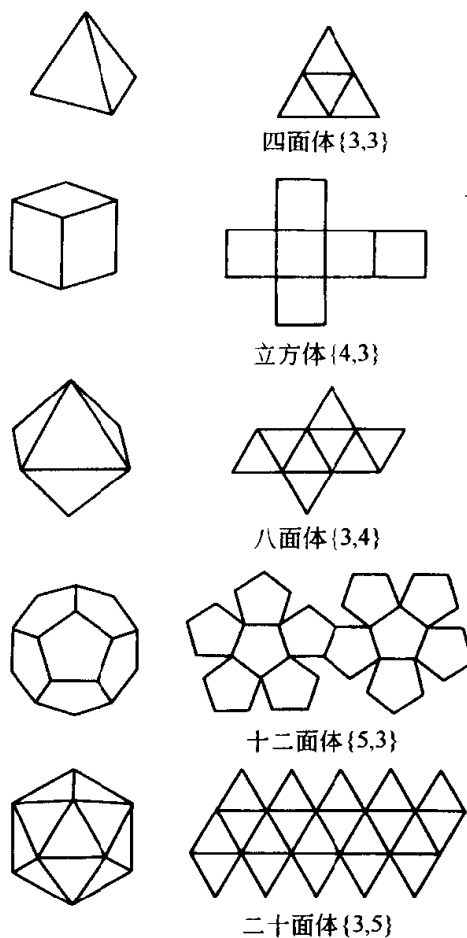


图 1-20 5 种柏拉图多面体及相应的折纸模型

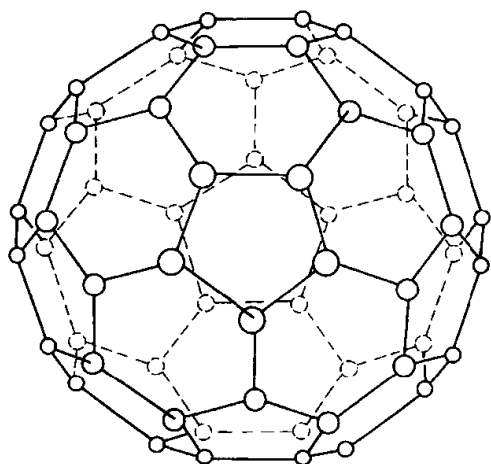


图 1-21 C_{60} 分子结构

而原先的 20 个三角形转变为 20 个正六边形，顶点总共 60 个，棱边共 90 根，仍然保留二十面体群的对称性（如图 1-21）。由于这种结构是由当代建筑奇才，R.Buckminster Fuller 首先提出来的建筑设想，因而发现这种分子的科学家将它命名为富勒烯（fullerene）。有关平移对称性和周期结构的问题，我们将在第三章中予以论述。

1.3 对称破缺——无序与有序

1.3.1 概念的引入

有序与无序这两个基本概念贯穿在物质结构的各个类型和层次之中。为了方便在概念上对有序与无序有个清楚的认识，我们不妨来看一个最简单的例子，即黑、白棋子在棋盘网格上的分布情况。

我们将黑、白两种棋子（各占一半）放在棋盘网格上，可以采用两种不同的方式来进行。一种是按精心设计好的方案，沿横行与竖行将白子与黑子彼此相间排列起来，如图 1-22(a)所示，黑白分明，井然有序；另一种是将白子与黑子掺和起来，盲目地抓一个子，放在网格的任意坐位上，直到填满为止，结果将如图 1-22(b)所示，黑白混淆，杂乱无章。这两种状态对照鲜明，判然有别，分别体现了有序态和无序态。在有序态图 1-22(a)，黑白棋子各有特定的坐位，在各自的坐位上，占有率分别为 100%；而在无序态图 1-22(b)，所有的棋盘坐位对于黑白棋子都一视同仁，毫无偏向。每个坐位上黑子与白子的占有率，就统计而言，都等于 50%，实际坐位为哪一种棋子所占，纯粹是随即行为的结果，并非于事先的策划。正因为彻底无序，统计的规律性就清楚地呈现出来。不妨设想一个黑白参半的棋子来代表黑、白子占有率分别为 50% 坐位，如图 1-22(c)所示的情况。

我们考察有序与无序两种图像的对称性问题。显然，图 1-22(a)具有严格的周期性，其横行与竖行的周期均等于网格间距的 2 倍。而图 1-22(b)中的图像则不复具有周期性了：从一个坐位到另一个坐位的平移，总不可能使图像复原，这表明严格的周期性业已丧失。但是如果我们将对称性的要求放宽，引入统计式的对称概念，即作对称操作之后，并不要求图像完全等同，而只要求大致相似，那么，将图像从一个坐位平移到另一坐位上，图像虽然不能全等，但棋子大致的无规分布情况大致相同，图 1-22(c)即可以作为反映了统计式的对称性的图像。这种无序图像具有统计式的周期性，棋盘的网格就是它的点阵，单元网格就是它的元胞。

从无序图像转变为有序图像，图像结构有变化，点阵类型也有变化。当然，还可以存在另一种黑白分明的有序排列，即黑白棋子分处棋盘两半，如图 1-22(d)所示。

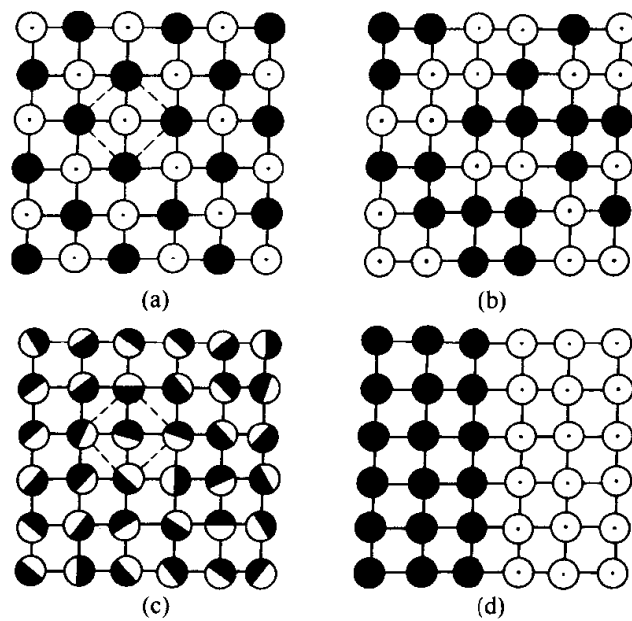


图 1-22 棋盘上黑白棋子排列的有序与无序
(a) 有序排列；(b) 无序排列；(c) 无序排列的统计表征；
(d) 另一种有序排列，两相分离

1.3.2 从概念到现实——有序-无序转变的实例

20 世纪 20~30 年代，科学家开始研究二元合金的有序-无序的转变。我们知道，完整晶体中每一晶格坐位上的原子都具有特定的品种，不可随意变更。假如我们意在保留晶体结构的骨架，但对晶格坐位上的原子以其他化学品种的原子来进行无规的替代，结果当然是破坏了晶体的严格的周期性，从而引入了替代无序。替代无序比较简单，物理图像清楚，所得结果对于更强烈的无序很有启发。

合金的结构有两类，一类是替代无序的，如通常的固溶体；另一类是替代有序的，如金属化合物。这两类合金俯拾皆是，但是观测替代无序到替代有序的转变，就需要选择合适的合金作为观测窗口。科学家们发现，CuZn 合金在 742K 和 AuCu₃ 合金在 665K 都存在有序-无序转变的迹象。CuZn 的无序相是体心立方结构，而有序相则为原点和体心位置分别为 Cu、Zn 原子所占的 CsCl 结构；AuCu₃ 的无序结构为面心立方结构，而有序相则是元胞原点为 Au 原子所占，3 个面心位置由 Cu 原子所占（参考图 1-23）。

X 射线衍射的实验结果证实了用图 1-22(c) 的模型来取代图 1-22(b) 无序结构的有效性。但是图 1-22(c) 的模型只是近似地成立，更精确的衍射实验，细致地测定衍射谱线之间的背景强度的分布，还会指出这一模型与实际无序结构之间存在的差异。

前面所讨论的合金的有序-无序的转变和原子在格点上的排列有关，而

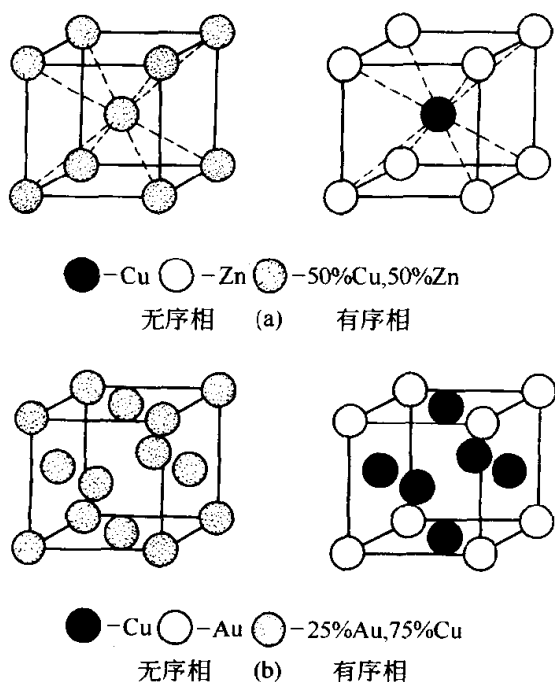


图 1-23 实际的有序-无序合金
(a) CuZn 合金; (b) AuCu₃ 合金

磁性的转变则涉及原子的自旋的排列的有序-无序转变，两者有不少相似之处。

法国物理学家皮埃尔·居里 (P. Curie) 在测量铁、镍、钴等物质的磁化率随温度变化的关系时，发现了临界温度 T_c ——这个温度被人们称为居里点。在高温下，原子自旋的取向完全无序，反映在磁性上表现为顺磁性 [参看图 1-24 (a)]；在居里点 T_c 以下，磁矩作顺向排列，就呈现了铁磁性 [图 1-24 (b)]。在 20 世纪 40 年代，法国科学家奈耳 (L. Néel) 发现另一种新磁有序结构，即磁矩作反平行排列的反铁磁性，对应的临界温度被称为奈耳点。由于自旋作反平行排列，反铁磁性物质虽然具有磁有序结构，但其宏观磁矩的总和为零。它不像铁磁性物质那样显示强的磁性，它的磁化率随温度变化的曲线的特色为在奈耳点为一尖峰，在图 1-24 (c) 上表现为一尖谷。后来又发现有些物质，虽然自旋作反平行排列，仍然可以具有强的磁性，只需反平行排列的自旋矢量大小不等即可。这类磁性被称为亚铁磁性，见图 1-24 (d)。铁氧体这类具有高电阻率的强磁性材料就具有亚铁磁性。而人类最早发现的吸铁石，即磁铁矿 Fe_3O_4 ，过去被误认为是一种铁磁性物质，事实上是亚铁磁性物质，也是铁氧体家族的一个成员。由于中子衍射对磁矩敏感，中子衍射就成为探测磁有序结构的强有力的实验方法，证实了过去间接地猜测出来的铁磁、反铁磁与亚铁磁等磁有序结构。

磁性转变是由磁矩之间的相互作用引起的，比二元合金的有序-无序转变要复杂。物理上经常采用一个同样的简化理论模型来处理千变万化的现象，这里我们介绍一个处理这类相互作用问题的最简单模型——伊辛 (Ising) 模型。伊辛模型认为自旋方向是单一的：或者向上，或者向下，而自旋之间的相互作用仅限于最近邻。这个假设虽然与磁矩的物理含义是有差异的，将磁矩的图像简化了，但是比较容易求解，因此这个模型可以用来解释磁性的相变，后来发现它也适用于合金的有序-无序转变，在凝聚态理论中扮演了相当重要的角色。一维伊辛模型能够求出严格解，得到的结果是一

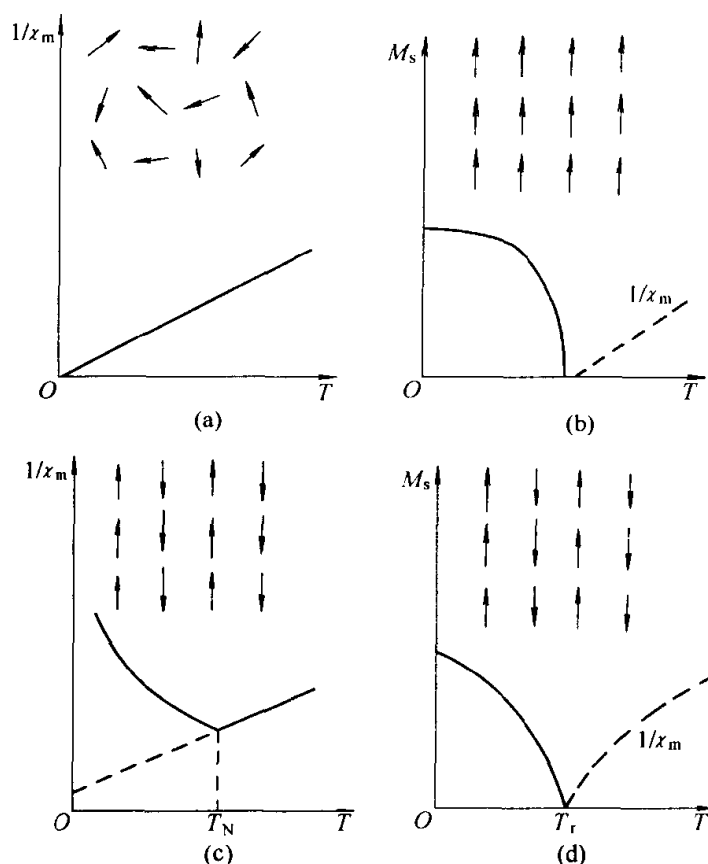


图 1-24 磁无序与有序结构的示意图与相应的磁化率 χ_m 及自发磁化强度 M_s 和温度的关系

- (a) 磁无序结构（顺磁相）；(b) 磁有序结构（铁磁相）；
(c) 磁有序结构（反铁磁相）；(d) 磁有序结构（亚铁磁相）

维情况下没有相变，即不发生磁性转变。但是三维的伊辛模型是无法严格求解的，在平均场近似下，才解得一个临界温度 T_c ，在 T_c 以上，是顺磁相；而在 T_c 以下，则为铁磁相。这表明近程的相互作用可导致长程的结构相。图 1-25 是伊辛模型及相应的物理系统。

在这里，我们探讨一下铁磁转变中的对称性的变化。在高温的顺磁相中，自旋取向是任意的，它的对称性是球面对称的（这是一种统计性的对称性），具有三维空间的全部旋转对称元素，当转变为铁磁相后，自旋具有特定的方向，也就一下子丧失了无穷多个对称元素。对于伊辛模型，在高温相，正向自旋和反向自旋同样地被激发，具有上下对称性。在 T_c 以下，单向的自旋占优势，就破坏了上下对称性。但上下对称性是离散的对称性，是和合金有序-无序转变中丧失某些特定的平移对称元素有相似之处，都是丧失了离散的对称性。

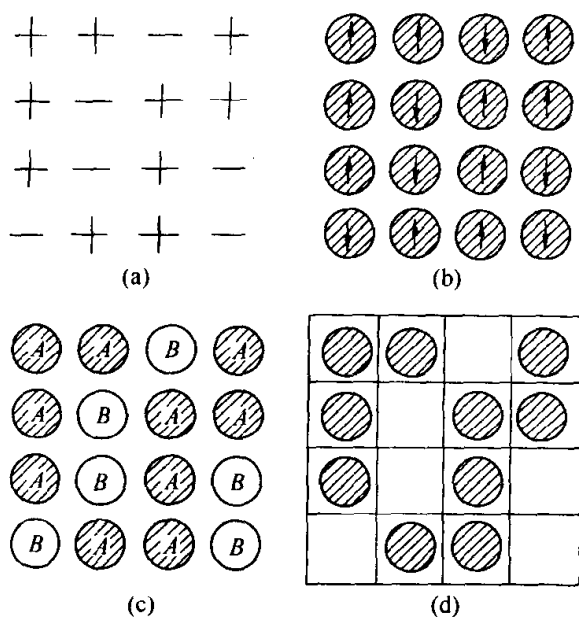


图 1-25 伊辛模型及相应的物理系统

(a) 伊辛模型; (b) 自旋系统;
(c) 二元合金; (d) 晶格气体

概括说来, 高温的无序相是高对称相 (虽然可能是就统计性的对称性而言的); 低温的有序相是低对称相, 相对于无序相丧失了一些对称元素, 可能包含连续对称元素或离散对称元素。而有序相的出现则是和丧失对称性相联系着的, 因此对称破缺导致了有序相的出现。前苏联物理学大师朗道 (L.D.Landau) 提出的对称破缺理论认为, 对称性的改变不可能是渐变式的, 因为对称性元素, 要么存在, 要么不存在, 其答案是不容模棱两可的, 因而对称破缺所导致的相变总是一种突变。这被另一位大师安德森 (P.W.Anderson) 称为是凝聚态

物理学的第一定理。对称破缺是凝聚态物理学最关键的基本概念之一, 不仅如此, 它在物理学的其他领域也影响深远。对称破缺有很多具体的事例, 甚至可以用来解释超导的转变——这种宏观的量子现象正是由于波函数位相的任意性的丧失而导致的。

1.3.3 能与熵的角逐——有序-无序转变的物理根源

上面我们对于有序-无序转变作了一个轮廓性的介绍, 要从物理上理解这一问题, 就需要求助于热力学和统计物理。在热力学中, 一个和周围环境处于热平衡状态的系统, 它的自由能 F 应为极小值。

自由能由两项构成:

$$F = U - ST$$

第一项是内能, 它是系统中各个原子的能量之和; 在第二项中, T 是绝对温度, 反映了系统中原子热运动猛烈的程度, 而 S 是熵, 是用来度量系统中混乱程度的物理量, 既无序度。按照玻耳兹曼 (L.Boltzmann) 著名的统计解释:

$$S = k_B \ln W$$

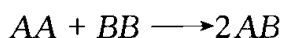
k_B 为玻耳兹曼常数, $k_B = 1.38 \times 10^{-23} \text{ J/K}$ 。 W 代表与一个宏观状态相对应的微观状态数, 而 $\ln W$ 为取 W 的自然对数。当合金作完全有序的排列, 对应于惟一的微观状态, 因而 $W = 1$, $S = 0$ 。合金作无序排列, 就有

难以胜数的不同微观排列状态与之对应，因而 S 很高。物质的平衡态就取决于能量和熵相互竞争的结果。

在有序-无序转变的问题上，一切和原子在格坐位上排列无关的能量可以完全忽略不计，只需讨论与 A 、 B 原子排列有关的能量和熵。进一步只考虑最近邻的原子之间存在相互作用。设最近邻为 AA 对、 BB 对和 AB 对，原子间相互作用能为 W_{AA} 、 W_{BB} 和 W_{AB} ，那么参量

$$W = W_{AB} - \frac{1}{2}(W_{AA} + W_{BB})$$

将对能量起关键作用。 W 的物理意义可以这样来理解：如果设法改变合金中的原子分布，使 AA 对中的一个 A 原子和 BB 对中的一个 B 原子交换位置，形成两个 AB 对，即



这一过程的能量变化就等于 $2W$ 。很显然， W 是决定合金中原子排列的基本能量参量。如果 $W < 0$ ，则异类原子作最近邻在能量上有利；反之， $W > 0$ ，则同类原子作最近邻在能量上占优。在 $T = 0$ 时，内能 U 即等于自由能 F 。如果 $W < 0$ ，合金将导致完全的有序相；反之 $W > 0$ ，则将分解成纯 A 相和纯 B 相，这就是相分离的现象出现。这样，能量的效果是在低温产生有序相，超结构和相分离是不同形式的有序相，对应于磁有序结构中的反铁磁性和铁磁性。

当 T 很大，即温度高，在自由能中举足轻重的就是 $(-TS)$ 这一项，熵愈大则自由能愈小，因而平衡相将是无序程度大的无序相。高温倾向于形成无序相，低温倾向于形成有序相，这是能量和熵相互竞争的结果。可以推断，存在一个临界温度 T_c ，在 T_c 以下是有序相，在 T_c 以上是无序相。 T_c 的具体数值是能量与熵相折中的结果，如果能量项起的作用大， T_c 就高。这里要特别说明的是，近年来所研究的软物质（如液晶，胶体等）中由于相互作用能甚小，可能导致熵致有序的出现（参看第 21 章）。

在二元合金中可能存在有两种情况：一种是完全有序的情况；另一种是完全混乱的情况，实际情况可能处在两极之间。因此，需要更加细致地来考虑坐位上的原子分布的统计规律。例如在二元合金中如果作为最近邻比同类原子要在能量上有利，将导致在异类原子近邻数比统计平均值增大，这样将导致短程序的出现。

要定量来描述这一现象，不妨来数一下系统中最近邻为 $A-B$ 型的总数 N_{AB} ，而每一个原子坐位有 Z 个最近邻，在一个有 N 个原子的系统中最近邻对的总数应为 $(1/2)ZN$ ，这样最近邻对为 $A-B$ 型的概率：

$$P_{AB} = \lim \frac{N_{AB}}{(1/2)ZN}$$

如果每一座位独立被占的几率分别为 C_A 或 C_B (A 、 B 分别为合金的原子成分), 那么这一概率应等于 $2C_A C_B$, 所以表征最近邻的关联性的近程序参量

$$\eta_s = \frac{1}{2} P_{AB} - C_A C_B$$

在具有长程有序的情况下, 两组格坐 α , β 就是不等价的。在完全有序时, α 座位为 A 原子占据, β 座位为 B 原子占据。如果偏离了完全有序, 一部分 A 原子坐对了, 另一部分坐“错”了。在 α 坐上坐对的分数为 r_α , 余可类推, 这样长程序参量 η_l 定义为

$$\eta_l = \frac{r_\alpha - C_A}{1 - C_A} = \frac{r_\beta - C_B}{1 - C_B}$$

$r_\alpha = 1$ 就表示完全有序, $\eta_l = 1$; 如果 $r_\alpha = C_A$ 即表示原子作无规分布, 长程序消失, $\eta_l = 0$ 。对于铁磁体系也可以作类似的表述, 自发磁化强度 M 可以作为长程序参量。

要具体计算出临界温度和序参量随温度变化的关系, 涉及复杂的理论问题, 这里不予进行深入的讨论。图 1-26 表示了关于铁磁相变的理论计算结果。在 $T = T_c$ 处, 长程序 η_l 连续地迅速下降为零, 但近程序 η_s 延伸到 T_c 以上的区域。这样的转变属于相变, 而这种相变可以连续地进行, 序参量连续地降为零, 或从零连续地上升, 因此属于连续相变, 也称二级相变 (序参量有突变的相变是一级相变)。

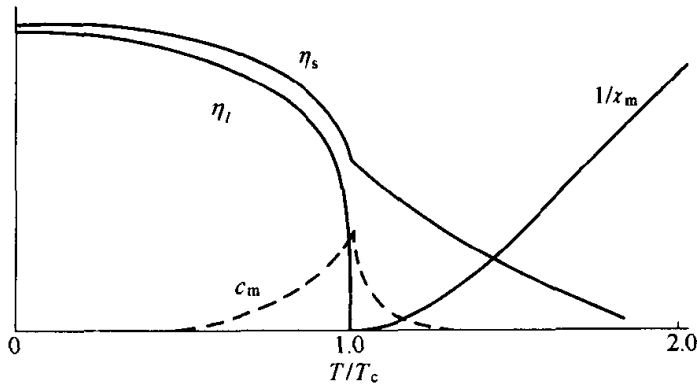


图 1-26 典型铁磁体中若干物理量随温度变化的关系

η_l —长程序参量; η_s —短程序参量; c_m —比热容; χ_m —磁化率

1.3.4 实际结构——有序相中的缺陷

媒质中序参量, 不仅会随温度而变化, 也可能随位置而变化, 后一特征往往出现在有序相之中。不妨以铁磁性媒体为例: 在居里温度以下, 作为序参量的自发磁化强度在空间分布上并不是均一的, 即使在热力学平衡态, 媒

质内会出现许多畴区，只有在畴区内部序参量是均一的，在分割畴区的畴壁上，序参量有突变，这就是序参量出现奇异性的场所，也就是缺陷。缺陷可以按其维数分为零维的点缺陷，如晶体中空位、填隙原子、杂质原子等；一维的位错对晶体的力学性质有重大影响。两维的如层错、畴壁与晶界等。有些缺陷在热平衡态即能存在，如晶体中空位，铁畴媒质中的畴界；有些缺陷只能存在于非平衡态，与材料所经历的生长过程或相变过程有关，如位错、层错、杂质原子等。由于各种缺陷存在，实际结构对材料性能影响甚大，因而我们将在第 4~6 章予以论述。

1.4 结构与信息

1.4.1 信息

过去人类使用的材料是以无机材料为主，如石头、青铜、铁等，一直到现在的钢铁与硅。有机材料是 20 世纪，在以高分子材料普遍使用以后，方始受到重视。当然从生物中取用的木材、蚕丝一直没有受到材料界的重视。现代材料日益向功能化、智能化的方向发展，这样一来，生物材料几乎提供了理想的模拟对象，材料仿生学应运而生。生物材料的特点在于它受到结构中蕴含的信息的控制，因而特别灵巧，是充分智能化的。因而阐述一下生物结构中信息及其复制、转录，乃至与功能的关系，将为更大规模利用生物材料或仿生材料提供理论的视野。

材料结构的各个层次上都蕴含了信息，在原子层次上，在分子层次上，乃至尺度更大的层次上……，无不如此。研究材料结构的科学家也习惯于将 X 射线、电子与中子衍射图中几何信息和强度分布的信息转换为晶体结构中各种原子位置的信息。上节我们讨论了有序与无序的问题。在有序态只要有一小组数字就可以明确地表明原子位置所有信息，而在无序态则无法一一标定原子的位置，因为哪一种原子处在哪一个格位上需庞大无比的数字（约 10^{24} ）。因而只好采取统计的方法来应付。而早在 1943 年著名物理学家薛定谔就在《生命是什么？》的报告中指出生命的奥秘在于非周期晶体中存在遗传密码，需要认真对待这一非周期性的信息，要从简单的统计理论脱颖而出。

如何科学地定义信息，1948 年 C. Shannon 提出了信息的统计理论：首先，考虑存在有 P 种可能性，其几率是相等的。例如，一个莫斯电码 $P=2$ ；一个拉丁字母的 $P=27$ （26 个字母加上一个空白）；一旦在 P 种可能性之中选定其一，我们就取得了信息。 P 愈大，相应地作出了选择之后的信息量也愈大，于是，信息 I 被定义为

$$I = K \log_e P$$

这里的 K 为比例常数。

由于相互独立的选择可能性（或概率）是相乘的，对应的信息量按此定义就具有相加性。

考虑一个信息量是一连串几个相互独立的选择的结果，其中每一个选择都是在 0 或 1 之间作出，因为总的 P 值应为 $P=2^n$ ，于是

$$I = K \log_e P = nK \log_e 2$$

如果令 I 与 n 等同，则

$$K = \frac{1}{\log_e 2} = \log_2 e$$

这样定出的信息量的单位，就是计算机科学中普遍使用的比特（bit）。如果令 K 等于玻耳兹曼常数 k_B ，那么信息量就用熵的单位来度量。在布里渊的理论中，信息相等于物理系统中总熵中的一个负值的量，转换成了负熵，即信息与负熵是等同的。

1.4.2 遗传信息的载体——DNA 分子

19 世纪，孟德尔总结出了遗传律，对自然科学产生了重大的影响。20 世纪，摩尔根提出生物遗传的基因学说，随后科学家们致力于寻找基因的物质载体。1944 年 O. Avery 发现了细菌转移现象，第一次直接证实了人们寻觅已久的控制生物遗传的物质基因，不是别的，正是脱氧核糖核酸（DNA）。即细胞核内 DNA 是遗传的物质基础，遗传信息就蕴藏在 DNA 分子结构里。

1953 年，J. D. Watson 和 F. H. C. Crick 确立了 DNA 分子的双螺旋结构（图 1-27），揭示了遗传信息极其复制规律。DNA 分子的基本结构单位是脱氧核苷酸。脱氧核苷酸含有碱基、磷酸和脱氧核糖，其中碱基有 4 种：腺嘌呤（A）、鸟嘌呤（G）、胞嘧啶（C）和胸腺嘧啶（T）。DNA 双螺旋结构主要由两条互补的多聚脱糖核苷酸链由氢键的作用配对在一起。碱基的配对是固定的：A-T 相配，G-C 相配（如图 1-28），因此 DNA 中 A 和 T，G 和 C 的数目是一致的，而 DNA 的碱基的序列就构成了遗传信息，它的不同排列则反映了各种生物遗传性的千差万别。

DNA 结构中存在 A，T，G，C 4 种碱基的序列。如果在这 4 种中任选两种来排列，共有 $4^2 = 16$ 种不同的排列；任选 3 种，则有 $4^3 = 64$ 中排列顺序。如果一条多核苷酸链上有 100 个碱基，那么则应有 4^{100} 种不同排列顺序，这个数目十分庞大，不仅远远超过了历史上所有物种的总数（约 4×10^9 ），而且也超过了太阳系中原子的总数。现在已知一个基因是 DNA 分子链上的一个区段，其平均尺寸约 1000 对碱基，对应地可能有 4^{1000} 种不同的排列顺序。如此巨大的信息量足以说明 DNA 结构具有充分的多样性，用来

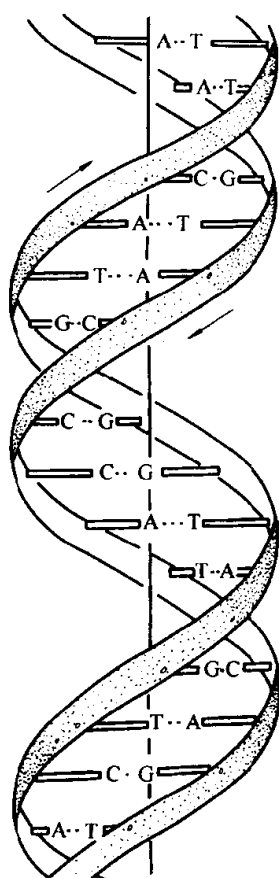


图 1-27 DNA 的双螺旋结构 (示意图)

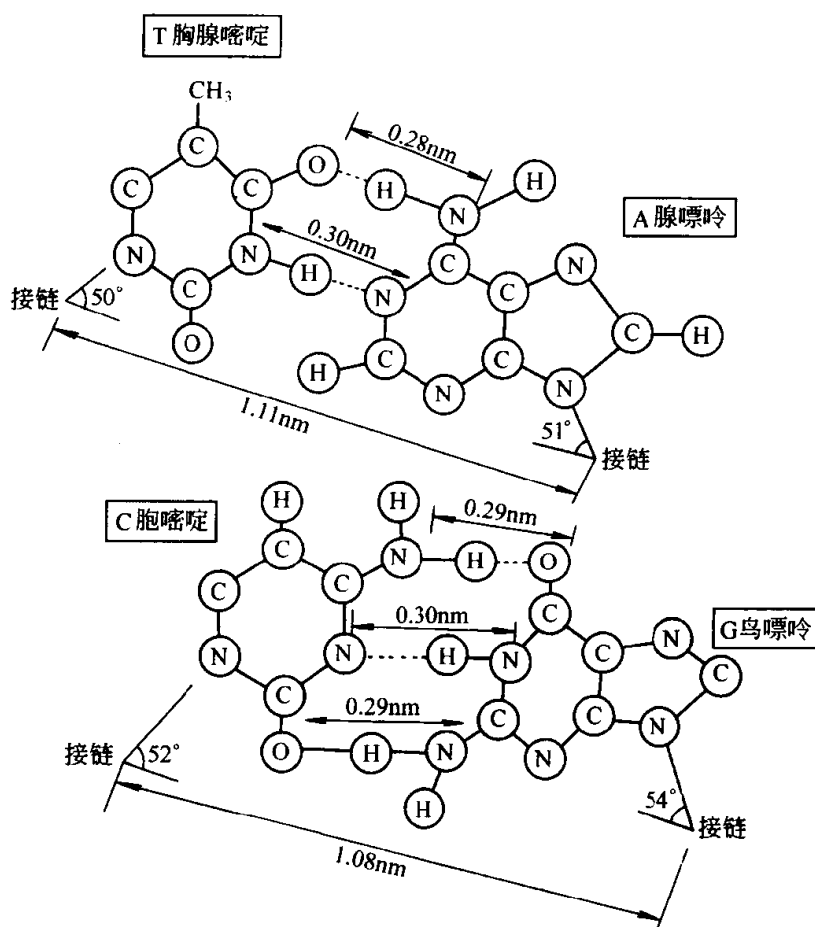


图 1-28 DNA 的 4 种碱基配对 (虚线为氢键)

解释物种的千差万别。记录一个大肠杆菌的碱基序列需要用一本千页的大书，而人体的所有碱基顺序数约为 2.9×10^9 。为了识别它们，科学界完成了一项庞大的人类基因组国际合作研究计划 (1991~2001 年)。对于 21 世纪生命科学的意义十分重大，可以比拟于 19 世纪化学家制出了元素周期表，促进了 20 世纪物质科学的发展。

生物特征的遗传，在分子水平上就是通过 DNA 复制来实现的：就是 DNA 双链松解下来，每一条链再与一条新链按碱基配对关系连接，结果相当于原子的双链衍生为两条等同的双链。由于碱基的顺序与原来的相同，实现了生物体中信息的复制 (见图 1-29)。这一过程说明信息序列在生物学中的重要意义。而且复制信息的能量消耗是非常低的。据估计复制一个比特的信息消耗的能量仅为 $100k_B T$ (k_B 为玻耳兹曼常数)，仅为当代最先进的微电子元件的百万分之一。而复制一个比特信息消耗能量的理论极限为 $k_B T$ 。

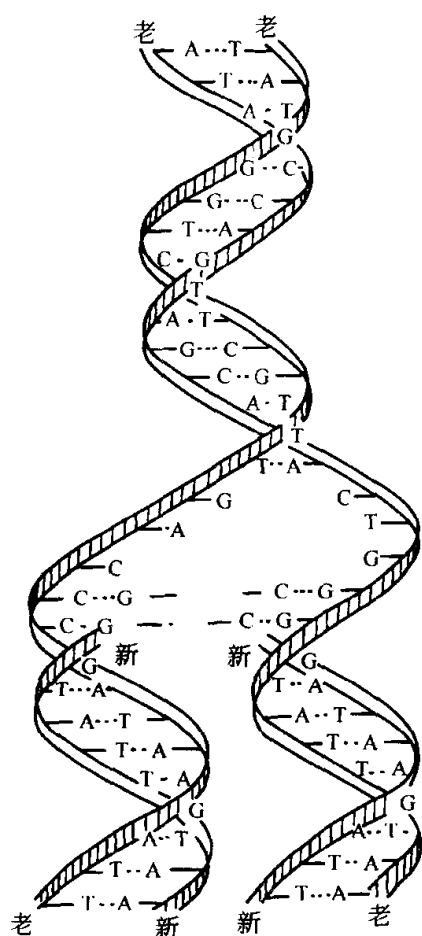


图 1-29 链上的碱基顺序
具有互补性质为前提的
DNA 复制方式示意图

1.4.3 蛋白质分子

生物的另一重要的分子是蛋白质分子。生物的所有功能都是通过蛋白质来实现的，因此蛋白质分子可以说是生物体的一个小工厂。

蛋白质分子异常复杂，比 DNA 分子复杂得多。最简单的蛋白质分子是肌红蛋白 (myoglobin)，它由许多股氨基酸的链折叠而成 (参看图 1-30)。不同的氨基酸序列导致蛋白质分子的不同折叠形态，从而得到不同蛋白质功能。

蛋白质和 DNA 表达内容用的“语言”不同：蛋白质用 20 个“字”，即 20 种氨基酸，而 DNA 用 4 个“字”，即 4 种核苷酸。DNA 既然要对蛋白质的氨基酸序列编程，必须找到方法用 4 种字的序列来表达 20 种氨基酸——编码。用一个核苷酸作密码，则 4 种核苷酸显然只能编码 4 个氨基酸。两个核苷酸组成密码，则 4 种核苷酸可以构成 16 个密码子 (codon)，仍然不足以编码蛋白质种的 20 种氨基酸。如果用所谓三联码 (triplet code)，则可以有 64 个密码子，超过了 20。DNA 的双螺旋结构发现以后，盖莫夫等物理学家立即提出了这种三联码。在 64 个密码子种，有 3 个终止密码子 (termination codon)，61 个密码子代表了 20 种氨基酸。因此大多数氨基酸有不只一个密码子，这就是密码子的简并 (degeneracy) 现象。几乎所有的生物都使用相同的密码子，这就是密码子的通用性。表 1-2 给出了通用密码子表。

遗传信息的表达首先是在细胞核内将它们转移到 mRNA 上，在 mRNA 中 T 由 U 来代替，这个过程称为转录 (transcription)。转录的产物还有 tRNA 和 rRNA。tRNA 能识别特定 mRNA 密码子，又能识别氨基酸。蛋白质的合成是在细胞质内由 mRNA 作模板，tRNA 负责运送氨基酸到核糖体，在核糖体上进行的。在核糖体中还发现了 rRNA。DNA 的两种催化功能和遗传信息的流向如图 1-31。

具有信息的分子在生命的演绎过程中起了关键性的作用，通过遗传密码

具有信息的分子在生命的演绎过程中起了关键性的作用，通过遗传密码

的转录，最后得到功能性的蛋白质分子。可见信息可以转换成功能，现代仿生学就是建立于这个基础上。这对材料科学来说既具有启发性，还具有挑战性。

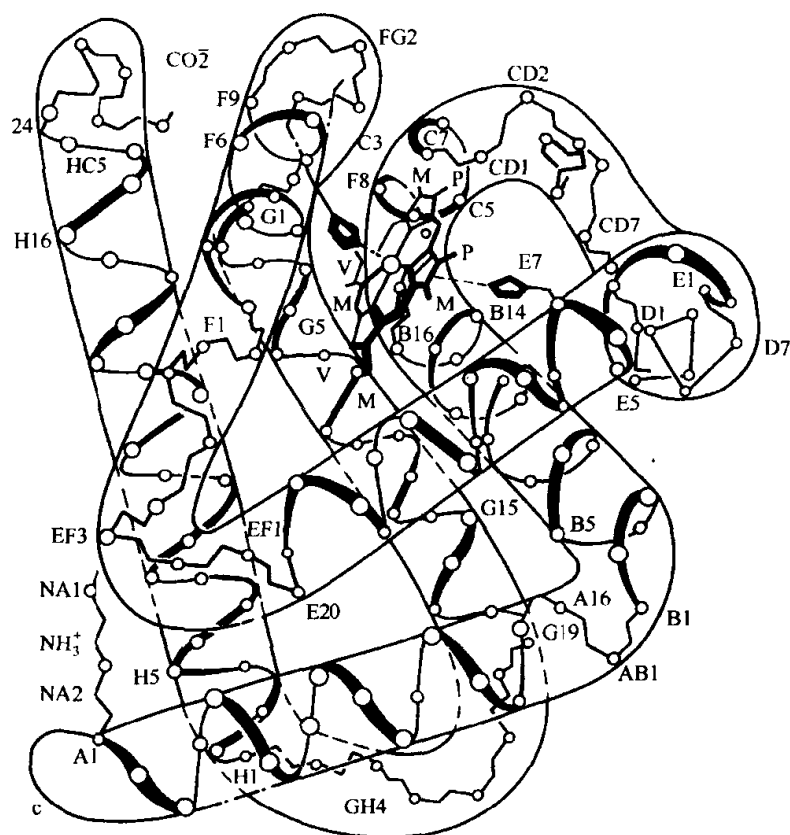


图 1-30 肌红蛋白分子的结构

表 1-2 通用密码子表

$\begin{matrix} y \\ x \end{matrix}$	U	C	A	G	z
U	Phe Phe Leu Leu	Ser Ser Ser Ser	Tyr Tyr Term Term	Cys Cys Term Trp	U C A G
C	Leu Leu Leu Leu	Pro Pro Pro Pro	His His Gln Gln	Arg Arg Arg Arg	U C A G
A	Ile Ile Ile Met	Thr Thr Thr Thr	Asn Asn Lys Lys	Ser Ser Arg Arg	U C A G
G	Val Val Val Val	Ala Ala Ala Ala	Asp Asp Glu Glu	Gly Gly Gly Gly	U C A G

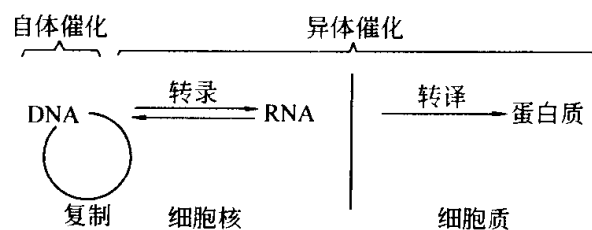


图 1-31 DNA 的两种催化功能和遗传信息的流向图

参 考 文 献

- 1 Allen M. and Thomas. E. The Structure of Materials. New York; John Wiley & Sons, Inc., 1998
- 2 Vainstein B. K. 等著, 吴自勤译. 现代晶体学. 第一卷及第二卷. 合肥: 中国科技大学出版社, 1992
- 3 冯端, 冯步云. 漫谈凝聚态物质. 第一册与第二册. 长沙: 湖南教育出版社, 1992
- 4 于渌, 郝柏林. 相变与临界现象. 北京: 科学出版社, 1984
- 5 Kittel C. Introduction to Solid State Physics. New York: John Wiley & Sons, Inc., 1997
- 6 Ball, P. Made to Measure, New Material for 21 Century. Princeton: Princeton University Press, 1997
- 7 冯端, 冯步云. 嫡. 北京: 科学出版社, 1992

第 2 章 晶 态

撰稿人 冯 端 凌惠琴

晶态是由于原子间相互作用导致的固体的热力学平衡状态。其特征就在于其中的原子的排列呈现三维周期结构。在材料中，金属与陶瓷通常是晶态，即使是高分子材料或生物材料也有相应的晶相。因而，对材料科学家来说，了解和掌握晶体结构的基本概念十分重要。

本章我们主要介绍晶体的周期结构、点阵、点群、空间群等基本概念，并举一些晶体结构的具体实例。最后介绍一点有关无公度调制结构和准晶的概念。

2.1 周期结构和点阵

规则的晶体多面体一直都以其美丽、晶莹和完美的表面形状令观察者着迷。17 世纪初开普勒就提出球形粒子密堆成雪晶的设想，随后矿物学家对晶体的多面体外形进行了详细的研究，揭示了晶面角守恒定律和晶面的有理指数定律。到 18 世纪末，Haüy 描述了如何从分子基块（平行六面体）构成各种晶体的直观设想（图 2-1）。

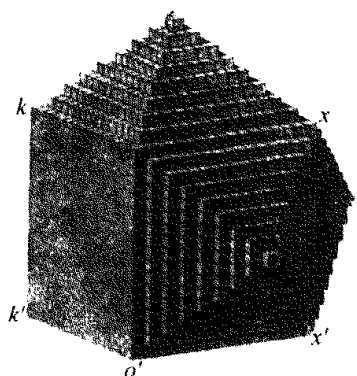


图 2-1 Haüy 设想的晶体形貌

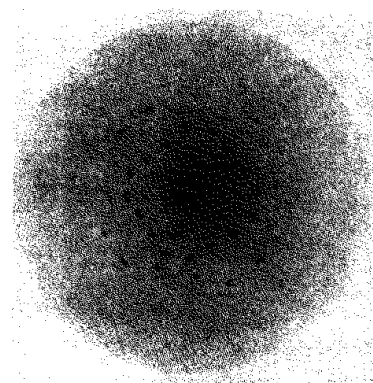


图 2-2 由 Laue 和 Friedrich 和 Knipping 得到的早期 X 射线衍射照片，闪锌矿的衍射花样反映了沿这个立方晶体的一个主轴的四次旋转对称

19 世纪出现了布喇菲的空间点阵学说。这一学说能解释有理指数定律和晶面角守恒定律，但它只是合理的猜想，其正确性到 1912 年才被劳厄等人的 X 射线衍射实验证实（图 2-2）。

几十年来,大量的研究探明了成千上万的晶体结构,肯定了晶体的周期性性质。现在用高分辨电子显微镜和扫描隧道显微镜等已能直接显示原子的周期排列图像,见图 2-3。

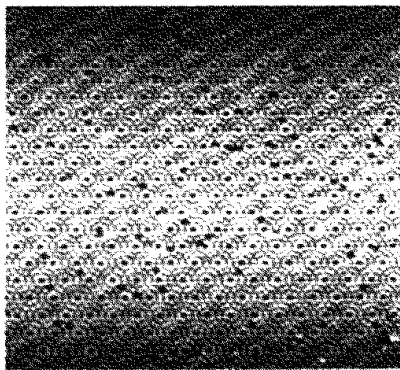


图 2-3 硅(111)表面 7×7 再构的扫描隧道显微镜像,显示了原子尺度上的周期性
(中国科技大学侯建国教授提供)

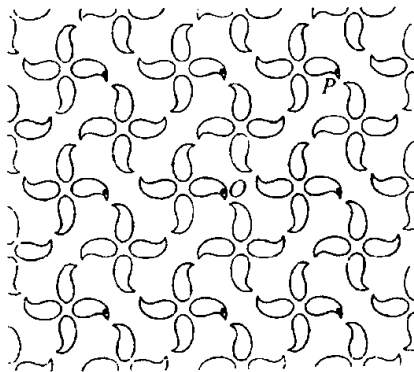


图 2-4 墙纸与其相应的点阵类型

2.1.1 点阵

点阵体现了晶体的平移对称性。我们可以通过墙纸来理解点阵的概念。图 2-4 是一幅典型的墙纸图像,在这里我们要运用一些想像力,忘却边框的存在,令图像延伸到无限远处。首先在图形上任意选定一点 O ,然后在图形上寻找与 O 点环境全同的其他点子,例如图中的 P 点或其他点子,这些环境相同的点子所构成的阵列就构成了墙纸的点阵。点阵是对实在图形进行抽象化的产物,用以说明其平移对称性。从 O 点到 P 点的平移矢量 t 被称为点阵平移。原图作了点阵平移后,将可以完全恢复。值得注意的是,平移对称性只适用于无限大的图形,有限图形的边框会导致图形不能完全恢复。从 O 点出发可以在非共线的两个方向求出最短平移矢量 a 与 b ,称为基矢,而任意的 t 可以表示为:

$$t = n_1 a + n_2 b \quad (n_1, n_2 \text{ 为整数}) \quad (2-1)$$

三维情况完全类似:

$$t = n_1 a + n_2 b + n_3 c \quad (n_1, n_2, n_3 \text{ 为整数}) \quad (2-2)$$

我们仍以图 2-4 为例来引出晶胞的概念。两个基矢所确定的平行四边形被称为单胞。应该注意,点阵原点的选择,基矢和相应单胞的选择都存在任意性,不是惟一的,容许有无限多的选择,如图 2-5 所示,图中 1、2、3 是最小的周期性单元,4 不是最小的周期性单元。但和图形相对应的点阵却是惟一的,全面地反映了图形的平移对称性或周期性。三维情况完全类似。

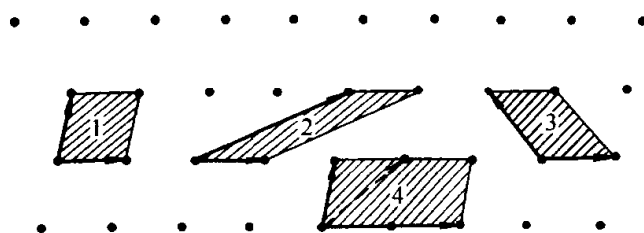


图 2-5 元胞示意图

为了反映对称性，单胞往往不是最小体积，晶胞中的阵点数可大于 1，如图 2-5 (4)。含有一个阵点的晶胞称为初基单胞，含有两个以上阵点的称为非初基晶胞。例如面心立方和体

心立方点阵的立方单元都不是最小的周期性单元，它们的初基晶胞如图 2-6 所示。可以证明含有 n 个阵点的晶胞体积为初基的 n 倍。如面心立方中，若 a 为立方单元的边长，则可以验证这个元胞的体积为 $a^3/4$ ，而立方单元中包含 4 个阵点（顶角处的阵点为共顶点的 8 个立方单元共有，面上的阵点为共面的两个立方单元所共有）。

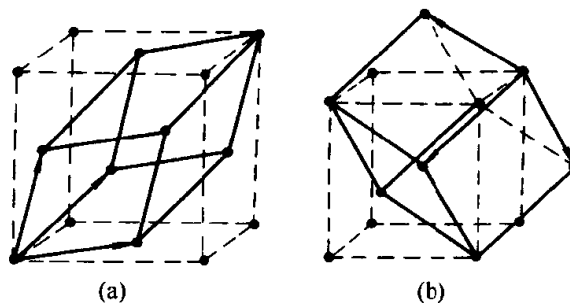


图 2-6 初基晶胞

如果赋予阵点以具体的内容，即它们是实际的原子或原子团，则整个晶体是晶胞在三维空间的重复堆砌。在这里，含有原子或原子团的晶胞已不是抽象的点阵，它相当于前面提到的“结构基石”。如果知道了晶胞的大小和形状，并且知道晶胞内部各原子的坐标，也就确定晶体的结构了。

2.1.2 点阵与点群的类型

点阵的类型是按点阵的对称性来分类的。早在 1848 年布喇菲 (A.Bravais) 就正确导出平面点阵只有 5 种，而空间点阵只有 14 种。如何来理解这一事实呢？需要对点阵可能有的对称性作一考查。点阵除了平移对称性外，还可能存在其他的对称性，如旋转、反映、反演等，乃至平移和反映等对称性的组合。

首先要明确一个问题：有些旋转轴是和周期性不相容的，因而可以排除。周期结构中只存在 $n=1, 2, 3, 4, 6$ 几种旋转轴，不存在 5 次及 6 次以上的旋转轴。这是周期性对旋转对称性的制约所致。图 2-7(b) 是五边形拼不出平面点阵的例子。在数学上很容易证明轴次限制的存在。

设想有一任意对称操作，转角为 θ ，在垂直于该转轴的晶面内，晶格常数为 a ，有一点 A, B ， $AB=a$ ，如绕 A 转 θ 角，则使 B 格点转到 B' ，由

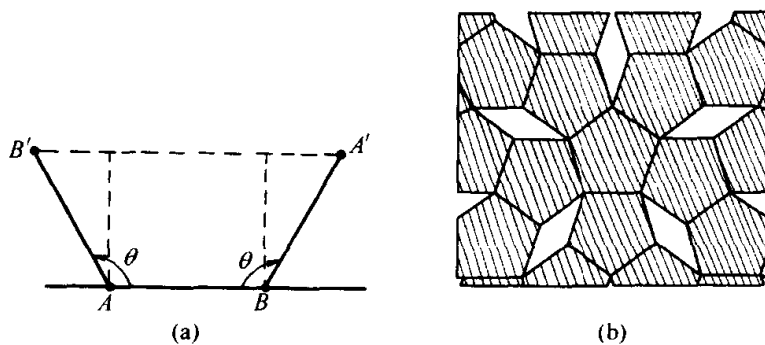


图 2-7 周期性对旋转轴次的制约

(a) 旋转变换示意图; (b) 正五边形不可能无隙缝地拼砌平面

于转动不改变格子, 在 B' 处必定原来就有一格点。因为 B 和 A 完全等价, 所以转动同样也可以绕 B 进行, 同样使 A 转到 A' , 所以 $B'A' = nAB$ [图 2-7(a)] 其中 n 为整数。根据图形的几何关系得:

$$B'A' = AB(1 - 2\cos\theta) \quad \text{或} \quad n = 1 - 2\cos\theta \quad (2-3)$$

因为 $\cos\theta$ 必须在 1 到 -1 之间, n 只能有 -1, 0, 1, 2, 3 五个值, 相应的 $\theta = 0^\circ, 60^\circ, 90^\circ, 120^\circ, 180^\circ$ 和周期结构相容, 其余均被排除在外。

有的点阵为了充分反映其旋转对称性, 在选取晶胞时不得不放弃初基的要求, 此时在面心或体心位置也可有阵点, 这些晶胞称为有心晶胞, 可由初基“加心”操作得到, 它们组成的点阵称为有心点阵。加心操作以不破坏点阵条件为原则, 要求保持晶系不变又能产生新的对称性与原点阵有区别。于是, 二维晶体有 5 种点阵类型(图 2-8)分属 4 个晶系, 三维晶体有 14 种点阵类型(表 2-1)分属 7 个晶系。

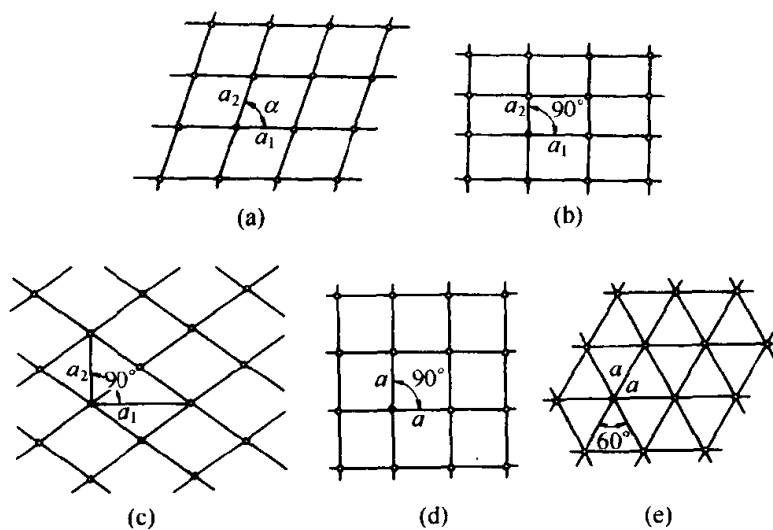
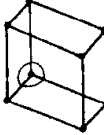
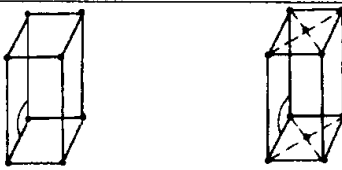
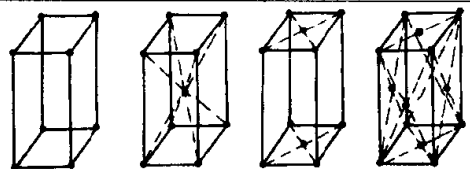
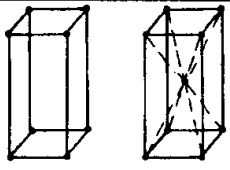

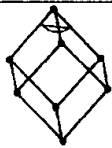
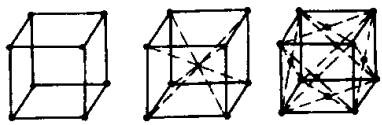


图 2-8 5 种平面点阵

表 2-1 晶系和点阵类型

晶 系	点 阵 类 型	相关的点群
	P (初基) I (体心) C (底心) F (面心) R (菱面体)	
三斜(triclinic) $a \neq b \neq c$ $\alpha \neq \beta \neq \gamma$		$1, \bar{1}$
单斜(monoclinic) $a \neq b \neq c$ $\alpha = \beta = 90^\circ \neq \gamma$		$2, m, \frac{2}{m}$
正交(orthorhombic) $a \neq b \neq c$ $\alpha = \beta = \gamma = 90^\circ$		$222, 2mm, \frac{2}{m}$ $\frac{2}{m} \frac{2}{m} \frac{2}{m} (mmm)$
四方(tetragonal) $a = b \neq c$ $\alpha = \beta = \gamma = 90^\circ$		$4, \bar{4}, \frac{4}{m} 422,$ $4mm, \bar{4}2m,$ $\frac{4}{m} \frac{2}{m} \frac{2}{m}$ $(4/mmm)$
六角(hexagonal) $a = b \neq c$ $\alpha = \beta = 90^\circ; \gamma = 120^\circ$		$6, \frac{3}{m}, \frac{6}{m},$ $622, 6mm,$ $\bar{6}2m, \frac{6}{m} \frac{2}{m} \frac{2}{m}$ $(6/mmm)$
菱面体(rhombohedral) $a = b = c$ $\alpha = \beta = \gamma$		$3, \bar{3}, 32, 3m,$ $\bar{3} \frac{2}{m} (\bar{3}m)$
立方(cubic) $a = b = c$ $\alpha = \beta = \gamma = 90^\circ$		$23, \frac{2}{m} \bar{3} (m\bar{3}),$ $432, \bar{4}3m,$ $\frac{4}{m} \bar{3} \frac{2}{m} (m\bar{3}m)$

点阵加心问题的关键在于加心之后点阵对称性有无变化。例如，二维矩形点阵加心之后，其初基晶胞具有菱形（两边相等）的特征。对称性确实有所改变，因而被认为一独立的点阵类型。与此相似的是三维点阵中面心立方，体心立方，体心四方等点阵。值得注意的是我们若在简单四方点阵的底面上加心，结果并不形成新的点阵类型，而是方向转过 45° 底面积缩为一半的简单四方点阵。另外菱面体点阵，实质上是种加心的六角点阵，加心的方式（见图 2-9）为在 $(2/3, 1/3, 1/3)$ 和 $(1/3, 2/3, 2/3)$ 处加上两个阵点，其初

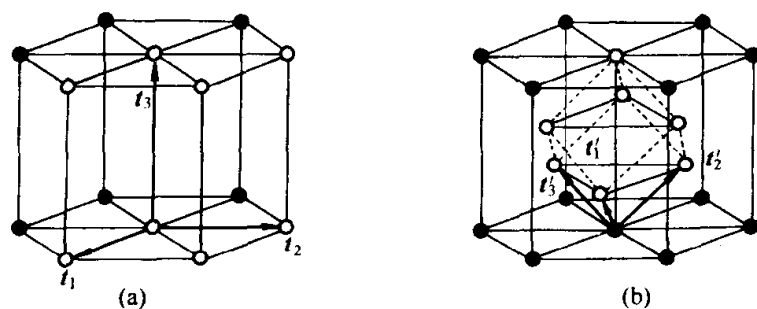


图 2-9

(a) 初基六角点阵；(b) 菱方体点阵（基矢为 t'_1, t'_2, t'_3 ）作为非初基六角点阵

基晶胞正是菱面体。

在国际符号中的点阵符号为：简单用 P ，面心用 F ，体心用 I ，底心用 B ，菱面体用 R 。立方晶系对称性最高，是高级晶系（有一个以上高次轴）；六方、四方、三方属中级晶系（只有一个高次轴）；正交、单斜、三斜属低级晶系（没有高次轴），三斜晶系对称性最低。由于菱形点阵可视为有心的六角点阵，因而有的科学家就将三方晶系归属为六角晶系，使晶系数目减为 6 种。同样二维晶体可分为斜方、矩形、正方、六角四种晶系。

晶体的宏观对称性又称为点对称性，因为宏观对称操作中空间至少有一点不动（点对称操作）。由于晶体具有周期结构，因而反映晶面外形的点对称性必然要受到周期条件的制约，使其旋转对称性也受到限制。

点对称元素有恒等（国际符号 1）、旋转（国际符号 n ，前面已证明轴次限制）、镜面（国际符号 m ）、对称中心（国际符号 $\bar{1}$ ）、反轴（先进行 n 次旋转，再进行中心对称操作，国际符号 \bar{n} ）。除去重复的元素，晶体的宏观对称元素只有以下 8 种是基本的，即 1, 2, 3, 4, 6, $\bar{1}$, m , $\bar{4}$ 。

8 种对称元素按照一定的规则组合构成了晶体的 32 种宏观对称类型，即 32 种结晶学点群，在表 2-1 中分列于相应的晶系之中。点群的国际符号将其各个轴的对称元素明确的表示出来，通常最多有三个位置，这三个位置的意义随所属晶系而有所不同，可以参阅表 2-1。

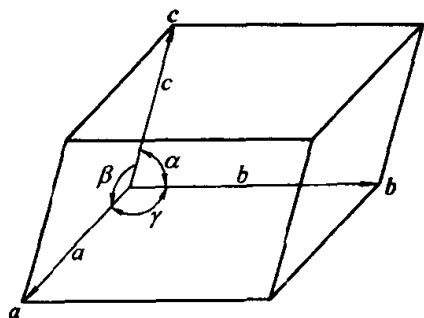


图 2-10 晶胞参数

晶胞是小平行六面体，可以用一组几何参数来描述，一般需要六个参量，即三个边长 a, b, c 及三个夹角 α, β, γ ，如图 2-10 所示，晶胞在三个方向上的长度矢量 a, b, c 就是点阵的基矢。通常称 a, b, c 及 α, β, γ 为点阵参数或晶胞参数，其中 a, b, c 又称为点阵常数。根据晶胞形状与点阵类型可将晶体分属为 7(或 6) 个晶系(参看 2.2.3 节

中的表 2-3)。

2.1.3 点阵的几何关系

下面讨论点阵、阵列和阵面的几何表示方式。设想以基矢 \mathbf{a} 、 \mathbf{b} 、 \mathbf{c} 确定坐标系 x 、 y 、 z ，指向任何阵点矢量为：

$$\mathbf{r} = u\mathbf{a} + v\mathbf{b} + w\mathbf{c} \quad (2-4)$$

这里的 u 、 v 、 w 都是整数。和此矢量平行的任何晶列就可以用指数 $[u \ v \ w]$ 来代表取向。如果存在由对称性相联系的一族晶向，就可以用 $\langle u \ v \ w \rangle$ 来代表。例如立方晶体中的 $\langle 110 \rangle$ 代表的是 $[110]$ 、 $[\bar{1}\bar{1}0]$ 、 $[101]$ 、 $[10\bar{1}]$ ……等的晶向族。

设想某一阵面和 x 、 y 、 z 三轴交截于 na/l 、 nb/k 、 nc/l 处，因而阵面的方程式可表示为：

$$h \frac{x}{a} + k \frac{y}{b} + l \frac{z}{c} = n \quad (2-5)$$

这里的 n 为整数， $n = 0, \pm 1, \pm 2, \pm 3 \dots$ 。当 $n = 0$ ，此面通过原点；当 $n = 1$ ，此面在 x 、 y 、 z 轴上的截距分别为 a/h 、 b/k 、 c/l 。这里的 $(h \ k \ l)$ 为一组不大的整数（消去了它们的公约数），被称为阵面（或晶面）的密勒指数 (Miller indices)。要注意 h 、 k 、 l 的值是截距的倒数，可以是 0，表示阵面和该轴向平行；也可以是负值，数值上加横杠。例如 $(h \ k \ l)$ 与 $(\bar{h} \ \bar{k} \ \bar{l})$ 为同一组阵面，其法线方向正好相反。在一晶体中由对称性相联系的阵面族用 $\{h \ k \ l\}$ 来表示。例如立方晶体的 $\{0 \ 2 \ 1\}$ ，则包括 $(0 \ 2 \ 1)$ 、 $(0 \ \bar{2} \ 1)$ ……等阵面。为了突出晶面之间或晶向之间的对称性关系，六角晶向常采用四指数的表示方法：坐标轴取为 \mathbf{a} 、 \mathbf{b} 、 $-(\mathbf{a} + \mathbf{b})$ 、 \mathbf{c} ，晶向指数表示为 $[u \ v \ t \ w]$ ，而晶面指数表示为 $(h \ k \ i \ l)$ ，显然 $t = -(u + v)$ ， $i = -(h + k)$ 都不是一个独立的指数。这种表示法的好处在于突出了有对称性相联系的晶向族。用密勒指数 $(h \ k \ l)$ 所表示的实际上是一组间距为 d 的平行无限多阵面组成的阵面族，这些阵面的法线方向完全相同，因而可以用其法线方向矢量 \mathbf{G} 来代表(图 2-11)，而 \mathbf{G} 的大小为正比于面间距的倒数，即 $1/d$ 或 $2\pi/d$ 。矢量 \mathbf{G} 被称为倒易矢量。密勒指数 $(h \ k \ l)$ 是没有公约数的，如果将一整数 n 乘上 h 、 k 、 l ，就使矢量 \mathbf{G} 增长 n 倍，而此阵面族的面间距则由原来的 d 缩为 d/n 。这样一来每个 \mathbf{G} 矢量都可以延伸

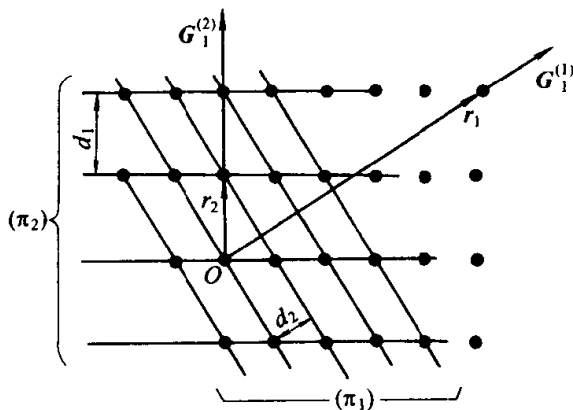


图 2-11 标志格点平面族的晶面基矢 \mathbf{G}

为一阵列，而所有的 G 矢量对应的阵点构成了倒易点阵。倒易点阵在衍射物理和固体物理中起了重要作用。两个阵面 $(h k l)$ 与 $(h' k' l')$ 相交截于一阵列 $[u v w]$ ，称为晶带轴，含有晶带轴的任意阵面 $(h k l)$ 应满足晶带定律：

$$hu + kv + lw = 0 \quad (2-6)$$

2.1.4 对称性对物性常量的制约

材料的许多物理性质表明了材料所受刺激与响应之间的关系，而这些刺激，响应和它们之间的比例常数可用矢量和张量来表示。例如对导电材料加电场强度矢量 E ，引起电流密度矢量 J ，它们之间的比例常数为电导率 σ ，可用下式来表示：

$$J_i = \sum_{j=1}^3 \sigma_{ij} E_j \quad (2-7)$$

电导率这一物性张量是二阶张量，将上式具体写出来，即为：

$$\begin{aligned} J_1 &= \sigma_{11} E_1 + \sigma_{12} E_2 + \sigma_{13} E_3 \\ J_2 &= \sigma_{21} E_1 + \sigma_{22} E_2 + \sigma_{23} E_3 \\ J_3 &= \sigma_{31} E_1 + \sigma_{32} E_2 + \sigma_{33} E_3 \end{aligned} \quad (2-8)$$

这些方程也可以写成矩阵形式：

$$\begin{pmatrix} J \\ J_2 \\ J_3 \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix} \quad (2-9)$$

物性张量是晶体本身所固有的属性，如介电常数、极化率为二阶张量，压电系数、非线性光学系数是三阶张量。

由于物性张量是晶体的属性，因此用物性张量描述的物理性质的对称性必然与结构对称性之间存在一定的关系，这二者的关系服从诺埃曼原理：晶体的任一物理性质所拥有的对称要素必须包含晶体所属点群的对称要素。这个原理有两层含义。第一是指二者对称要素的种类和数量方面的包含，即

- ① 晶体的物理性质可以而且经常具有比晶体所属点群更高的对称性。
- ② 晶体物理性质的对称性不能低于晶体所属点群的对称性。

另一层含义是指有些对称要素的取向方面的包含。

一个二阶张量的非对角分量若存在 $T_{ij} = T_{ji}$ 关系，此张量就是对称的。因而一般二阶物性张量是对称的，只有 6 个独立分量。而且当 q_j 与 p_i 换向时， T_{ij} 各分量保持不变，即存在：

$$p_i = \sum_{j=1}^3 T_{ij} q_j \quad (2-10)$$

这一普适关系。因而所有材料的二阶物性张量都是中心对称的，即使有些晶

体本身没有对称中心。这吻合于诺埃曼原理，即物性张量的对称性等于或高于晶体点群的对称性。这一结果可以推广到偶数阶的物性张量都是中心对称的，与晶体的所属点群有无对称中心无关，即 32 种点群都可能具有偶数阶张量描述的物理性质。比如电极化率二阶张量不仅存在于 11 种有对称中心的晶体中，在 21 种无对称中心的晶体中也都存在。

经过大量的实验证实诺埃曼原理是完全正确的，它是讨论晶体结构的宏观对称性与物理性质对称性的关系的基本准则。由诺埃曼原理可以推测晶体物理性质：

① 只有极性晶类才可能有一阶张量描述的物性常数。10 种极性晶类的结构特点是具有特殊方向的极轴，根据诺埃曼原理，物性常数中只有矢量（一阶张量）才可能与这个特殊极轴方向一致。例如具有自发极化的铁电晶体一定是极性晶体。

② 凡是具有中心对称的晶体都不可能具有奇数阶张量描述的物理性质。

假设三阶张量具有对称中心，由于对称中心的坐标变换矩阵为：

$$(\alpha_{ij}) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (2-11)$$

将其作用在三阶张量上，根据三阶张量变换法则，变换前后所有对应的分量应相等，由此得到，三阶张量所有分量都为零，这说明三阶张量乃至所有奇数阶张量都不具有中心对称性。由诺埃曼原理，物理性质中没有的对称要素，结构中肯定没有，所以凡是具有中心对称的晶体都不可能具有奇数阶张量描述的物理性质。因此晶体是否有对称中心可用压电效应或非线性光学效应来鉴定。

2.2 空间群

2.2.1 晶体的微观对称性

晶格的周期性，也称平移对称性，平移对称性是最基本的微观对称性。由平移对称操作得到的对称群称为平移群。前面讨论的 4 种布喇菲点阵就是 14 种平移群。

晶体的点对称元素和平移相结合能产生新的对称元素。如旋转轴 + 平移 → 螺旋轴；镜面 + 平移 → 滑移面，操作顺序并不重要。

(1) 螺旋轴 螺旋轴国际符号用 n_m 来表示。一般定义 τ 方向的 n_m 次螺旋轴对称操作由 $\frac{2\pi}{n}$ 旋转和 $\frac{m}{n}\tau$ 平移组成，其中 τ 是基本平移矢量。重复这样的操作 n 次，晶胞中的一点可螺旋前进到相邻第 m 个晶胞中的等效点位

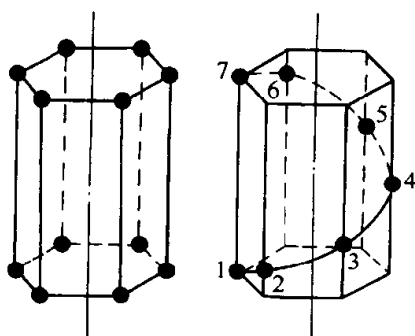


图 2-12 六重轴和六重螺旋轴的比较
(右图经旋转 $2\pi/6$ 后再平移 $a/6$)

置。螺旋轴的示意图如图 2-12 右图。

(2) 滑移面 滑移面是对镜面反映后再沿某一方向平移，操作的顺序并不重要，平移量为点阵周期的一个分数距离。图 2-13 所示具有典型的滑移对称性。

有三种类型的滑移面：

轴向滑移——沿 a , b , c 轴作滑移，轴向滑移的平移平行于镜面，平移量为该方向平移周期的一半；

n 滑移——沿面对角线滑移到一半处；

d 滑移——亦称“金刚石”滑移，系沿体对角线滑移到 $1/4$ 处。

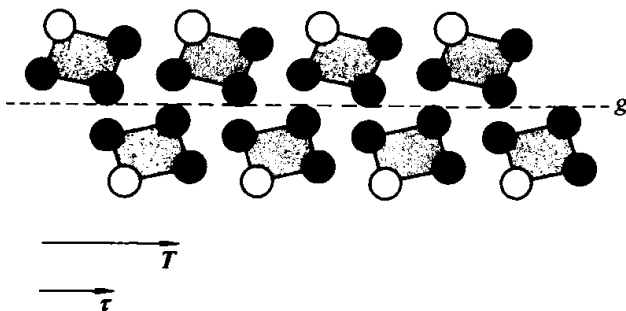


图 2-13 滑移反映面的示图
 T 为点阵矢量； τ 为滑移矢量； g 为滑移反映面

2.2.2 空间群

空间群能提供晶体的全部对称性信息。知道了空间群和每个晶胞中含有几个化学式单位，常常可以推测出原子在晶胞中的位置。通过空间群还有助于了解晶体结构的变化如相变等。特别是《结晶学国际表》的建立，知道了空间群，通过查表我们即可获得该晶体的全部对称性资料。

空间群涉及到一个给定的点群、布喇菲点阵以及这个点群作用在这个点阵上的结果。将各种点阵与其相容的点群一一予以考虑，可以推导出 73 种点式空间群。考虑了滑移面与螺旋轴这两类对称元素，再推导出 157 种非点式空间群，使空间群的总数到达 230 种。

这个工作是俄国的弗德洛夫与德国的熊夫利于 1891 年左右分别独立完成的。弗德洛夫是晶体学家，在他的理论中，多面体的空间拼砌起了关键性的作用；而熊夫利是数学家，数学论证更严密，常为后人所引用。230 个空间群的导出标志了经典晶体学的顶峰，成为固体理论的一个支柱。

和三维空间群类似，二维空间群也有点式和非点式两类。5 种二维点阵

与二维点群直接组合成 13 种点式空间群，再考虑到周期结构中还可含有平移的对称操作，又可导出 4 个平面群，这样共有 17 种平面群，包括了二维周期结构所有可能的对称类型，见图 2-14。

对于二维周期结构来说，共有 10 种点群，5 种是具有单一旋转轴的，另外 5 种是和镜面相结合的（表 2-2）。至于三维空间群表就太繁复了，本书就不一一列出，请读者参看“International Table for Crystallography”一书。

2.2.3 空间群的国际符号

空间群有两种常用的表示符号。较早的是熊夫利，而材料学界常用的是国际符号，这里我们仅介绍国际符号。

国际符号的第一位符列出布喇菲点阵类型—— P （初基）， A 、 B 或 C （底心）， I （体心）， F （面心）， R （菱形）；然后根据对称元素对于晶体学轴的位置列出它们的符号，符号位置所代表的轴向对不同的晶系并不相同。表 2-3 列出了空间群国际符号的顺序。

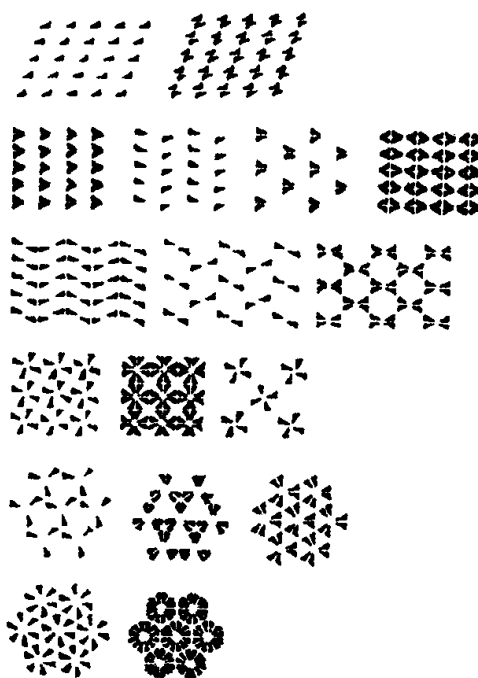


图 2-14 17 种平面群对应的图像

表 2-2 二维布喇菲点阵、点群和空间群表

二 维 空 间 群		二维点群	晶系和布喇菲点阵类型
序 号	国际符号	国际符号	
1	$p1$	1	斜形 ($a \neq b, \gamma$ 任意) p
2	$p2$	2	
3	pm	m	矩形 ($a \neq b, \gamma = \pi/2$) p 及 c (有心)
4	pg		
5	cm		
6	$p2mm$	$2mm$	
7	$p2ng$		
8	$p2gg$		
9	$c2mm$		
10	$p4$	4	正方 ($a = b, \gamma = \pi/2$) p
11	$p4mm$	$4mm$	
12	$p4gm$		

续表

二 维 空 间 群		二维点群	晶系和布喇菲点阵类型
序 号	国际符号	国际符号	
13	$p3$	3	六角 ($a = b, \gamma = \pi 2/3$), p
14	$p3m1$	$3m$	
15	$p31m$		
16	$p6$	6	
17	$p6mm$	$6mm$	

表 2-3 国际符号的顺序

	点阵类型	位置 1	位置 2	位置 3
三斜	P	✓		
单斜	P	✓		
	$B(或 C)$	✓		
正交	P	a	b	c
	C	a	b	c
	I	a	b	c
	F	a	b	c
四方	P	c	$a+b$	$a-b$
	I	c	$a+b$	$a-b$
六角	P	c	a, b, \cdots	$2a+b, \cdots$
	R	c	a, b, \cdots	$2a+b, \cdots$
立方	P	a, b, c	$a+b+c, \cdots$	$(a+b), (a-b) \cdots$
	I	a, b, c	$a+b+c, \cdots$	$(a+b), (a-b) \cdots$
	F	a, b, c	$a+b+c, \cdots$	$(a+b), (a-b) \cdots$

举例来说， $P2/m$ （2 在 m 上）表示单斜初基点阵，具有垂直于镜面的二次旋转轴。值得注意的是国际符号的简略形式，如 $Pnma$ 表示垂直于三个正交晶轴分别有 n 滑移面、 m 镜面和 a 滑移面。根据这些对称元素的组合，在三个方向上必产生三个 2_1 螺旋轴，因此它的完整符号是 $P \frac{2_1}{m} \frac{2_1}{m} \frac{2_1}{m}$ 。

2.2.4 晶体结构的描述

材料科学家往往需要对有关材料的晶体结构有完整的了解。这就要掌握

晶格参量，以及晶胞内每一个原子的品种与坐标。许多常见的晶体结构有资料可查，但是一些已测定的晶体结构的结果一般都用简略的语言（例如 Wyckoff 符号）来描述。要使这些结果具体化，往往需要查阅 International Table for Crystallography vol. I 中有关空间群的图页。现在就以 LiNbO_3 晶体为例作一说明。结构测定的资料如下：

化学式： LiNbO_3 ；

点阵：菱方体，如采用六角晶胞 $a = 0.5148\text{nm}$ ， $c = 1.3863\text{nm}$ ；

晶胞中化学式数：6；

空间群： $R3c$ (No.161)；

结构分析结果表明：Nb 处于 $6(a)$ 位置： $0, 0, w_1$ ， $w_1 = 0.0186$ ；Li 处于另一组 $6(a)$ 位置： $0, 0, 1/3 - w_2$ ， $w_2 = -0.0318$ ；O 处于 $18(b)$ 位置： x, y, z ， $x = 0.0492$ ， $y = 0.3446$ ， $z = 0.0833$ 。有了这些资料，我们就根据晶体学国际表上 $R3c$ 空间群这一页选择六角坐标系，查出等效位置。

格点位置为 $(0, 0, 0; \frac{1}{3}, \frac{2}{3}, \frac{2}{3}; \frac{2}{3}, \frac{1}{3}, \frac{1}{3}) +$

Wyckoff 符号： a 6 个位置： $0, 0, z; 0, 0, \frac{1}{2} + z;$

b 18 个位置： $x, y, z; \bar{y}, x - y, z; y - x, \bar{x}, z;$

$\bar{y}, \bar{x}, \frac{1}{2} + z; x, x - y, \frac{1}{2} + z; y -$

$x, y, \frac{1}{2} + z。$

将这里的 a 、 b 位置分别加上三种格点位置，从而将 Nb 与 Li（占 a 位置）及 O（占 b 位置）原子全部位置一一得出。结果如图 2-15 所示。

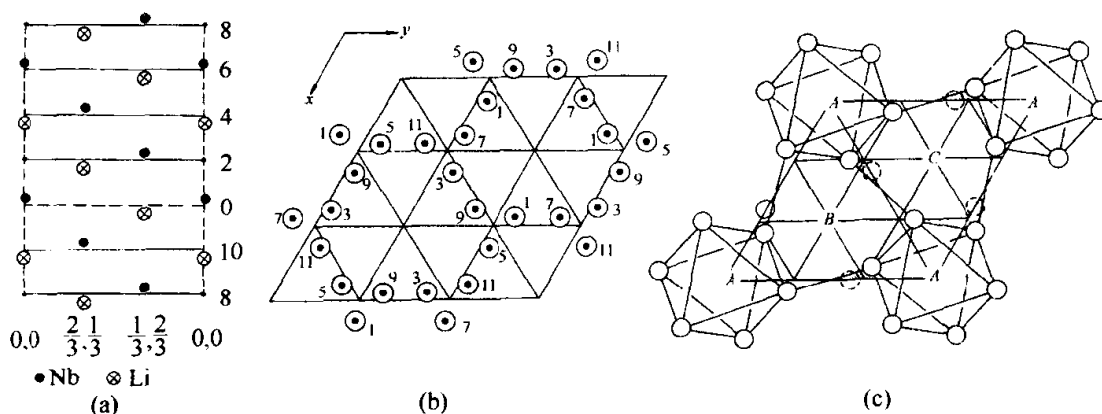


图 2-15 LiNbO_3 的结构

(a) Nb, Li 原子在 $(11\bar{2}0)$ 截面；(b) O 原子在 (0001) 面上的投影，数字为 $c/12$ ；

(c) 在 $z=0$ 面附近的氧八面体，粗圈为 $z=c/12$ ，细圈为 $z=-c/12$

2.3 若干晶体结构的实例

2.3.1 堆积结构

金属键、离子键和范德瓦耳斯键这几种方式键合的晶体，其结构可以由原子或离子的堆积结构 (packing structure) 来理解。

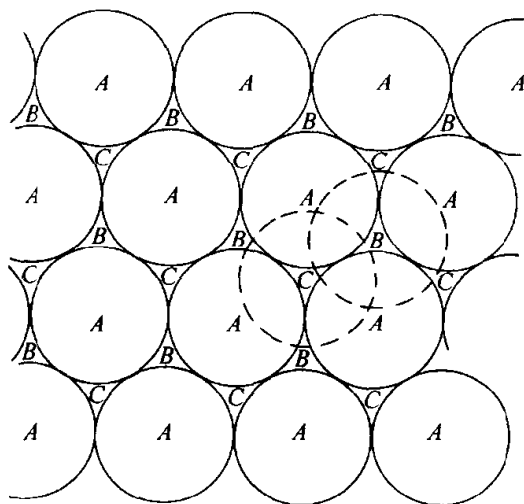


图 2-16 刚球密排面的堆积

首先我们来看一下空间的球体堆积。因为原子的形状基本是球形对称的，因此由球形对称的原子构成的晶体的一个最简单的假设是硬球堆积，而硬球堆积只有在密堆积，即原子间距最小的情况下能量达到极小值，所以硬球的密堆积问题对理解晶体结构是十分关键的。对于二维密排问题，设想一个无限大的平面用一系列圆盘排列的情况，规则的六角排列是最密的排列 (图 2-16)。其堆积系数 (packing fraction) 可以用圆盘的面积与圆盘所占的面积之比来表示。二维

堆积系数相当大，为 0.9069。

三维空间的密堆情况是球体密堆。1611 年，开普勒首先提出了球体的三维密堆，即球体的密排面的密堆，这种球体密堆有两种基本堆法 (见图 2-16)。图中每一层小球都是二维密堆，第二层的最密堆垛是它的球落在第一层球之间的空隙上。从图可以看到，如第一层处于 A 位置，第二层可堆在 B 或 C 处，即第一层的空隙 b 或 c 上。第二层位置的符号要等放上第三层后才确定下来 (图中取为 B)，第三层则堆在 B 层的空隙之上。这样任何更多的堆垛可以用 A、B 和 C 的序列表示，密堆系数为 0.74048。密排面堆垛的层序如按 ABABAB……排列，则形成六角密堆结构，如按 ABCABC……堆垛，则为面心立方结构 (见图 2-17)。

下面我们讨论一下密堆结构的空隙问题。在最密堆垛球之间有两种空隙。在最密排列的原子层上堆第二层时，一种情况是下层的空隙上堆一个上层原子，即空隙周围有 4 个球，它们的中心形成四面体，这就是四面体间隙如图 2-18(c)。第二种类型是空隙上下各有三个球，形成八面体间隙如图 2-18(b)。在所有三维密堆垛中四面体间隙的数目是原子的 2 倍，八面体间隙数等于原子数。如球半径为 R ，空隙中可以填充更小的球。在四面体间隙中小球的半径是 $0.225R$ ，八面体中为 $0.415R$ 。由于四面体间隙体积小，如

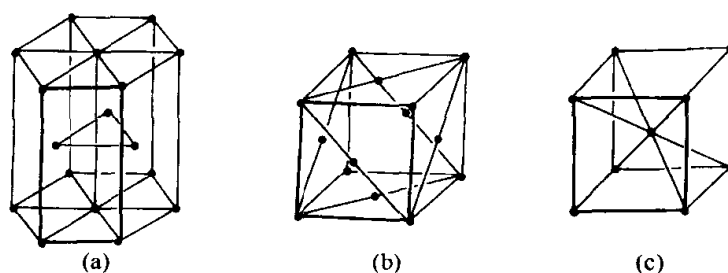


图 2-17 三种典型的堆积结构
(a) 六角密堆；(b) 面心立方；(c) 体心立方

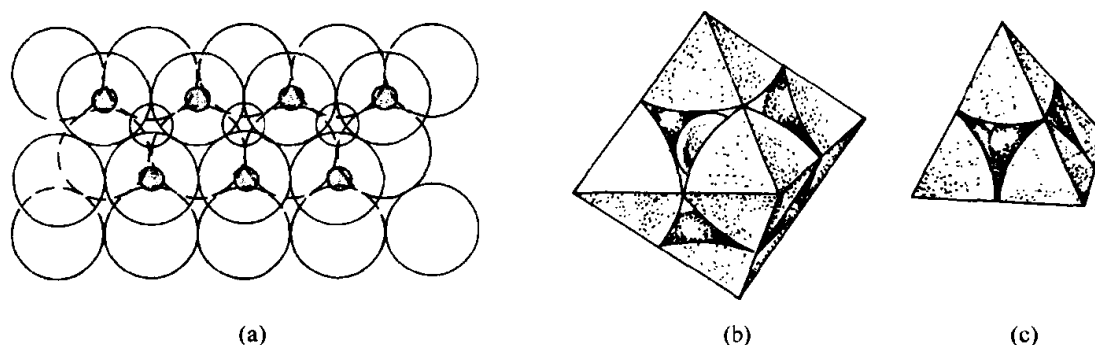


图 2-18 密堆结构中的间隙空洞
(a) 显示两类间隙位置；(b) 八面体间隙；(c) 四面体间隙

果完全按四面体方式堆积起来，空间占有率还可能进一步提高，但因为纯四面体间隙式的堆积和周期结构是不相容的，只有插入一定数目的八面体间隙才能形成晶体。

数学家曾经证明过，理论上三维结构中开普勒最密堆垛是等径球体空间占有率的最高极限，即 0.7404，无从突破。这在追求高的空间占有率上受到了挫折。

另外有些复杂的合金(金属间化合物)具有四面体密堆结构(TCP—tetrahedrally close-packed structure)。前面我们已经讲过，四面体对于堆积是有利的，因为八面体堆积空隙大。四面体堆积结构的合金往往有一种大原子一种小原子，构成高配位的环，有 12 个或 14 个原子环绕一个原子。若四面体用同样的方式堆积不能填满整个空间，利用合金结构，有些四面体发生扭曲，来满足这个要求。这就是四面体密堆结构，又称作 Frank-Kaspar 相，这是以两个科学家的名字命名的。这种 TCP 结构的合金局部区域往往有 20 面体对称性，这和后来发现的准晶有血缘关系。图 2-19 是一种典型的 TCP 结构。

元素和许多化合物常常具有体心立方堆垛(bcc)结构，如果把它形式上看作相同的球，堆垛系数为 68.01%，比最密堆系数小。就不能用硬球堆积

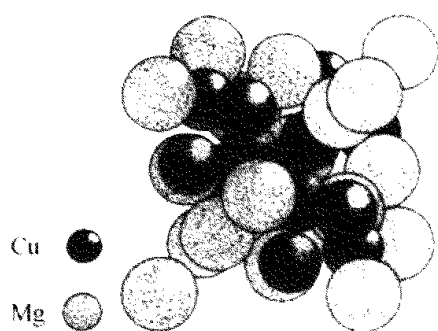


图 2-19 典型的 TCP 结构：
cF24-Cu₂Mg 结构

来解释了。

我们知道原子是有相互作用的球体，严格来说不是绝对的硬球，而是软球。但对各种元素而言，球体的软硬程度上是有差异的，如丰满型金属 Cu、Ni、Al 等，离子半径较大，离子间空隙小，采用 fcc 结构自然有利；而开放型金属，如 Li、Na、K 等，离子半径较小，离子间空隙大，就采取了 bcc 结构。因此 bcc 结构对碱金属这类开放型金属比较有利。对于软

球模型的排列情况，不一定追求最高的堆垛系数。典型的离子晶体，如 NaCl、KCl 等也采用堆积结构，阳离子较小，处于八面体间隙位置，如图 2-20 所示。

2.3.2 键联结构

如果原子间的键合方式是共价键，那么键的重要性就突出来了。晶体结构主要取决于共价键的要求。最典型的键联结构(linkage structure)是四配位的金刚石结构，每个原子有 4 个最近邻，分处在以该原子为中心的正四面体的顶角的位置上。前面在 1.1.3 节中论述过电子

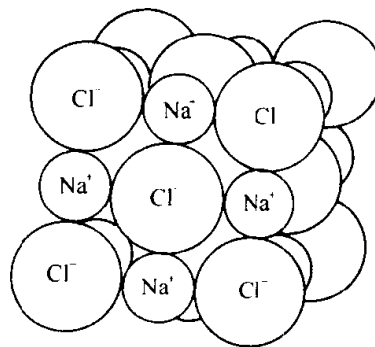


图 2-20 NaCl 结构

处于 sp^3 杂化轨道上与相邻原子的电子构成了共价键。金刚石结构 100% 地实现了 sp^3 杂化键的要求。金刚石结构也可以看作两个面心立方嵌套，其中一个沿体对角线平移四分之一得到的结构。图 2-21 是金刚石的键联结构，

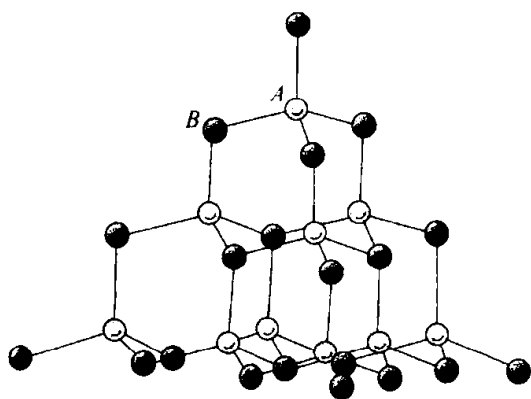


图 2-21 金刚石的键联结构，若 A、B
位置由不同的原子占据
得到闪锌矿结构

硅、锗等元素半导体就采用这种结构。键联结构以金刚石结构为主，闪锌矿(zincblende)是从金刚石结构衍生的，两种原子分别占据了近邻格位。Ⅲ~Ⅴ族化合物半导体 GaAs 等就采用这种结构。

另一种四配位的结构是六角金刚石。六角金刚石是金刚石的一种变体，可以看作密集六角的一组四面体空隙被填充而得到。这种结构实际也有两组位置，如由不同的原子所占则为辉锌矿结构(wurtzite)，

图 2-22 即为六角金刚石的结构示意图。有些半导体采用这种结构。六角金刚石可以理解为冰结构的骨架。冰的 O 位置基本为六角的四配位结构，冰即相当于六角金刚石的原子位置上填入 O，每个氧原子周围有两个 H，这就是冰的主要结构—— I_h 结构（冰有许多种结构）。

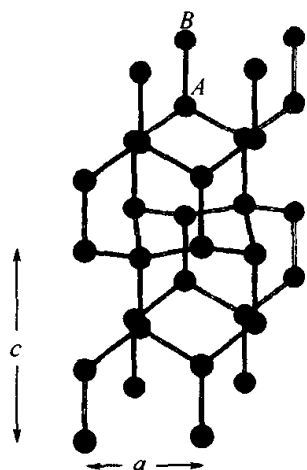


图 2-22 六角金刚石，若 A，B 位置由不同原子占据就得到辉锌矿结构

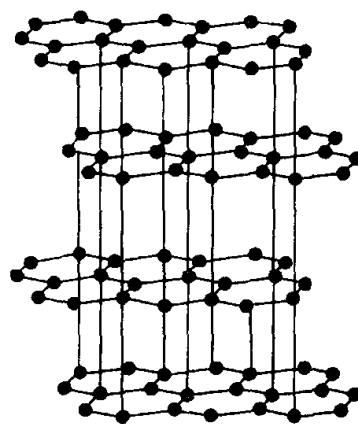


图 2-23 石墨的层状结构

下面我们介绍一下三配位结构。石墨是典型的三配位结构，如图 2-23 所示，常温下稳定的石墨是层状结构， sp^2 杂化构成平面六角网络。 s 电子和 2 个 p 电子组合成夹角为 120° 的键，称为 σ 键，另一个 p 电子离域化构成 π 键。原子层之间是弱的范德瓦尔斯键。层状结构中原子团由强键结合在一起并在二维上无限延伸。

除石墨以外，还有许多具有层状结构的化合物与硅酸盐，是天然的准二维系统。

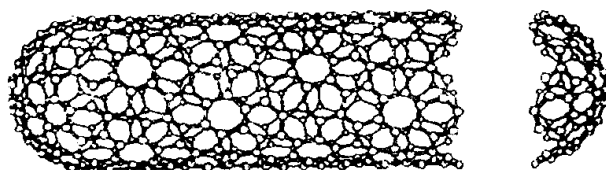


图 2-24 单层纳米碳管示意图。由 C_{140} 卷成
 c_h 为 (15, 5); $d = 1.036\text{nm}$; $\theta_h = -19.11^\circ$

石墨结构之后，具有明显的层状结构的单个石墨片构成六角网络，可以设想将这六角网络卷成直径为纳米量级碳的碳管（图 2-24）。其特征可由联结两个网点 A、A' 的蜷矢(chiral vector)：

$$\mathbf{c}_h = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 \quad (2-12)$$

来确定。这里 \mathbf{a}_1 及 \mathbf{a}_2 分别表示六角网的两个基矢 (图 2-25)。石墨片是沿图中的虚线 (与 \mathbf{c}_h 正交) 蜷接起来形成小管。 \mathbf{c}_h 矢量确定了蜷角 θ_h 和碳管的直径 $d = |\mathbf{c}_h|/\pi$ 。碳管可分为单层管和多层管, 多层管由若干个层间距约为 0.34nm 的同轴圆柱面套构而成。碳纳米管的径向尺寸较小, 管的外径一般在几到几十纳米; 管的内径更小, 有的只有 1nm 左右。而碳纳米管的长度一般在微米量级, 相对其直径而言是比较长的。因此, 碳纳米管被认为是一种典型的一维纳米材料。这种纳米碳管由于引人注目的物理性质, 如高强度、特异的电子输运性质而成为科学家们研究的热点。

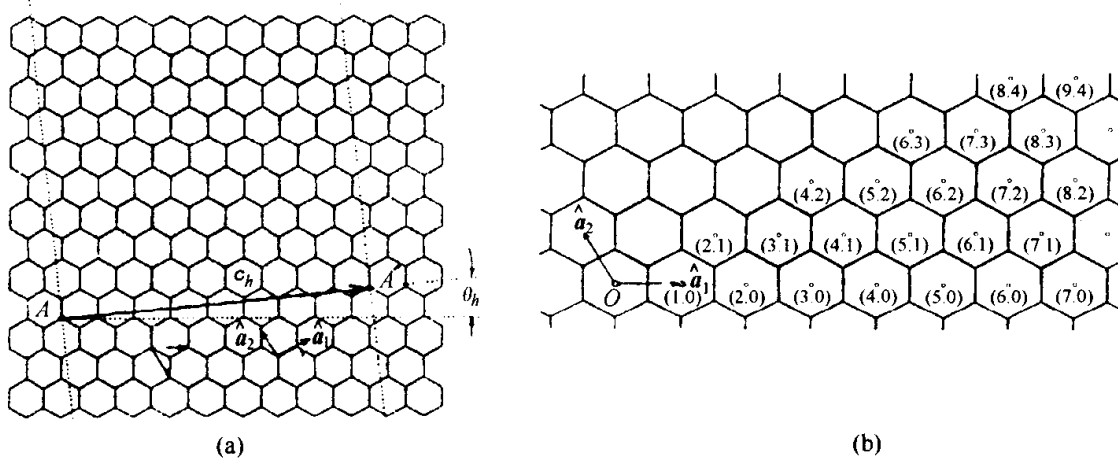


图 2-25 纳米碳管的蜷矢与蜷角

(a) 蜷矢 \mathbf{c}_h 、蜷角 θ_h 和六角网络的基矢 \mathbf{a}_1 、 \mathbf{a}_2 的示意图;

(b) 蜷矢 \mathbf{c}_h 和坐标 (n_1, n_2) 相对应

其他的键联结构还有岛状结构 (包括高聚物以外的所有分子晶体和由无机或有机配位体组成的金属络合物等) 和链状结构等, 我们就不一一细说了。这里, 我们主要介绍一下零维体系的 C_{60} 团簇 (由几个至几百个原子组成的微观和细小聚集体)。

如果将石墨片中一部分六边形用五边形取代, 可以构成封闭的富勒烯 (fullerene)。最新发现的这种笼状 C 的团簇可以用通式 C_{20+2n} 来表示, 由 20 个六边形和 n 个五边形构成。最早发现的是 C_{60} 团簇, 见图 1-21。根据计算, 五边形的最近邻都是六边形时系统的能量最低, 因此 C_{60} 是最小的满足此条件的团簇, 其次是 C_{70} 。纯 C_{60} 固体是绝缘体, 用碱金属掺杂后成为导体, 适当的掺杂成分可以使它成为超导体, 因此 C_{60} 成为人们研究的一个新热点。

2.3.3 钙钛矿结构及其家族

有许多结构构成家族，这类结构有一突出的特点就是能从典型的晶胞出发，原子有一些简单的位移或替代，就能得到许多变体，最著名的例子就是钙钛矿(perovskite)结构家族，有 100 多种变体。

钙钛矿结构有一标准结构 ABO_3 ，键合类型属于离子键与共价键的混合。如 $BaTiO_3$ 高温相为钙钛矿结构的原型，由位于八面体顶角的氧原子框架构成，Ti 在八面体中心，Ba 离子在八面体间的空隙中心(图 2-26)。

从钙钛矿结构原形出发，保持结构的基本骨架，但具体键长和键角允许变化，大体有三种结构变体。

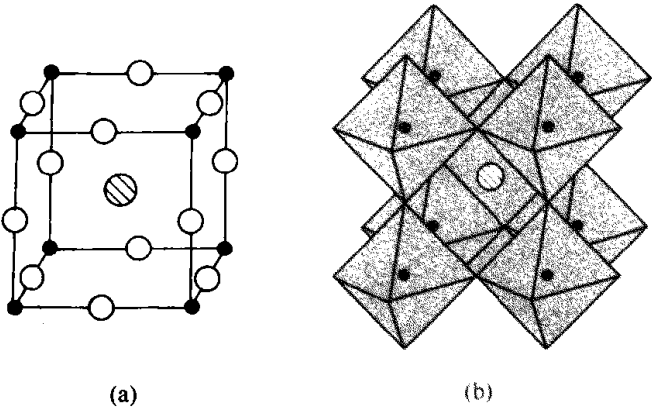


图 2-26 标准的钙钛矿结构
(a) 一个晶胞；(b) 结构中的氧八面体

① 位移型。阳离子偏离中心，沿某一方向位移。例如 $BaTiO_3$ 在 393K，通过 Ti 离子偏心位移而成为四方结构的铁电相。

② 扭转型。阴离子氧八面体对于各轴扭转，如图 2-27 所示的 $SrTiO_3$ 低温结构。

③ 混合型。以上两种效应都有，这样一来，其对称类型可能从立方下降到四方、三角或正交。

钙钛矿结构晶体不仅种类繁多，而且有许多特殊的物理性质。如 $BaTiO_3$ 等具有铁电性，钙钛矿结构的铜氧化物有超导

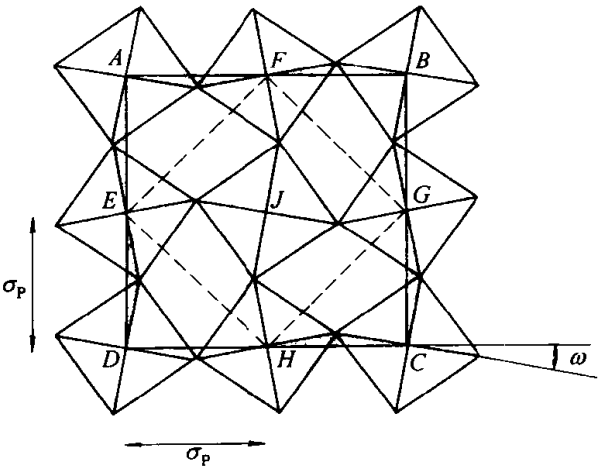


图 2-27 $SrTiO_3$ 的低温结构

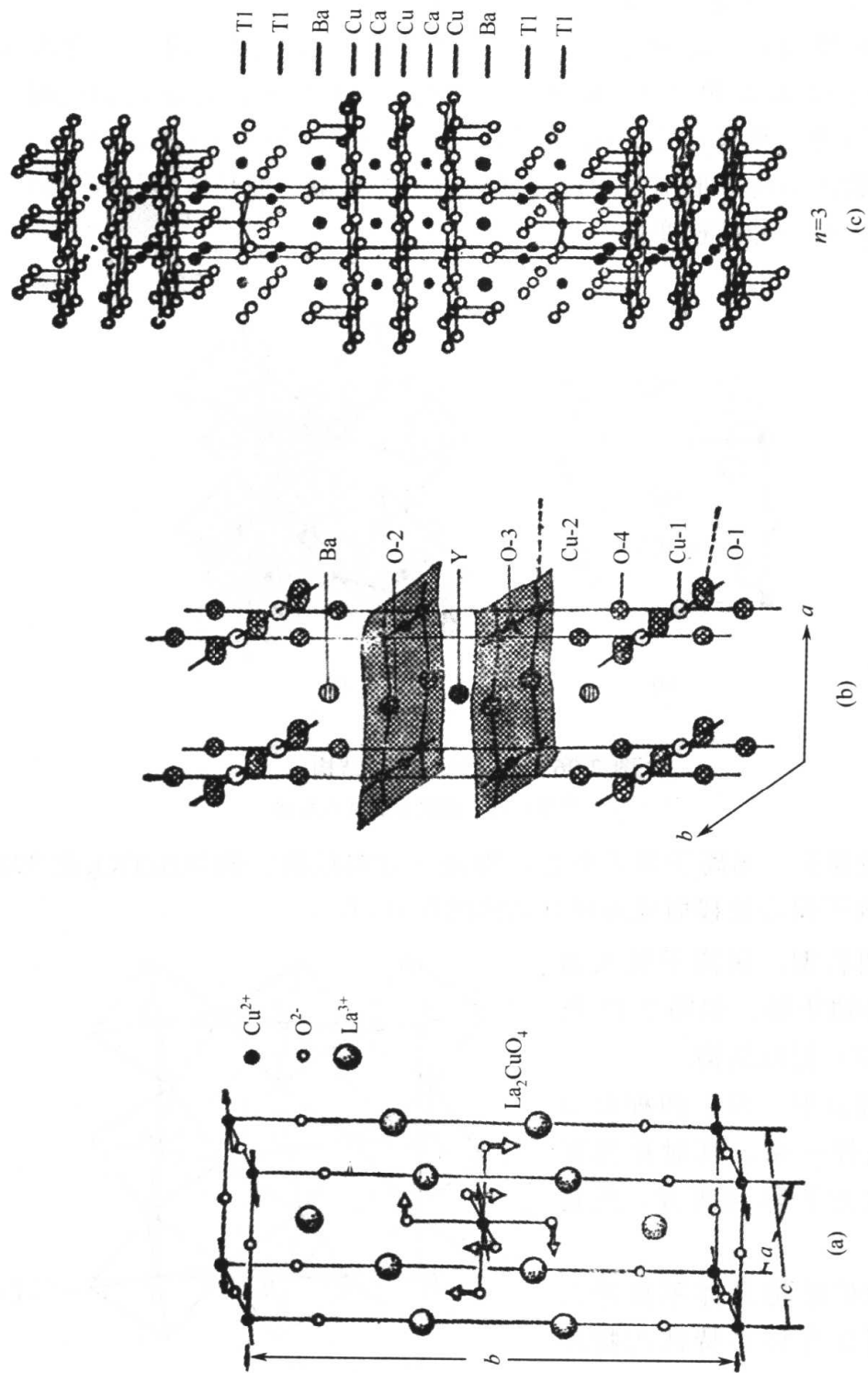


图 2-28 不同的钙钛矿结构
(a) La_2CuO_4 ; (b) $\text{YBa}_2\text{Cu}_3\text{O}_7$; (c) $\text{Tl}_2\text{Ba}_2\text{Ca}_2\text{Cu}_3\text{O}_9$

电性，锰氧化物有庞磁电阻效应……。

1911 年发现超导现象到 1985 年，所得的最高临界温度为 23.3K。1986 年发现了 La-Ba-Cu-O 复合氧化物的 T_c 高达 52K，一个世界范围内的高温超导研究热从此揭开了序幕。X 射线和中子衍射研究表明，这些材料的晶体结构基本上都属于钙钛矿结构的层状变型。例如 La_2CuO_4 属于 K_2NiF_4 结构，它和钙钛矿结构的不同之处在于在两个 BO_2 平面之间有两个（而非一个）AO 平面，其余的有更多夹层。图 2-28 分别为 (a) La_2CuO_4 、(b) $\text{YBa}_2\text{Cu}_3\text{O}_7$ 和 (c) $\text{Tl}_2\text{Ba}_2\text{Ca}_2\text{Cu}_3\text{O}_9$ 的结构图。研究表明， $\text{YBa}_2\text{Cu}_3\text{O}_7$ 在高温（约 $>1000\text{K}$ ）时是四方点阵，为非超导体，低温下是正交点阵，为超导体。

2.4 超出空间群的结构

2.4.1 色群和磁结构

我们来看图 2-29 所示的太极图，它有没有对称性呢？直观上它是有对称性的，但是如果验证一下，又发现它没有了。如果我们将图形沿垂直于图面的轴旋转 180° ，再引入一个新的对称操作：黑白颠倒，图形就可以复原。这种超越常规的对称操作是在 1930 年由 H. Heesch 首先提出来的。在 20 世纪 50 年代以后，前苏联晶体学家许勃尼柯夫对于黑白对称操作的对称群进行了系统的推导，为了纪念他的贡献，黑白空间群也被称为许勃尼柯夫群。

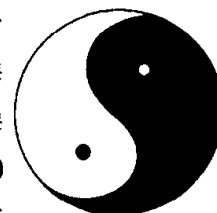


图 2-29 太极图——黑白对称的实例

黑白群也可以看作三维空间群朝四维的推广，而第四个维度限于两种值：黑与白，正与反。当然可以推广到多种颜色。还可以是波函数的相位、自旋、电荷符号等。这类广义的对称群被称为色群。

磁结构是由磁性材料的晶体结构加上磁性原子的磁矩构成的。磁对称群是一种色群，第四个变量为磁性原子的自旋。磁序包括铁磁、反铁磁和铁氧体三类（图 2-30）。图 2-31 为磁结构晶胞示意图，(a)、(b)、(c) 分别为铁磁、反铁磁、螺旋磁结构。在一般的对称操作基础上，加上使磁矩反转的操作，可把 230 种空间群增加到能描述铁磁和反铁磁性晶体对称性的 1651 个对称群，这还不包括螺旋磁结构。

2.4.2 无公度调制结构

无公度调制是指在基本晶格（周期为 a ）上附加一个周期为 λ 的某种调制， λ/a 为无理数，就得到无公度调制，得到的相为无公度相（图 2-32）。无公度相严格来讲是一种准周期结构。调制可以是一维的如 Na_2CO_3 ， NaNO_2 ；也可以是二维的，如 TaSe_2 、石英；甚至可以是三维的，如 Fe_{1-x}O 。在无公度相中，调制只对基本晶格产生另一周期的微扰，基本晶

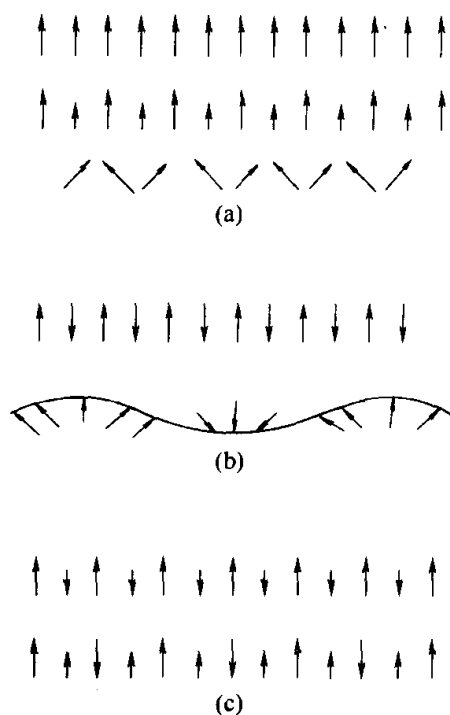


图 2-30 一维自旋排列示意图

(a) 铁磁; (b) 反铁磁; (c) 铁氧体

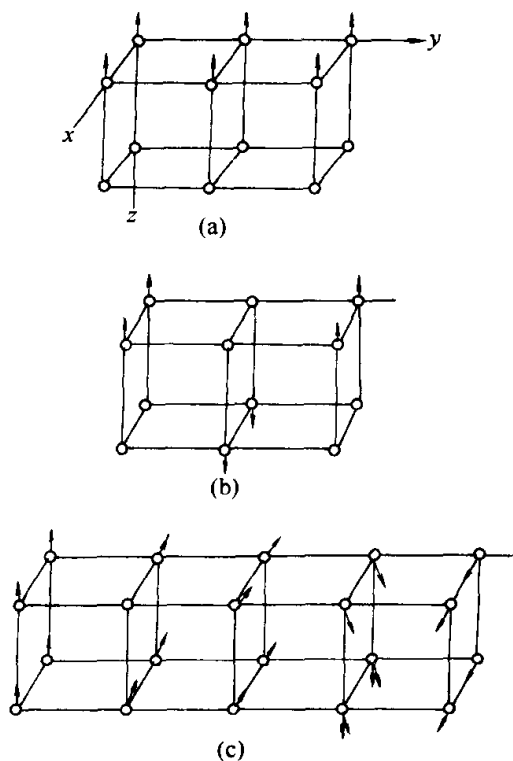
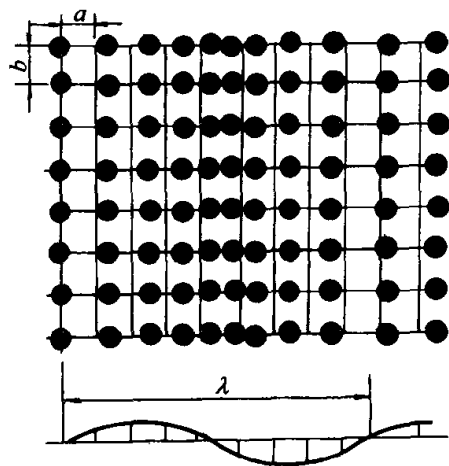


图 2-31 磁结构晶胞示意图

(a) 铁磁结构; (b) 反铁磁结构; (c) 螺旋磁结构

格的衍射图样仍然保留,但在正常衍射斑点之间偏离有理分数处出现卫星斑点。调制周期 λ 和温度及其他外界条件有关,在一定温度下它和公度结构间发生转变。

图 2-32 调制波长为 λ 的
位移调制型无公度相

2.4.3 准晶和 Penrose 拼砌

在无公度相受到人们关注的同时,数学家开始关注平面的非周期拼砌问题,采用形状不同的基本拼块,无空隙、不重复地布满平面。1974 年, R. Penrose 用两块胖、瘦 [图 2-33(a)] 两种菱形实现了这种非周期拼砌,称作 Penrose 拼砌。两种菱形边长相等,角度分别为 36° 和 72° ,而且这两种拼块还有自相似性,可以通过缩放规则来获得形状相似而尺寸不同的拼块 [图 2-33(b)],比例因子为 τ^{-1} , (τ 为黄金分割数)。利用这两种拼块可以构造出如图 2-34 表示的平面拼砌。从图上可以看出,拼砌的局域结构是相似的,即有局域同构性,而且存在晶体学禁止的五

重取向序。图 2-35 是 Penrose 拼砌的光衍射图。

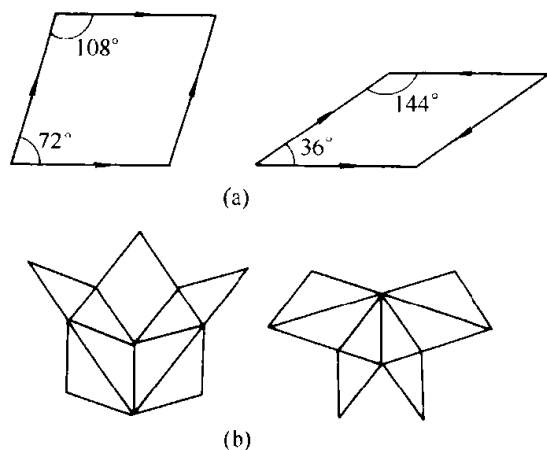


图 2-33 Penrose 拼砌的两种拼块

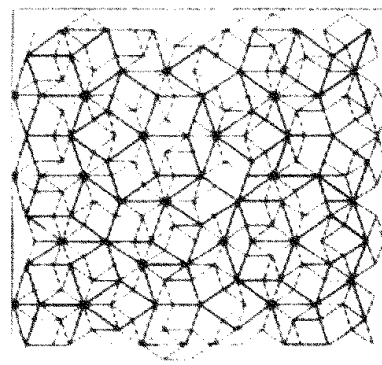


图 2-34 典型的 Penrose 拼砌图

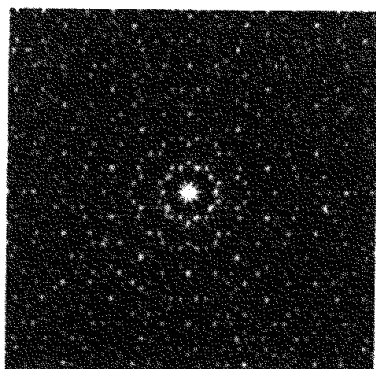


图 2-35 Penrose 拼砌的光衍射图

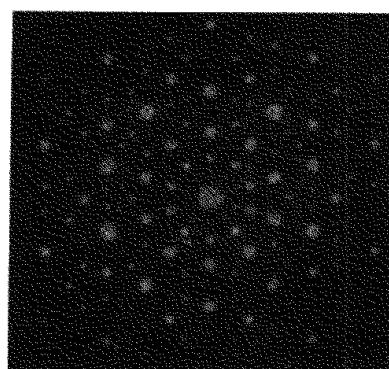


图 2-36 Al-Mn 二十面体相的电子衍射图

1984 年 Shechtman 等在急冷的 Al-Mn 合金中获得了具有二十面体对称性(包括五重对称轴)、斑点明锐的电子衍射图(图 2-36), 这与周期结构是不相容的。Levine 和 Steinhardt 认为, 这正是有非晶体学对称性的三维准周期结构, 相当于 Penrose 拼砌的三维推广, 可取名为准晶(quasicrystal), 作为准周期结构晶体的简称。郭可信等也独立地发现了五重对称性的电子衍射图和急冷 Ni-V 合金中的二十面体相准晶。

准晶是固态物质的一种新有序相, 同时具有长程准周期平移序和晶体学上不允许的长程取向对称。无公度晶体和准晶都具有准周期性, 但两者是有区别的。非晶体学取向对称不仅把准晶从周期和无公度晶体区别开来, 并且对准周期性加以限制。三维准晶虽然在多种合金系中观察到, 但都超不出二十面体相的范畴。另外科学家们发现了多种旋转轴的二维准晶, 其中包括八重、十重与十二重, 当然不能排除可能有其他的类型。

参 考 文 献

- 1 Allen S. M. and Thomas E. The Structure of Materials. New York: John Wiley & Sons, Inc., 1998
- 2 Vainstein B. K. 等著. 吴自勤译. 现代晶体学. 第一卷及第二卷. 合肥: 中国科技大学出版社, 1992
- 3 冯端, 冯步云. 晶态面面观. 长沙: 湖南教育出版社, 1992
- 4 Megaw H. D. Crystals Structures: A Working Approach. Saunders, Philadelphia, 1973
- 5 王仁卉, 郭可信. 晶体学的对称群, 北京: 科学出版社, 1990
- 6 肖序刚. 晶体结构的几何理论. 第二版. 北京: 高等教育出版社, 1993
- 7 Elliot S. R. The Physics and Chemistry of Solid. New York: John Wiley & Sons, Inc., 1998
- 8 Hahn T. International Tables for Crystallography, vol. A, 4th ed., Kluwer, Dordrecht, 1996

第3章 非晶态与液晶态

撰稿人 冯 端 万建国

在晶体中，原子的取向和位置均具有长程序，而在液体中，这两个长程序都消失了，原子的位置具有无序性和非定域性。如果再对液体进行快速冷却，就有可能获得一种特殊的非晶态，即玻璃态。因此，玻璃态实质上是一种过冷液体，其结构特征是短程有序，长程无序。本章主要讨论有关玻璃态的问题，包括各种无机玻璃和聚合物。另外，对一些具有取向序的液态，如液晶，本章也将进行讨论。

3.1 导向非晶态

3.1.1 熔化

我们可以给出晶体熔化过程的基本物理图像：从有序的晶格出发，原子的挪动不仅使晶格坐位上的原子发生重新排列，而且导致了晶格的彻底瓦解。这一过程可以用计算机来模拟，如图 3-1 所示。图 3-1(a)对应于高温下的晶体，虽然存在不规则的热运动，但原子仍停留在晶格的坐位附近。图 3-1(b)对应于液态，晶格已经不复存在，原子的位置分布是无规的，体现了位置无序性，而且原子并不定域于某个特定的坐位上。图 3-1 充分揭示了晶态与液态在结构上的差异，说明了液态结构中原子位置的无序性和非定域性。

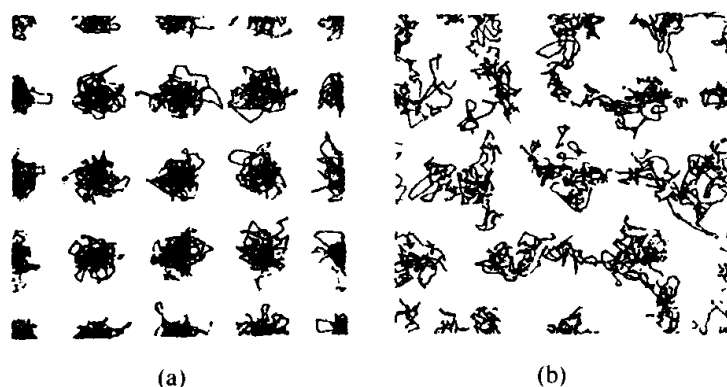


图 3-1 32 个硬球运动轨迹的计算机模拟(采用周期性边界条件)

(a) 晶相；(b) 液相

F. A. Lindemann 提出了晶体熔化的判据：当原子热振动的均方根位移与原子间距之比值超过一定限度 δ 后，晶体即产生熔化，即

$$\frac{(\langle u^2 \rangle)^{1/2}}{a} > \delta \quad (3-1)$$

δ 值因晶体结构和材料类型而有差异。对于体心立方金属 Li、Na、K 等， $\delta \approx 0.11$ ；对于面心立方金属 Cu、Ag、Au、Al 等， $\delta \approx 0.07$ 。

科学家通过计算机模拟来进一步揭示熔化问题。如图 3-2 所示，用硬球代表原子，可以定出固相的状态方程。当原子体积 V/V_0 (V_0 为密集状态的原子体积) 之比接近于 1.5 处就发生了熔化（这是金属熔化的通常情况，熔化导致体积膨胀。当然也存在一些反常情况，如冰、硅、锗等在熔化时体积反而收缩，后一种情况与强共价键有关，这里就不细说了）。然后就转到了液相状态方程曲线。如果对液相进行压缩，状态方程就会沿着另一支曲线变化。这一结果表明熔化是一级相变，在固相到液相的相变过程中，原子体积（也可以说密度）有一跳跃式的变化，基本上不存在过热状态，但在降温时，不会在熔点转变为固体，而是以过冷液体的方式存在。如果用软球模型，原子间作用势采取 $\Phi(r) = cr^{-n}$ ，进行计算机模拟，可以估计出熔化时的参数， δ 的临界值约在 $0.07 < \delta < 0.09$ 之间，与假定的势函数中的 n 值几乎无关，证实了 Lindemann 的设想。

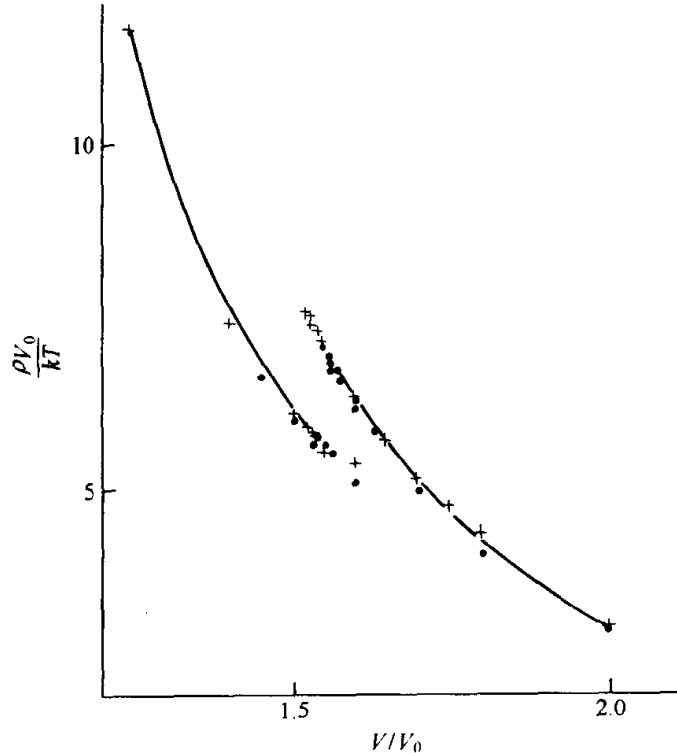


图 3-2 硬球系统的状态方程

(+ 计算机模拟, • 来自不同的模拟计算, V_0 为密集晶态的体积)

熔化是一级相变，在这一过程中必须吸收熔化潜热，体积发生不连续变化，导致熵也产生跃变。造成熔化熵的物理原因可以归结为原子排列无序度的增加，原子公有化相应的贡献，以及物质内部自由体积的膨胀。这些因素错综复杂，因此很难建立一个清晰的熔化理论模型。

3.1.2 玻璃化转变

当液体冷却到熔点 T_m 时，并不会立即凝固或结晶，而是先以过冷液体的形式存在于熔点之下。新的晶相形成，首先要经过成核阶段，即在局部形成小块晶核。由于晶核尺寸很小，表面能将占很大的比例，因而将形成能量的壁垒。因此，在熔点之上，成核是不可能实现的。只有当温度下降至熔点以下，即存在一定的过冷度 $\Delta T = T_m - T$ 时，成核的几率才大于零。晶核形成后，晶核的长大就主要依靠原子的扩散过程。因此，结晶的速率既和成核的几率有关，又和长大的速率有关。前者取决于过冷度的大小，后者则取决于温度的高低。图 3-3 显示了结晶的体积分数为 10^{-6} 的转变曲线，曲线的纵坐标为温度，横坐标为时间的对数。可以看到，曲线的形状类似于字母 C。不同的材料有不同的 C 曲线，但基本形状都是相似的。我们可以看出，如果从液相冷却下来的速率足够快的话，冷却曲线将不与 C 曲线的鼻尖相接触，这样过冷液相就将避免结晶而形成玻璃态。

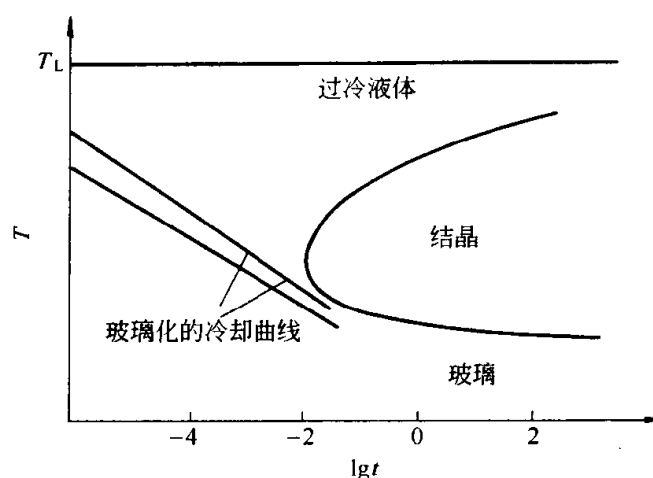


图 3-3 结晶曲线示意图

图上 C 形曲线代表结晶体积分数为 10^{-6}

人类使用玻璃态已经有几千年的历史了。我们通常所说的玻璃，一般是指以 SiO_2 为主要成分的氧化物玻璃。这些氧化物玻璃具有相当复杂的晶体结构，在液态时就有很强的粘滞性，造成原子的扩散相当困难，因此在冷却过程中晶核的形成和长大速率很低。所以一般的冷却速率 ($10^{-4} \sim 10^{-1} \text{ K/s}$) 就足以使这些液态的氧化物避免结晶，而形成玻璃。

但对于金属或合金，情况就完全不同了。由于原子的扩散速率很大，因此一般的冷却速率是无法形成玻璃的。1959年，P. Duwez 在实验室发展了一种泼溅淬火(splat quenching)技术，将液滴泼溅在导热率极高的冷板上，使冷却速率高达 10^6 K/s ，首次将 Au_3Si 合金制成了玻璃态，开创了金属玻璃的新纪元。后来人们又发展了熔态旋淬(melt spinning)，将熔融的合金喷注在高速旋转的冷金属圆筒上，形成金属玻璃的薄带，以 1000 m/min 的速率甩出，使金属玻璃的生产工业化。另外还发展了激光玻璃化的技术，以激光束产生快速熔化和淬火，冷却速率可高达 $10^{10} \sim 10^{12} \text{ K/s}$ ，甚至可以形成玻璃态的硅，而通常的非晶硅则是用气相沉积的方法来制备的。

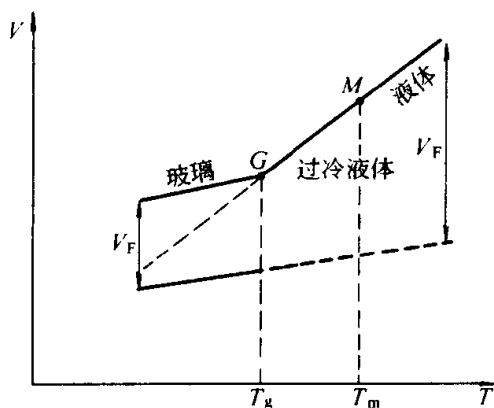


图 3-4 过冷液体冷却过程中体积随温度的变化

T_m 为熔点， T_g 为玻璃化转变温度， V_F 为自由体积

目前，除了少数金属元素以外，几乎所有元素和化合物都可以用熔态淬火的方法来制备玻璃态。可以设想，进一步提高冷却速率，将可能导致所有的物质都可以制备成玻璃态，计算机模拟淬火表明，这一设想是完全可行的。

图 3-4 是液体冷却过程中体积随温度的变化关系曲线，从图中我们可以看出结晶和玻璃化过程的差异。前者在凝固时体积有一跃变（正好相反于晶体熔化时的体积跃变）；而后的体积变化则是连续的，但在体积连续变化过程中在温度 T_g 处斜率产生了明显的转折，这一转折称为玻璃化转变，对应于过冷液体转变为玻璃态。

有的科学家采用自由化体积理论来描述玻璃化转变。当自由体积不再随温度的变化而变化，即达到一临界值时，对应于该值的温度即为玻璃化温度。此时的自由体积分数 f_F 为

$$f_F = f_g + (T - T_g)(\alpha_L - \alpha_g) \quad (3-2)$$

式中， α_L 、 α_g 分别表示液相和玻璃态下的热膨胀系数， f_g 对应于 T_g 的自由体积分数。

玻璃化转变的实质是什么呢？前面我们已经讲过晶体与液体的区别。晶体在结构上是有序的，原子停留在晶格坐位附近，具有定域性。而液体具有流动性，在结构上是无序的，原子是非定域的。原子的定域性是固体的特征。玻璃化转变对应于液体原子非定域性的丧失，原子被冻结在无序结构中，这就是玻璃化转变的实质，即结构无序的液体变成了结构无序的固体。

这个过程和液态结晶过程是不同的。在液态结晶过程中，存在两种类型的转变：结构无序向结构有序的转变以及原子非定域化向原子定域化的转变，这两种转变是耦合在一起同时实现的。而在玻璃化转变过程中，这两种转变却脱耦了，只实现了原子非定域化向原子定域化的转变，而结构无序却仍然存在。

玻璃化转变造成了材料粘滞系数的急剧变化。在 T_m 到 T_g 之间过冷液体的粘滞系数可以用经验公式来表示：

$$\eta(T) = Ae^{\frac{B}{T-T_0}} \quad (3-3)$$

式中， T_0 为小于 T_g 的一个特征温度。可以看出，在这一温度范围内，粘滞系数急剧上升。在 T_g 附近超过了 $10^{12} \text{Pa} \cdot \text{s}$ 。粘滞系数急剧上升使流动性丧失，从而转变成为玻璃态。

我们还可以从熵的变化角度来解释玻璃化转变。液体的冷却过程同时也是熵的排除过程。图 3-5 是 $\text{H}_2\text{SO}_4 \cdot 3\text{H}_2\text{O}$ 的过剩熵随温度的变化曲线。可以看出，如果 T_g 不存在玻璃化转变，那么沿原曲线外推，在 $T = T_0$ 处将出现过剩熵为零的这一物理上不合理的结果。为了避免这一结果

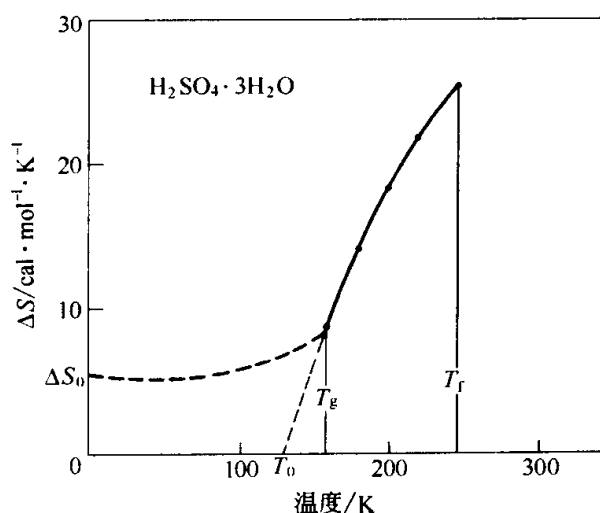


图 3-5 熵的变化与玻璃化转变

注：1cal = 4.18J

的出现，在 T_g 处将发生玻璃化转变，从而使曲线下降的趋势变缓，避免了一场熵的灾难。降至 $T = 0$ 处，仍保留一定的剩余熵，这和热力学第二定律要求平衡系统 $T = 0$ 处熵为零的结果并不矛盾，因为玻璃态是非平衡相。

玻璃态的非晶态固体材料已经在许多领域中获得了应用，表 3-1 列出了部分非晶态固体应用的实例。

3.1.3 位置无序的统计描述

由于非晶态的长程无序特征，因此对非晶态的结构描述，一般都是用统计的方法来描述，即用它的径向分布函数来描述。

从每个原子的中心作出引向邻原子中心的连线，这些连线的中分平面可以形成一系列无规的多面体，称为伏龙诺伊 (Voronoi) 多面体。伏龙诺伊多面体的面数相当于原子的配位数 z 。在位置无序的原子系统中， z 值随原子的种类而异，是有起伏的。但 z 的平均值是描述结构的很有意义的一个参量。

表 3-1 非晶态固体应用的实例

非晶态固体的类型	代表性的材料	应 用	所用的特性
氧化物玻璃	$(\text{SiO}_2)_{0.8}(\text{Na}_2\text{O})_{0.2}$	窗玻璃等	透明性, 固体性, 形成大面积的能力
氧化物玻璃	$(\text{SiO}_2)_{0.9}(\text{GeO}_2)_{0.1}$	用于通讯网络的纤维光波导	超透明性, 纯度, 形成均匀纤维的能力
有机聚合物	聚苯乙烯	结构材料, “塑料”	强度, 质量轻, 容易加工
硫系玻璃	$\text{Se}_{1-x-y}\text{As}_x\text{Ge}_y$	静电复印技术	光导电性, 形成大面积薄膜的能力
非晶半导体	$\text{Te}_{0.8}\text{Ge}_{0.2}$	计算机记忆元件	电场引起非晶 \leftrightarrow 晶化的转换
非晶半导体	$\text{Si}_{0.9}\text{H}_{0.1}$	太阳能电池	光生伏打的光学性质, 大面积薄膜
金属玻璃	$\text{Fe}_{0.8}\text{B}_{0.2}$	变压器铁心	铁磁性, 低损耗, 形成长带的的能力

为了对位置无序结构进行定量描述, 引入原子分布函数。引出矢量 R_{12} , 在矢量端点处微小区域 $dv_1 dv_2$ 内一对原子的出现几率为 $g(R_{12})$ 。 $g(R_{12})$ 即被称为原子分布函数。由于液态和玻璃态均表现为各向同性, R_{12} 的方向可以任取, 因而可以完全忽略其方向性, 故原子分布函数可以写成 $g(R)$ 的形式:

$$g(R) = \frac{1}{\langle \rho \rangle} \frac{dn(R, R+dR)}{dv(R, R+dR)} \quad (3-4)$$

这种形式的两体分布函数称为径向分布函数 (radial distribution function, 简称 RDF)。 $g(R)$ 的物理意义是什么呢? 我们不妨来看一个具体的例子。从图 3-6 所示结构的中心原子出发, 以半径 R 和 $R+dR$ 的球面壳层内的平均原子数就等于 $g(R) 4\pi R^2 dR$ 。很显然, $R < R_0$ 处 (R_0 : 一个原子的硬球直径), $g(R) = 0$; 从 R_0 起, $g(R)$ 开始上升, 到第一峰值处 ($R = R_1$), 又重新下降, $g(R)$ 的第一个峰对应于中心原子周围的第一个配位层, 第一峰下的面积就等于此结构的配位数 z 。由于结构存在无序, z 就不一定是整数。类似地可以定出近邻的第二壳层, 但峰展宽了, 峰高也降低了。逐渐和其他的峰合并, 终于在 $R \rightarrow \infty$ 时, $g(R) = 1$ 。如图 3-7 所示。这里, 我们可以引入关联函数

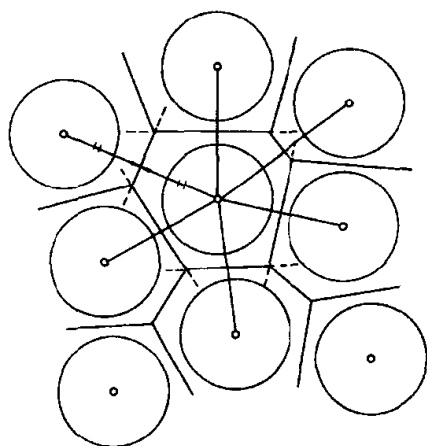


图 3-6 无序系统的伏龙诺伊多面体的二维示意图

$J(R) = g(R) - 1$ (3-5)

这正好可以用来反映局域中偏离统计均匀性的情况。我们也可以定义序的范围 L , 即当

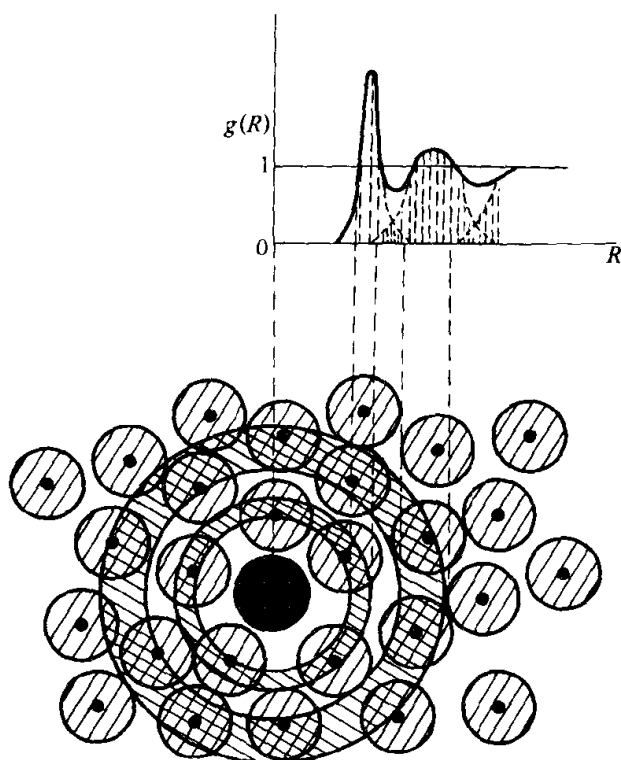


图 3-7 揭示径向分布函数含义
的二维示意图

$R > L$, $h(R) \approx 0$ 。与具有长程序 $L \rightarrow \infty$ 的晶体结构不同, 液体和玻璃的结构都只有短程序, 即 L 为数个原子间隔。为了和 X 射线 (或电子; 或中子) 散射实验的结果相对照, 可以定义 $J(R) = 4\pi R^2 g(R)$ 为径向分布函数的另外一种表达形式, 通过实验测出散射强度 $I(\theta)$ 对散射角 2θ 的散射, 可以通过傅里叶变换, 求得 $J(R)$ 或 $g(R)$ 。

图 3-8 给出了晶体、玻璃和气体典型的径向分布函数图及相应的结构示意图。可以看出, 晶体的 $g(R)$ 是敏锐的峰, 而气体的是平坦的直线; 非晶和液体的则介乎其间, 在短程范围内是振荡式的, 到长程范围就趋于平坦。

通过径向分布函数, 我们可以获得表征液体和玻璃态结构的信息, 诸如有关短程序和化学键的性质, 同时也是检验不同结构模型从而进行甄别的关键。

径向分布函数是对所有原子统计平均的结果, 仅仅描述固体中一个平均原子的周围环境, 并不能给出非晶态原子分布的全貌, 在统计过程中抹去了不少结构信息, 如对不同种原子组成的固体, 就漏掉了一些有关化学关联及原子键键合性质的有价值的信息。因此, 利用径向分布函数来描述非晶态结构信息存在一定的局限性。要获得结构的细节, 还得乞助于构造模型。下面

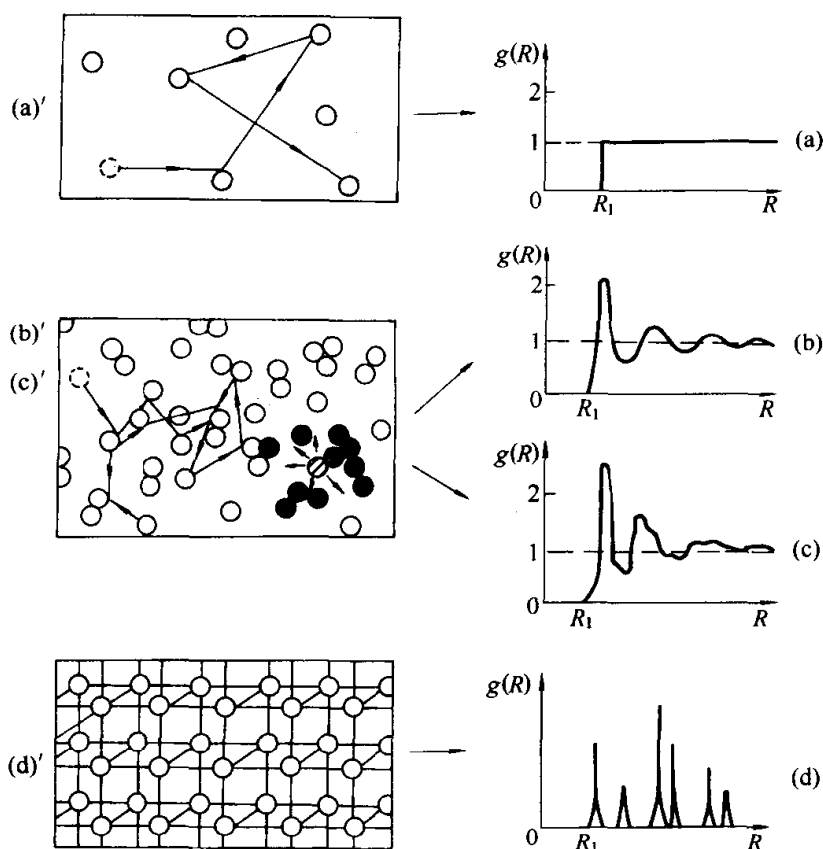


图 3-8 四种状态的物质

(a)气体；(b)液体；(c)非晶；(d)晶体的径向分布函数 $g(R)$

(a)'(b)'(c)'(d)'分别为它们对应于某时刻的原子分布状态

我们将介绍两类著名的模型。

3.2 无机玻璃

3.2.1 无规密堆模型

无规密堆模型是英国晶体学家 J. D. Bernal 于 20 世纪 50 年代末期为阐述液态结构问题所提出的,现在已经成为较成功地说明金属玻璃结构的模型。这一模型的基本思路是:将液体视为均匀的、相干的而且基本上是无规的分子集合,其中并不包含晶态的区域,或在低温下存在大到足够容纳其他分子的空洞。为了回避分子位形所造成的复杂性,只考虑了球形原子的堆集问题。为了获得无规密堆模型, Bernal 制备了多种模型进行研究,如球杆模型,橡皮泥球体的压结,大量滚珠的堆集。他将许多表面粘有粉笔屑的橡皮泥球放在足球胎中,抽去空气,再将它们压紧,然后取出来观察其几何形状。发现有少量十二面体存在,但总的情况比较复杂,而且并不规则,占主导地位的是包含许多五边形的多面体,如图 3-9 所示。于是 Bernal 进一步利

用了上千个滚珠进行密堆实验并设计出表示无规密堆的球杆模型, 通过计算机模拟获得了相应的伏龙诺伊多面体。在模型中球体的空间占有率为 $(63.66 \pm 0.004)\%$, 显然低于晶体密堆集的对应值 74.05% ; 多面体的面数平均值为 14.251 , 各个面的边数平均值为 5.158 , 相当接近于五边形。后来这一模型得到进一步发展, 变得更加精确, 并引入了球体的相互作用势 (即所谓的软球) 来取代原始的硬球, 如图 3-10 所示, 使模型进一步接近实际情况, 结构也得到了调整, 产生了类似于非晶体中观察到的结构弛豫。



图 3-9 球体的二十面体密堆积显示了五次对称性

无规密堆模型还可以从间隙多面体的分布情况来加以表征, 如图 3-11 所示。在 Bernal 的原始工作之中, 孔隙的分布比较杂乱, 虽然以四面体间隙为主, 还向八面体间隙连续延伸, 间隙的形状也不太规整, 但在引入相互作用势后, 产生了结构弛豫, 间隙形状也得到了调整, 主要只剩四面体和八面体两种间隙。

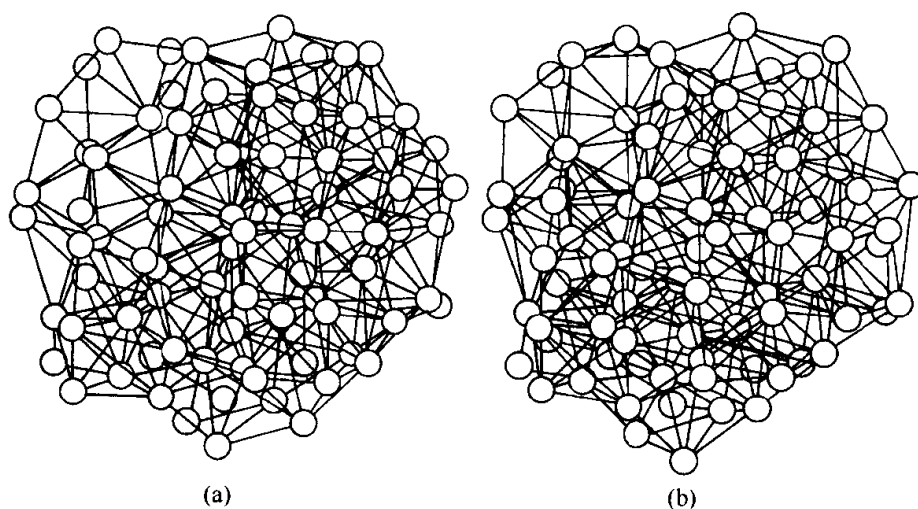


图 3-10 无规密堆的球杆模型 (100 原子)
(a) 硬球; (b) 软球 (引入相互作用势)

金属玻璃的模型接近于 Bernal 的硬球无规密堆模型, 将根据 X 射线测定的径向分布函数的实验结果与理论计算结果进行对比, 发现两者基本吻合。

3.2.2 无规网络模型

1932 年美国物理学家 W. H. Zachariasen 在研究 SiO_2 玻璃的结构问题时

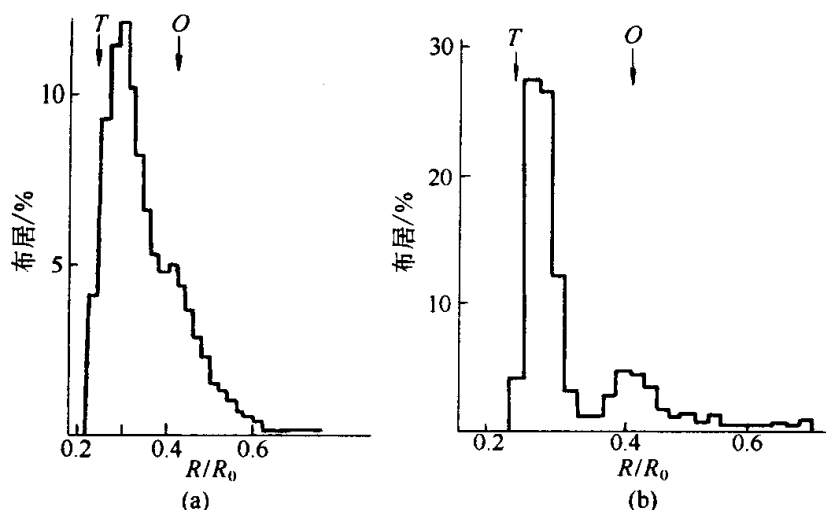


图 3-11 无规密堆模型的间隙分布
(a) 硬球; (b) 软球 (引入相互作用势)
(R_0) 为球半径; T 为四面体间隙; O 为八面体间隙

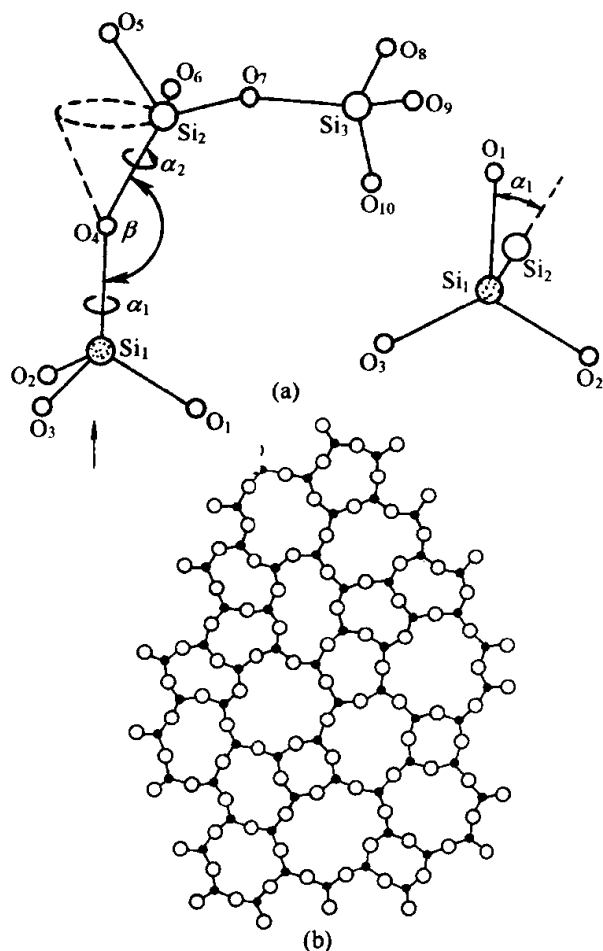


图 3-12 无规网络模型
(a) Si—O 四面体由共顶点氧原子所联结;
(b) Zachariasen 提出的二维无规网络示意图

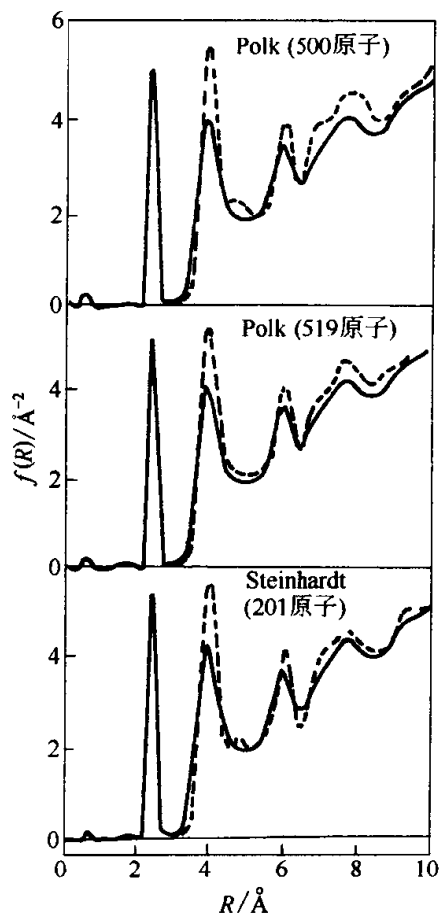


图 3-13 Ge 的径向分布函数
 $1\text{\AA} = 0.1\text{nm}$
(实线为实验结果, 虚线为根据连续
无规网络模型的计算曲线)

提出了连续无规网络模型。其基本思路是：结构的基本单元为 4 个氧原子构成的四面体，并与处于中心处的四价硅原子键合；而相邻的四面体是共顶点的，因而无限结构形成后，化学式保持为 SiO_2 。这样可以形成一个无规网络结构，无规性的引入使 Si—O—Si 键角可以对平均值产生偏离，键长也可以相应地予以伸缩，还可以沿 Si—O 键来旋转四面体的方位。如图 3-12 所示。

科学家也利用计算机来模拟无规网络，构筑了几百个四面体单元的 SiO_2 球杆模型， Si—O—Si 键角平均分散度约为 $\pm 20^\circ$ 。对非晶 Si 与 Ge 的模型，则约 12%，键长平均伸长了约 10%。这些模型是按照 Bernal 提出的“基本上无规的集合”方式构筑的，同时继承了 Zachariasen 的维持化学键近程序的传统，与无规密堆积模型一样，根据连续网络模型也可以导出其原子坐标，密度以及构成闭合圈的成员数目的统计数据。图 3-13 显示了连续无规网络模型与 Ge 的实验结果对比，大致还是令人满意的。

3.2.3 硫系玻璃

硫系玻璃是由硫、硒、砷、锗等元素构成的一类具有无规网络结构的玻璃。

一般硒的晶体结构是由八个原子构成一个分子，进而再构成晶体。但是，如果在硒晶体中加入少量的砷和锗，则硒将形成无规网络结构，并呈现非晶态，这有点类似于无机聚合物。 $\text{Se}_{1-x-y}\text{As}_x\text{Ge}_y$ 的无规网络结构如图

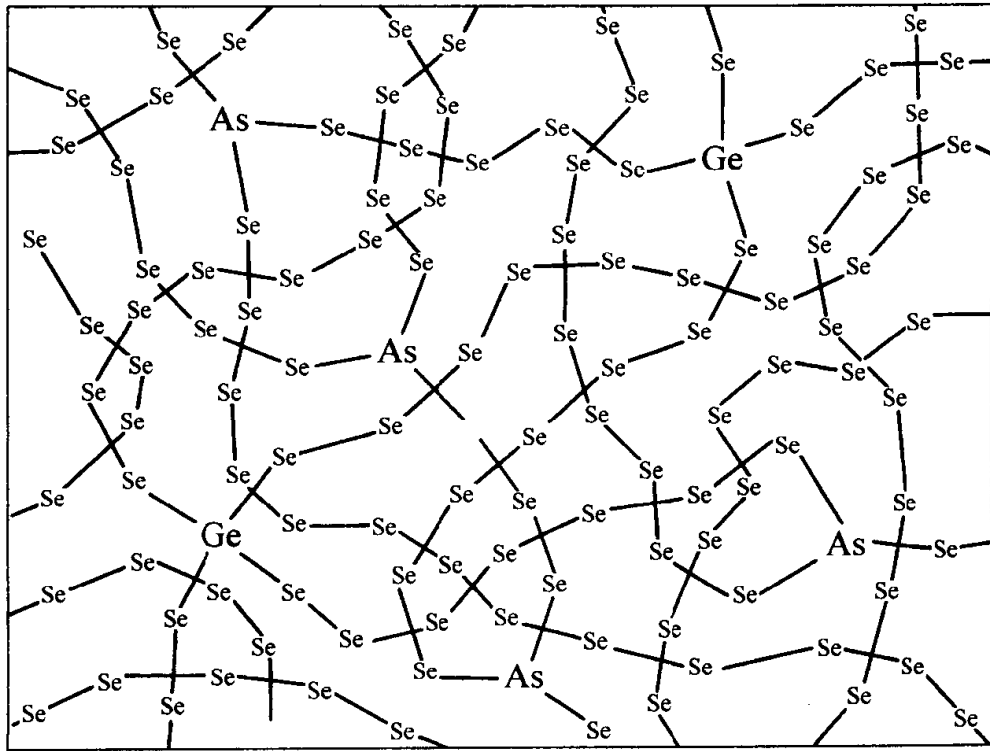


图 3-14 $\text{Se}_{1-x-y}\text{As}_x\text{Ge}_y$ 的无规网络结构模型示意图

3-14所示。作为一种光伏材料,这类玻璃在静电复印中有重要应用。

3.3 聚合物

3.3.1 聚合物的结构和构型

聚合物是由长链大分子(又称高分子)所构成的。构造的高分子的结构单元是单体,高分子中单体的数目一般在 $10^2 \sim 10^5$ 之间。图 3-15 给出了若干高分子单体的结构式。图 3-16(a)给出了聚乙烯中的一 CH_2 — 的空间结构,图 3-16(b)给出了聚乙烯中一连串 C—C 链的连接情况。

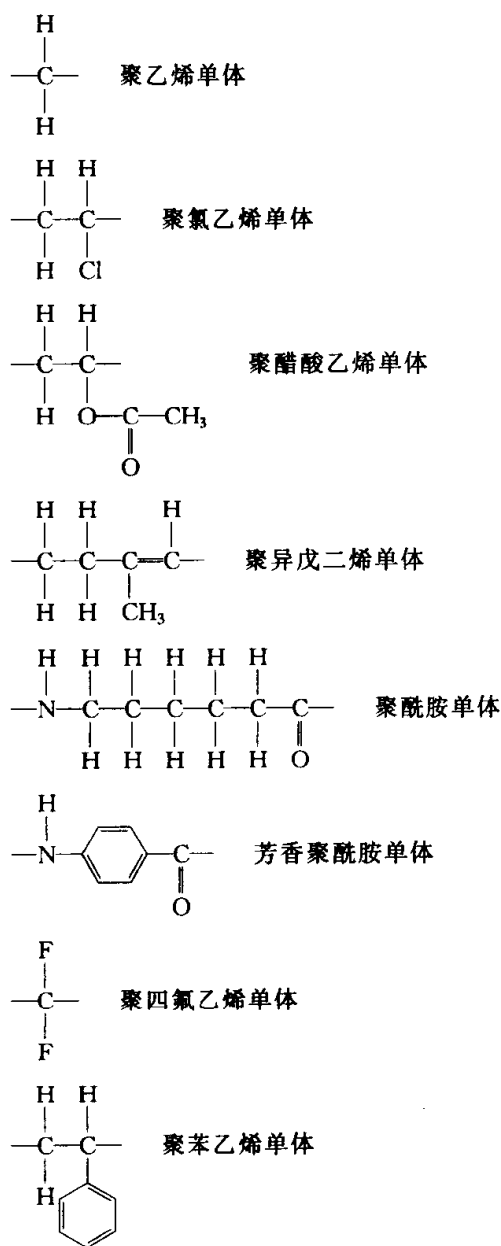


图 3-15 若干形成聚合物的单体结构式

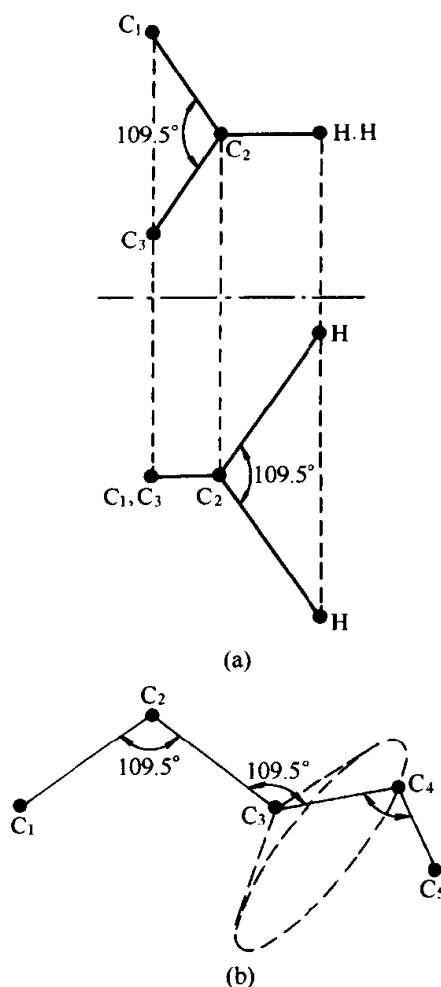


图 3-16 聚乙烯的结构

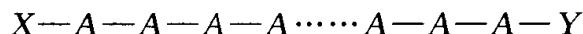
(a) 聚乙烯中的 CH_2

(上部为俯视图,

下部为侧视图);

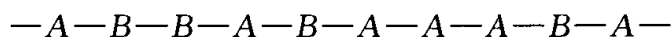
(b) 5 个碳原子的键链

单体可以简单地重复来构成高分子。例如：



其中 A 代表重复单元， X 、 Y 代表单体终端单元，在聚乙烯中， $X=Y=H$ 。

长链中单体的结构和类型不一定完全相同，往往会出现许多种不同的变型，单体的排列方式也可以千变万化。例如：



其中的单体 A 和 B 就是两种不同的单体。

共聚物(copolymer)是由两种或多种不同的单体按照一定的排列方式构成的。共聚物存在多种类型。按照单体的不同排列方式，共聚物一般可以分为无规共聚物(random copolymer)、嵌段共聚物(block copolymer)等类型，如图 3-17 所示。

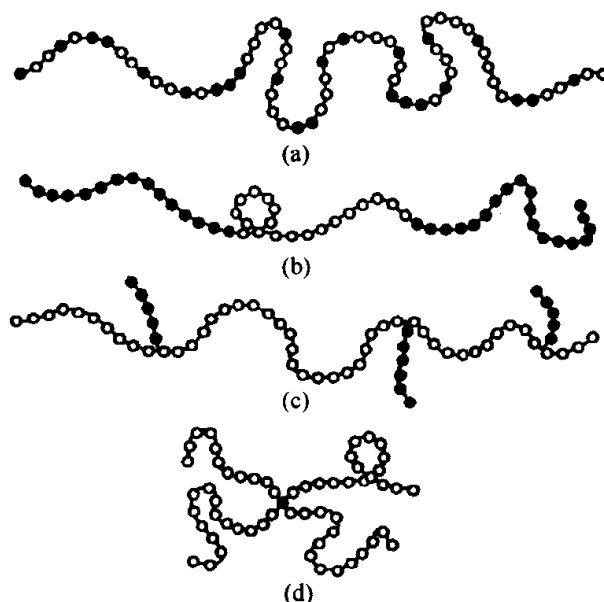


图 3-17 各种不同类型的共聚物

(a) 无规共聚物；(b) 三嵌段共聚物；(c) 接枝共聚物；(d) 四臂星型均链聚合物

而在生物体聚合物中，各个单体几乎都不一样，是更加复杂的非周期聚合物。但不管高分子链的具体结构如何，长链的存在本身就意味着强的关联性和一定的有序性。

聚合物的构象是十分复杂的。高分子链大体上可以分为三类：柔性链、刚性链和螺旋链。在主链中含有苯环或杂环的高分子一般是刚性链，呈棒状，如图 3-18(a)所示。而像聚乙烯这样的高分子，主链是由 $C-C$ 链所构成，由于存在能量相差不多的几种基本构象之间可以通过键内旋转来实现，

从而使分子链看起来千姿百态，像一团团无规的线团，这就是柔性链，如图 3-18(b)所示。还有一种本质上是柔性链，即主链的链内旋转势垒并不太高，再加上分子内的相互作用，可以形成较稳定的螺旋构象。蛋白质中的 α -螺旋及核酸中的 DNA 等属于这一类。

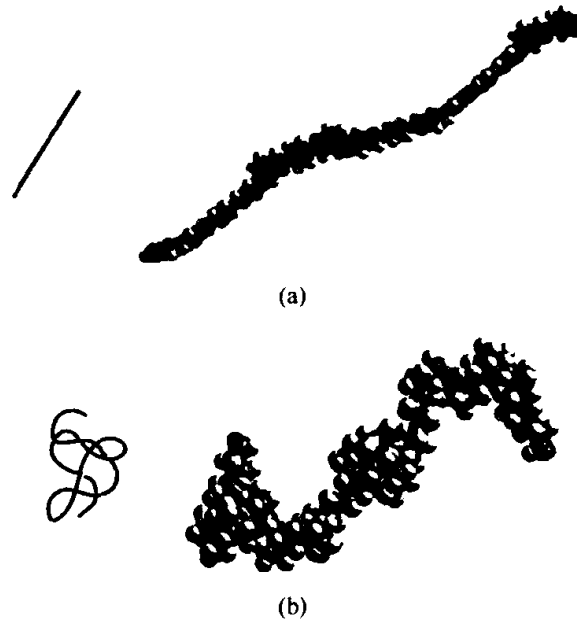


图 3-18 两种典型的聚合物长链结构
(a) 刚性链；(b) 柔性链

3.3.2 无规行走的线团模型

在本节中我们将介绍以柔性的长链分子为构造基元的无序高聚物的结构模型——无规行走的线团模型。这个模型是 P. T. Flory 于 1949 年在布朗运动粒子轨迹的启发下首先提出的。图 3-19 是 J. Perrin 于 1909 年对布朗运动粒子轨迹的观测结果。仿造液体中粒子的无规运动，Flory 将柔性的高分子长链分为若干段，每段对应于一个布朗运动粒子的运动轨迹。这样聚合物在空间任意地、不断地改变方向，按照无规行走的方式从原点进行无规行走。在行走了 N 步之后 (a_1, a_2, \dots, a_N)，可以求出从原点到所在位置的距离 r

$$r = \sum_{n=1}^N a_n \quad a_1 + a_2 + \dots + a_N \quad (3-6)$$

由于 a 方向是任意的，因此对 r 求平均值总等于零。有意义的是 r^2 的平均值：

$$\begin{aligned} \langle r^2 \rangle &= \sum_{n=1}^N (a_1 + a_2 + \dots + a_N)^2 \\ &= \sum_{i=j} (a_i \cdot a_j) + \sum_{i \neq j} (a_i \cdot a_j) = Na^2 = R_0^2 \end{aligned} \quad (3-7)$$

上式中有两类项，一类脚标相同的 a_i 与 a_i 的乘积，由于夹角为零就等

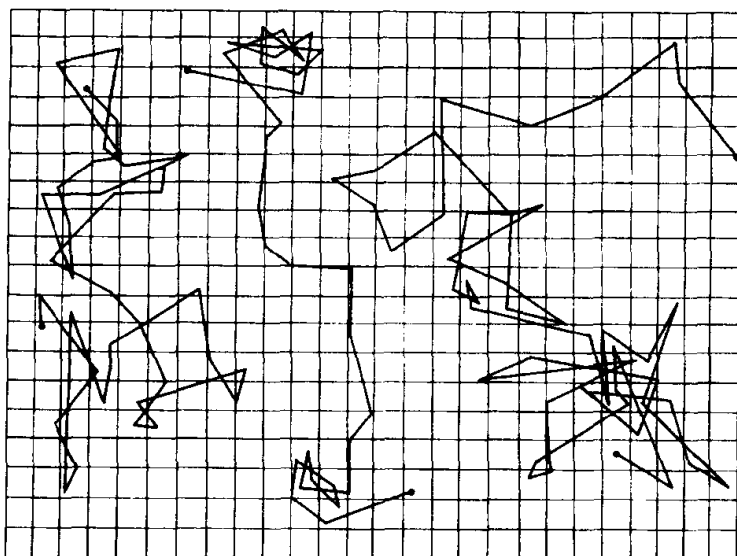


图 3-19 布朗运动粒子的轨迹 (Perrin 的实验结果)

于 a^2 ; 另一类, 脚标不同的 a_i 与 a_j 的乘积应等于 $a^2 \cos \theta_{ij}$, 这里 θ_{ij} 表示 a_i 与 a_j 的夹角, 由于 a_i 与 a_j 的方向都是任意的, $\sum_{i,j} a^2 \cos \theta_{ij} = 0$, 这样就得出

$$R_0 = N^{1/2} a \quad (3-8)$$

从式 (3-8) 可以看出, 无规线团的尺寸 R_0 与 $N^{1/2}$ 成正比。这一结果与维数无关, 对于一维、二维、三维甚至于更高的维数都能成立。

3.3.3 溶胀的线团

用无规行走的轨迹来模拟高分子的构型是一个最简单的模型。但是无规行走轨迹是可以相交的, 而高分子构型是无法与自己相交的, 应该自相回避。这样, 自回避行走 (self-avoiding walk, 简称 SAW) 模型就应运而生。它建立在一根高分子链的不同部分相互回避, 体现了所谓不可入性, 如图 3-20 所示, 圆圈部分体现了分子单体间存在相互排斥的效应。

求解自回避行走模型是一个复杂的数学问题, 比较困难, 至今尚未求得其解析解。当然, 可以采用计算机模拟的方法, 生成许多不同的构型, 然后进行统计平均。如果将 R 与 N 的关系表示为

$$R \sim a N^\nu \quad (3-9)$$

计算机模拟的结果为, 对 d 维的自回避行走

$$\nu = \frac{3}{d+2} \quad (3-10)$$

上式表明, ν 与维数 d 有关, 对 $d=1$, 既然长链分子不能与自身相交, 只能直线向前, 因而 $\nu=1$; 对于显示的三维空间中的高分子链, $\nu=3/5$ 。这

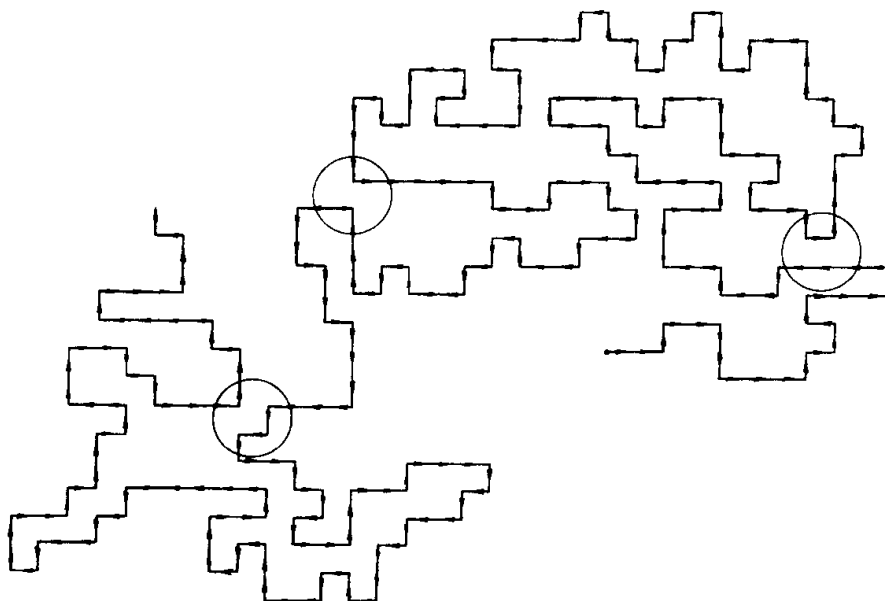


图 3-20 二维正方点阵上的自回避行走（圆圈处表明了自回避）

一结果也是 Flory 最早以唯象的近似方法导出的。

由于 $3/5 > 1/2$ ，对同样 N 个单体，自回避行走算出的线团尺寸要比无规行走的大，可以说线团胀大了。

可以将氘化的长链分子引入试样，从而用中子散射来测量无规线团的尺寸，实验结果表明，对聚合物的稀溶液而言，实验结果符合自回避行走的结果；而对于浓溶液，乃至熔体，实验结果反而和无规行走的计算结果吻合，有点出乎人的意料之外。但仔细考虑一下，这一结果并不奇怪。因为自回避导致线团胀大，是由于自回避排斥力的作用，使胀大后会降低系统的能量。在长链分子浓度甚高的熔体中，长链分子不仅受自身其他部分单体的排斥，也受到别的长链分子中单体对它的排斥。这样一来，排斥力被抹平了，线团胀大并不使它在能量上有利，因而其构形与无规行走线团模型的计算结果吻合。

总的说来，高分子的稀溶液和浓溶液（包括熔体）这两种极端情形下无规线团的构形问题，可以用上述两种模型来描述。但对居间的情形，如亚浓溶液，情况则要复杂得多。P. G. de Gennes 考虑到长链分子的线团相当于一种具有长程关联的无序体系，和相变临界点附近的涨落颇为相似，就借用处理临界现象的理论方法来解决这一问题，在此就不详细描述了。

上述的一些具有柔性长链分子的高分子材料，通称为热塑性塑料 (polymer thermoplastic)，是日常生活中经常碰到的材料。一般需要升温到其玻璃化转变温度 T_g 或熔点温度 T_m （对半晶态材料而言）以上来加工成型。表 3-2 列出了一些常用的热塑性塑料的化学结构与典型应用。

表 3-2 常用热塑性塑料的化学结构与典型应用

符 号	聚合物名称	单 体	典型应用
PE	聚乙烯(polyethylene)	$-(CH_2-CH_2)-$	包装膜、瓶, 电缆套
PP	聚丙烯(polypropylene)	$-(CH_2-\underset{\substack{ \\ CH_3}}{CH})-$	汽车零件, 引水管, 户外地毯
PS	聚苯乙烯(polystyrene)	$-(CH_2-\underset{\substack{ \\ C_6H_5}}{CH})-$	烧杯, TV 架, 玩具, 热绝缘和包装泡沫
PVC	聚氯乙烯(polyvinyl chloride)	$-(CH_2-\underset{\substack{ \\ Cl}}{CH})-$	管道, 人造皮革, 雨披, 椅套
PVDC	聚偏二氯乙烯 (polyvinylidene chloride)	$-(CH_2 \quad CCl_2)-$	屏障包装膜
PMMA	聚甲基丙烯酸甲酯(有机玻璃, polymethyl methacrylate)	$-(CH_2-\underset{\substack{ \\ COOCH_3}}{CH})-$	器械套壳, 招牌, 窗户玻璃
PTFE	聚四氟乙烯 (polytertrafluoroethylene)	$-(CF_2-CF_2)-$	电绝缘, 不粘涂层
POM	聚甲醛(polyoxymethylene)	$-(O-CH_2)-$	精密部件
PET	聚对苯二甲酸乙二酯(涤纶, polyethylene terephthalate)	$-[O-(CH_2)_2-O-CO-C_6H_4-CO]-$	纤维, 瓶
PA6	聚酰胺(尼龙 6, nylon 6, polyamide)	$-(NH-(CH_2)_5-CO)-$	纤维, 齿轮

3.3.4 交联

如果将长链高分子进行交联, 就将固化成热固性聚合物 (polymer thermostets)。交联的聚合物由于链间化学键的存在阻碍了分子链的运动, 从而增加了聚合物的耐热性。交联的聚合物不溶不熔, 除非在分解温度之上才能使结构破坏。

在天然橡胶中加入 0.5% ~ 3% 的硫, 就可以使橡胶在一定程度上交联固化, 从而极大地提高了其性能。其中, 硫的作用是在长链分子间引入许多交联点。交联后的橡胶弹性大大得到了改善, 可以拉伸到其原来长度的 10 倍, 而一般的固体在拉伸到 1% 左右就达到了弹性极限。两者的巨大差异归源于机制不同。

一般的固体作弹性拉伸时, 原子的间距增大, 固体的内能增大, 从而使自由能增大。但橡胶在拉伸时, 情况就不同。拉伸使交联点间的分子线段变直(图 3-21), 基本上不影响分子中的原子间距。弯曲的分子线团被拉直导致了分子线段的位形熵减小, 从而使自由能增大, 也就是说, 拉伸的结果使系统的有序度得到了增加。因此, 橡胶的弹性变形是和熵的变化联系在一起

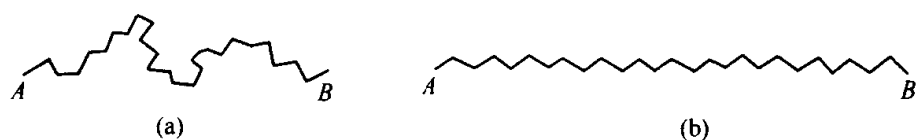
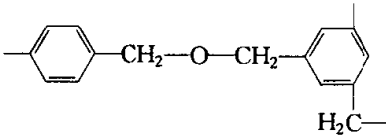


图 3-21 在交联点 A、B 之间的分子
(a) 拉伸前呈松塌状态; (b) 拉伸后呈完全张开状态

的,外力做功导致熵的减小。

由于聚合物交联固化后一些性能可以得到提高，因此，热固性聚合物的应用也十分广泛。表 3-3 列出了一些常用的热固性塑料的化学结构与典型应用。

表 3-3 一些常用的热固性塑料的化学结构与典型应用

符号	聚合物名称	单 体	典型应用
PUK	聚氨酯 (polyurethane)	$\text{—R—NH—COO—R'—OOC—NH—}$ 加 三官能团交联剂	汽车零件, 高弹体, 纤维, 泡沫
SIR	聚硅氧烷 (silicones)	$\begin{array}{c} \text{CH}_3 \\ \\ \text{—O—Si—} \\ \\ \text{CH}_3 \end{array}$ 加三官能团交联剂	高弹体, 植入物, 防水涂层
UP	不饱和聚酯 (unsaturated polyesters)	$\text{—O—(CH}_2\text{)}_2\text{—O—CO—CH—CH—CO—}$	储存罐, 船体
PF	酚醛树脂 (phenolic resins)		电绝缘, 刀柄
UF	尿素树脂 (urea resins)	$\text{—NH—CO—N—CH}_2\text{—N—CO—NH—}$	泡沫

3.4 液晶态

3.4.1 概述

一些物质的晶态结构受热熔融或被溶剂溶解之后，变成具有流动性的液体，分子位置无序，但结构上仍然保持有序排列，即分子取向仍具有长程序，从而在物理性质上呈现各向异性，形成一种兼有部分晶体和液体性质的过渡状态，这种中间状态称为液晶态，具有这种状态的物质

称为液晶。

形成液晶的有机分子通常是具有刚性结构的分子，分子量一般在 $200 \sim 500 \text{ g/mol}$ ，长度达几十个埃，长宽比约在 $4 \sim 8$ 之间。

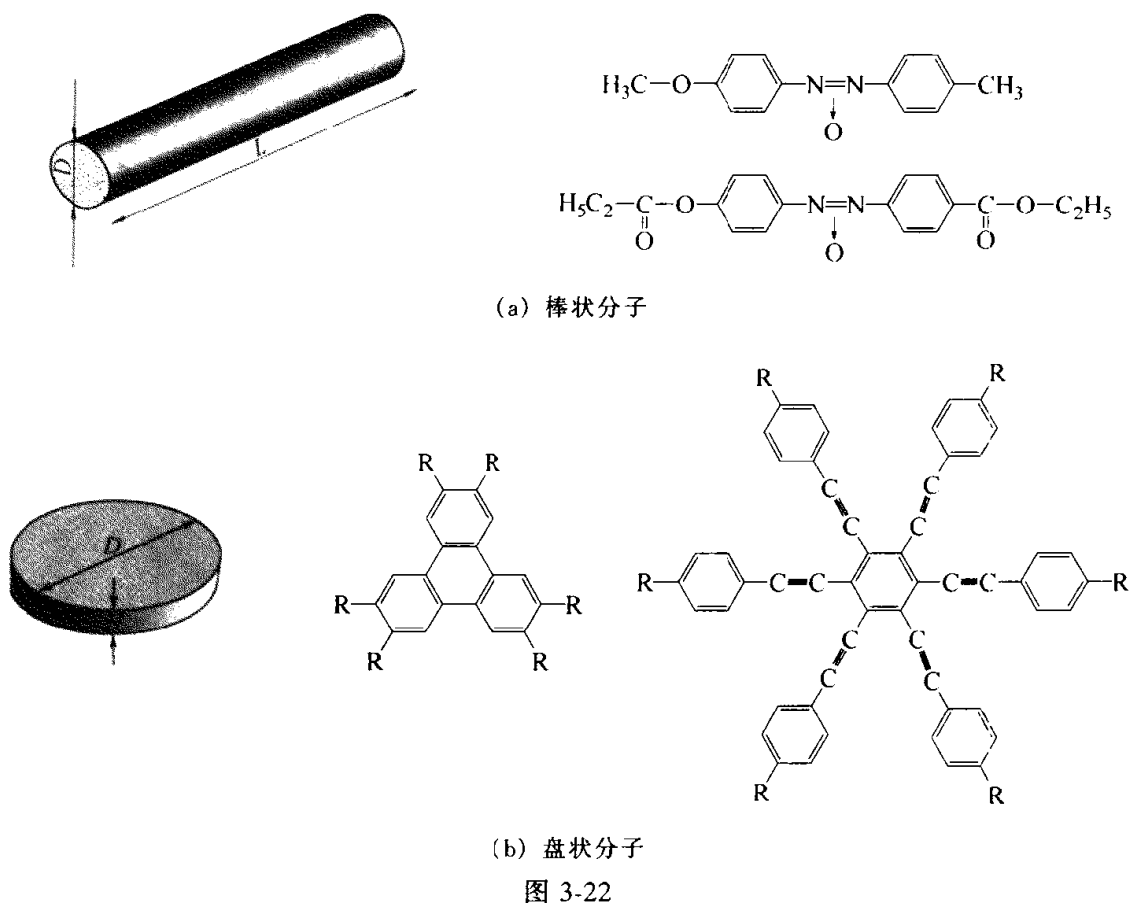
液晶态的获得一般有两种方法。一种是采用降温的方法，即将熔融的液体降温，当降温到一定程度后分子的取向有序化，从而获得液晶态，这种液晶称为热致液晶；另一种方法是有机分子溶解在溶剂中，使溶液中溶质的浓度增加，溶剂的浓度减小，可以使有机分子的排列有序，从而获得液晶态，这种液晶称为溶致液晶。

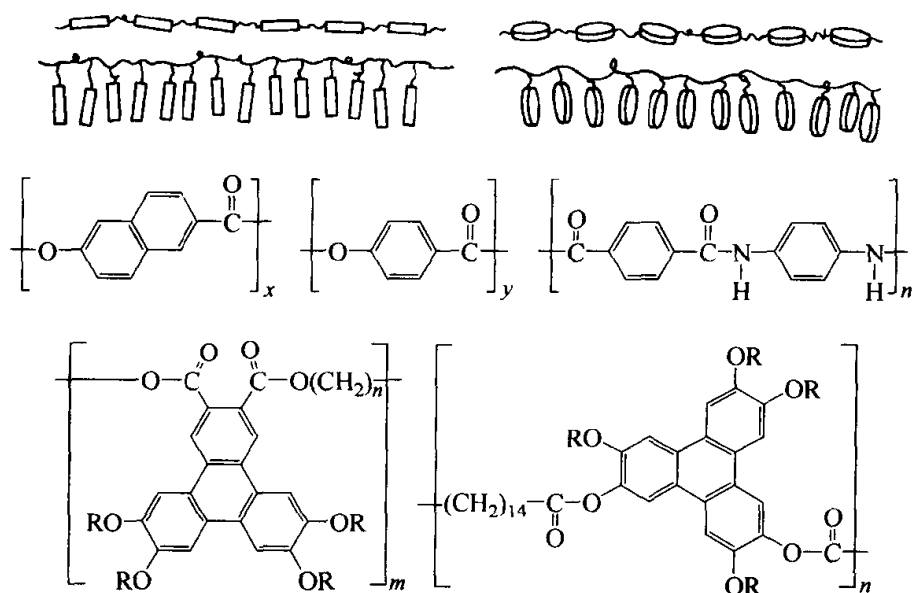
构成液晶态的结构单元大体上可分为四类：

- ① 棒状分子；
- ② 盘状分子；
- ③ 由棒状或盘状分子连接而成的（柔性）长链聚合物；
- ④ 由双亲分子自组装而成的膜。

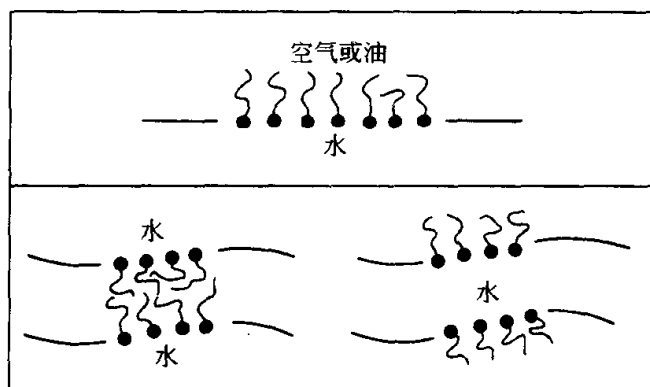
各种结构单元的形态和典型结构如图 3-22 所示。

按照液晶的对称性和有序性，热致液晶可以分为三类，即向列相、胆甾相和近晶相，如图 3-23 所示。下面将分别介绍这三种液晶。





(c) 由棒状或盘状分子连接而成的(柔性)长链聚合物



(d) 双亲分子膜

图 3-22 构成液晶态的四类结构单元

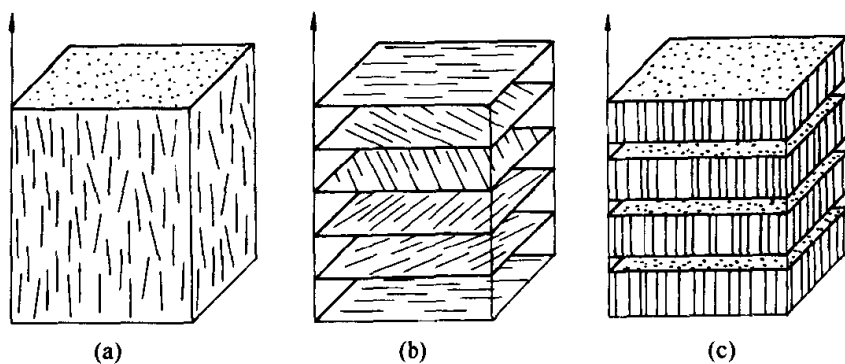


图 3-23 液晶分子的排列结构示意图

(a) 向列相; (b) 胆甾相; (c) 近晶相

3.4.2 向列相和胆甾相

向列相液晶是一类典型的液晶，它具有相当大的流动性。向列相液晶的基本特征在于一方面存在长程的方向序，分子倾向于沿特定的方向 n （指向矢）排列起来，而另一方面分子的质心的位置分布却是杂乱无章的，不存在长程的位置序，表现出液体的特征。而且，即使就分子的取向而言，也存在偏离，这里可以引入统计分布函数来描述。我们引入长程取向序参量 S

$$S = \frac{1}{2}(3\langle \cos^2 \theta \rangle - 1) \quad (3-11)$$

这里的 θ 为分子取向与向矢 n 之间的夹角， $\langle \rangle$ 符号表示取平均值。如果处于完全有序态， $\langle \cos^2 \theta \rangle = 1$ ，即 $S = 1$ ；如果处于完全无序态， $\langle \cos^2 \theta \rangle = 1/3$ ，则 $S = 0$ 。

图 3-24 给出了液晶的序参量 S 随温度的变化关系。可以看出，随着温度的下降，液晶的序参量 S 将不为零，这个变化是一个一级相变。

胆甾相液晶和向列相相似，缺乏长程平移序，也使胆甾相具有流动性。在胆甾相中，长型分子基本上是扁平的，依靠端基的相互作用，彼此平行排列成层状。但它们的长轴是在层片平面上，层内分子的排列与则向列相相似。而相邻两层间分子长轴的取向规则地扭转在一起，角度的变化呈螺旋型。因此，胆甾相的分子具有手征性，它造成了分子取向序在空间的方向不确定。由于扭转分子层的作用，反射的白光发生色散，透射光发生偏振方向旋转，使胆甾相液晶具有多彩的颜色和极高的旋光本领等独特的光学性质。

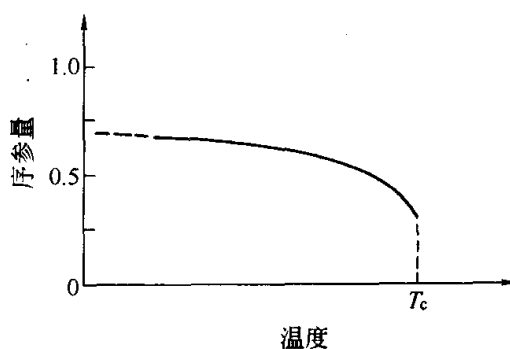


图 3-24 液晶的序参量 S 随温度的变化关系

3.4.3 近晶相和柱状相

近晶相中，棒状分子相互平行地排列成层状结构，分子的长轴垂直于层面。在层内，分子的排列方向具有二维有序性，分子的质心位置排列则是无序的，分子只能在本层内活动。而在层间具有一维平移序，层间可以相互滑动。近晶相分为两类。一类是近晶 A 相，分子的取向与层面垂直，如图 3-23(c)所示。一类是近晶 C 相，分子的取向按照特定角度与层面倾斜。

近晶相比上述两种液晶具有更高的有序性，它既有层内的二维取向序，又有层间的一维平移序。因此，描述近晶相除了要用取向序参量 S_0 之外，还要引入平移序参量 S_t

$$S_t = \langle \cos \frac{2\pi Z}{a} \rangle \quad (3-12)$$

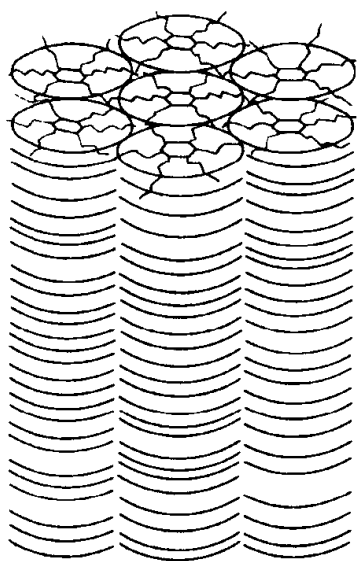


图 3-25 柱状相液晶

式中, Z 为分子重心的坐标。

如果所有分子的重心均在 $Z = 0$ 处, 则 $S_z = 1$; 如果分子的重心均匀分布, 则 $S_z = 0$ 。

柱状相液晶中, 盘状分子呈一叠叠的柱体, 具有六角形结构, 并具有二维的平移序, 如图 3-25 所示。

柱状相液晶在升温的过程中, 平移序首先消失, 接着取向序消失。近晶相的平移序消失后就变成了一般的向列相, 再变成各向同性的液体。

3.4.4 溶致液晶 (自组装膜的有序结构)

溶致液晶也是具有分子序的液相, 这和热致液晶是一样的, 但就导致液晶的相互作用以及液晶相的结构而言, 两者的差别是十分明显的。

肥皂液是一种典型的溶致液晶, 它呈现双折射现象。但这类液晶组分比较复杂, 相互作用更迂回曲折。肥皂 (学名月桂酸) 分子是一种典型的双亲分子, 它的头尾两端的亲和性截然不同。其分子头部是极性基团 $-\text{CO}_2-\text{K}^+$, 易溶于水, 具有亲水性; 而其尾部的碳氢链 $\text{CH}_3(\text{CH}_2)_7$ 则很不容易溶于水, 具有疏水性, 或亲油性。双亲分子可以构成表面活性剂, 在水溶液中, 它们倾向于群集于界面上, 亲水性集团使疏水性链不和水接触, 这样双亲分子群构成了液体内的界面——液膜。

通常晶体的构造单元是零维的原子或离子, 热致液晶的构造单元是一维的棒状分子, 而溶致液晶的构造单元则是二维的液膜。液膜本身的结构并不

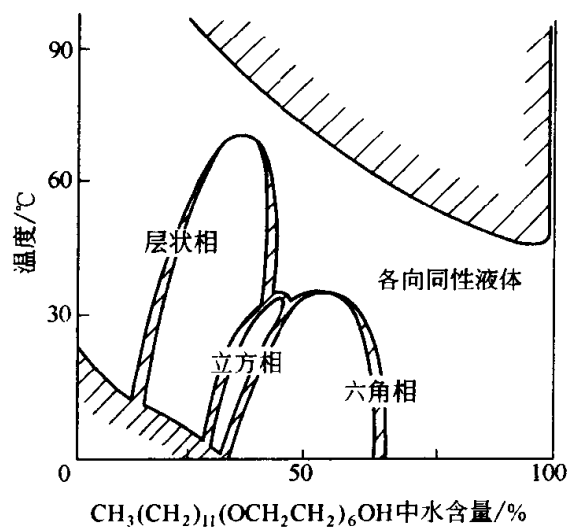


图 3-26 肥皂液的相图

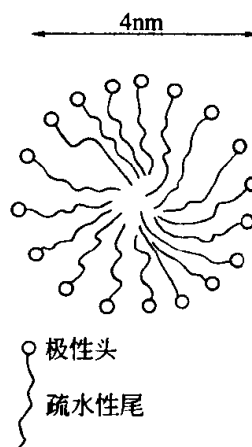


图 3-27 胶束的示意图

具有长程序，但以液膜为构造单元的溶致液晶却可能有明确的长程序。图 3-26 显示了肥皂液的相图，可以看出，随着肥皂液浓度的增大，将出现一系列不同结构的溶致液晶相。

在低浓度肥皂液中，液相呈现各向同性，双亲分子无规分布在其中并构成胶束，从而在液膜中形成了闭合的球面，如图 3-27 所示。胶束的大小和形状一般不定，和周围液体中杂散的双亲分子保持统计的平衡。如果加水稀释，胶束就会迅速消解；反之，增加溶液浓度，则会形成更加广泛的胶束，最终导致一系列溶致液晶的出现，按照肥皂浓度的次序，分别会出现层状相、六角相和立方相液晶，如图 3-28 所示。

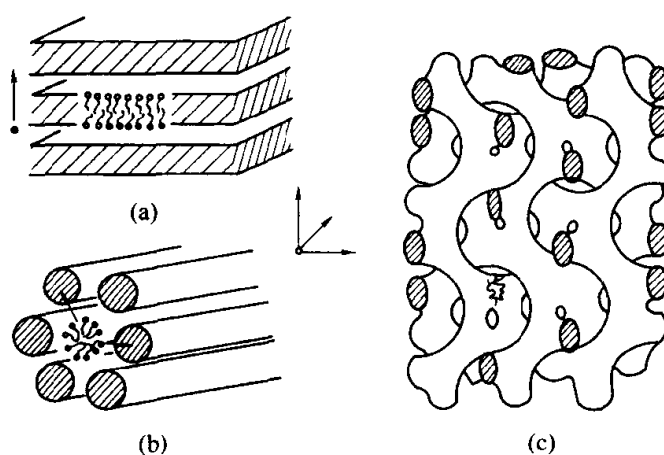


图 3-28 不同的溶致液晶结构

(a) 层状相；(b) 六角相；(c) “双连续”立方相

图 3-28 所示的三种类型的相是根据 X 射线衍射实验结果获得的。在小角度区域显示明锐的衍射线，表明在几个纳米的范围具有明确的周期结构；但在较大的角度区域却是模糊的，表明液膜本身（在十分之几纳米范围内）并无长程序。六角相和层状相都比较简单，复杂的是立方相。它是两组不相通的液膜管道，每根管道又按一定规范来分叉，像迷宫一样呈现出特殊的拓扑结构，但总体上具有立方对称性。

溶致液晶的研究属于当前研究中比较热门的软物质的自组装范畴。这种自组装过程的研究对材料的发展十分有益。构成这类物质的结构单元尺寸一般介于宏观与微观之间，在微观上分子的位置和排列是无序的，但在宏观上，却可以构成稳定的双亲界面和液膜，这类结构在生物膜中经常可以看到。因此，溶致液晶的研究不但有助于我们理解表面活性剂中的物理现象，在食品与制药工业、化妆品工业、石油回收、选矿等多种领域中获得广泛应用，而且还可以促进诸如生物膜等相关结构的研究。

3.4.5 聚合物的有序结构

很多高分子材料具有取向结构。聚合物的取向结构是指在某种外部条件下，分子链或其他结构单元沿着一定方向择优排列的结构。取向过程是一种分子的有序化过程，一般高聚物的取向有利于它的结晶。分子链的折叠就是高分子链的一种择优取向方式，图 3-29 显示了高分子薄片晶体的链折叠模型。按照这种方式结晶的区域一般不会超过 40%，总是晶化区和无规线团区共存，如图 3-30 所示。

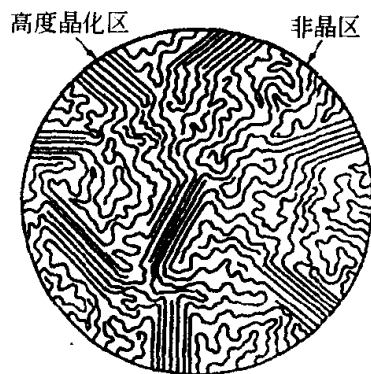
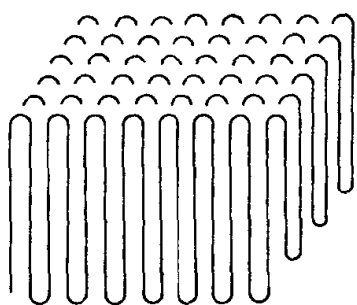


图 3-29 高分子薄片晶体的链折叠模型

图 3-30 部分晶化的聚乙烯的结构示意图

施加外力也可以使高分子择优取向，如定向结晶，定向加工，高分子溶液脱溶等。在外力作用下，无规的高分子线团将择优取向，沿着其长轴方向定向排列成一束束纤维并部分结晶，如图 3-31 所示。经过这种方式处理后，高分子可以表现出优异的力学和物理性质。

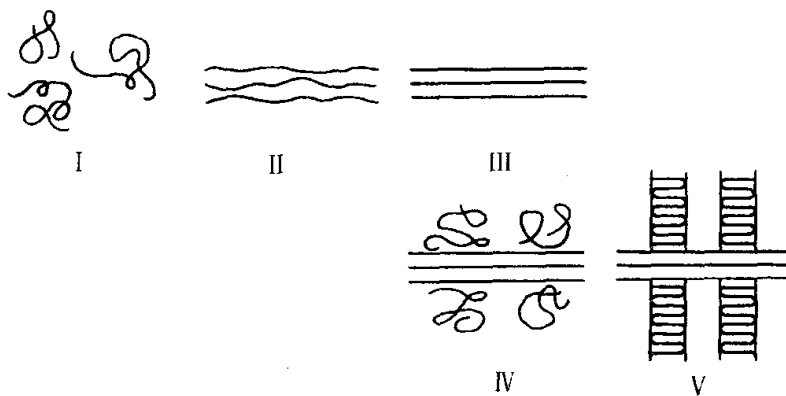


图 3-31 定向结晶的示意图

另外一种方法是形成聚合物液晶。聚合物液晶化是制备高强度聚合物材料的有效途径。在液晶态下，高分子链基本上都沿一定方向定向排列。由于高分子中的 C—C 键是强键，如果使这些 C—C 键定向排列，则在特定方向上将具有很高的强度。如聚酰胺分子在浓硫酸中形成液晶态，从硫酸中脱溶

出来后形成纤维（俗称凯芙拉纤维），其强度极高，甚至超过了金属钢琴弦，其强度密度之比则比钢丝高八倍，是目前制造防弹背心的理想材料。表 3-4 列出了部分金属丝和聚合物纤维的强度和密度。

表 3-4 部分金属丝和聚合物纤维的强度和密度

材 料	拉伸强度 $R/10^7 \text{N} \cdot \text{m}^{-2}$	相对密度 ρ	R/ρ
钢琴弦(钢)	200	7.8	26
铝丝	17	2.7	6.3
人造丝	30	1.56	19
聚酰胺纤维(商品名 Kevlar)	300	1.3	230
尼龙	100	1.15	87

3.3.1 节中已经提到了嵌段共聚物，即聚合物的长链分子是由性质不同的几段所构成的。现在考虑最简单的情况，即由 A、B 两段所构成的二嵌段共聚物。由同类分子构成的链段形成了介观尺度上的有序结构，如图 3-32 所示。这些结构和前述的溶致液晶结构很相似，这里 A、B 嵌段间的界面起了双亲分子界面的作用。而控制这些有序结构的基本参量为 A、B 两段的相对长度。如果镶嵌的段数更多一些，形成的有序结构会更丰富多彩。这样，嵌段共聚物为材料科学家制备各种尺度(从微米到纳米)的有序结构提供了新手段(参阅第 21 章)。

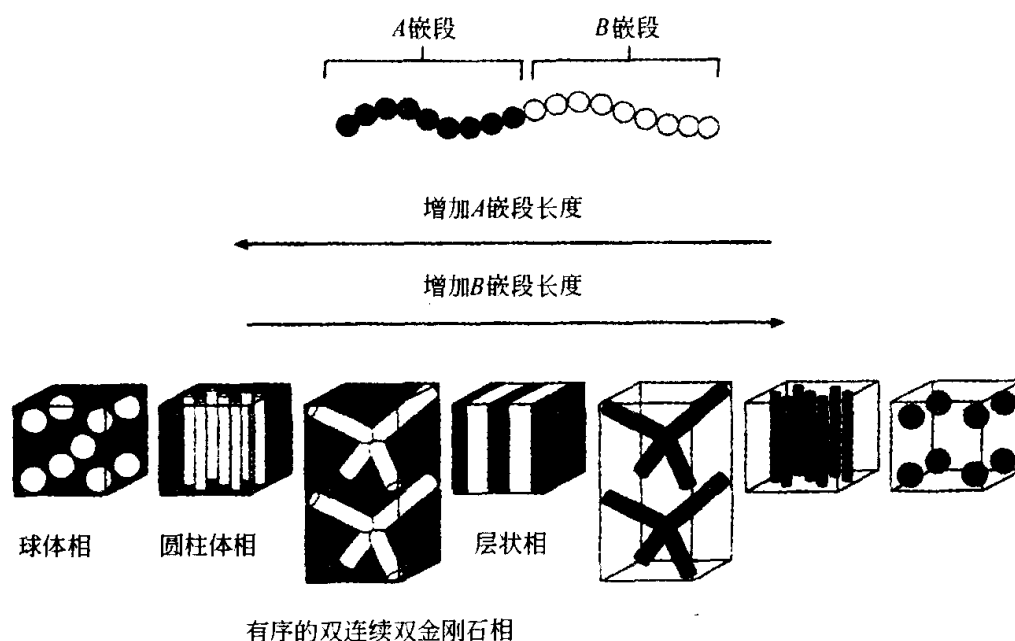


图 3-32 二嵌段共聚物的各种类型的有序结构^[10]

3.3.1 节中还提到接枝共聚物和星形聚合物，这一方面近年来进一步发展了枝状聚合物 (dendrimer)，如图 3-33 所示。从一定的中心出发，形成了

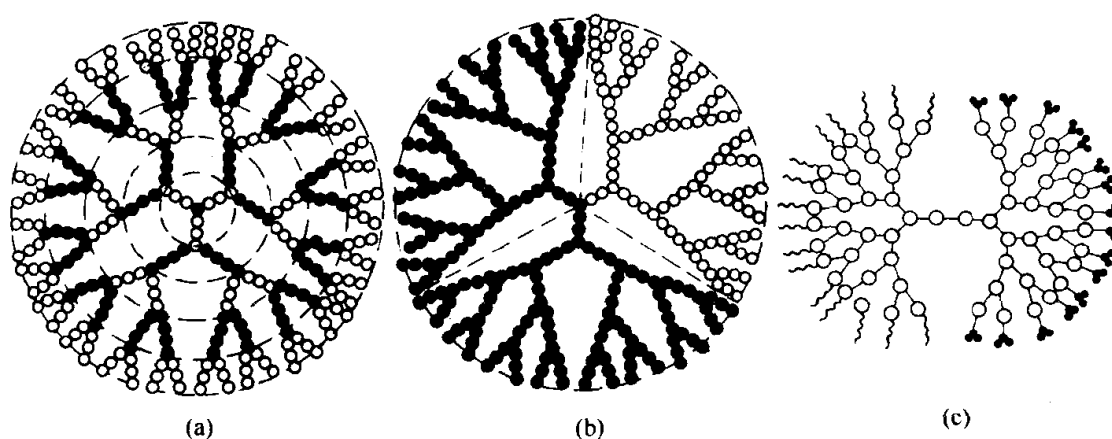


图 3-33 枝状聚合物的几种类型

- (a) 多层圈 (圆内分子结构相同); (b) 多块体 (块内分子结构相同);
(c) 双亲性 (两个半球分别具有亲水性和疏水性)

多重分枝的结构。这类枝状聚合物的合成与应用也是一个尚待开拓的领域。

高分子材料虽然已经获得了广泛的应用,但近年来对嵌段共聚物、枝杈聚合物的研究表明,通过对高分子材料的介观结构的剪裁,可以设计和制备出性能优异的新型高分子材料,因此蕴藏了巨大的研究和应用价值,尚有待于人们进一步的发掘。

参 考 文 献

- 1 Allen S. M., Thomas E. L. The Structure of Materials, New York: John Wiley & Sons, Inc., 1998
- 2 冯端, 冯步云. 放眼晶态之外. 长沙: 湖南教育出版社, 1993
- 3 Zallen. R. The Physics of Amorphous Solids. New York: Willey-Interscience, 1983 中译本: 非晶态固体物理学. 黄昀等译. 北京: 北京大学出版社, 1988
- 4 郭贻诚, 王震西 (主编). 非晶态物理学. 北京: 科学出版社, 1984
- 5 Ziman. J. M. Models of Disorder. Cambridge: Cambridge University Press, 1978
- 6 Gennes P. G. de, Prost. J. The Physics of Liquid Crystals. 2nd ed. Oxford: Oxford University Press, 1993; 初版中译本: 孙振民, 王新久译. 液晶物理学. 上海: 上海翻译出版公司, 1980
- 7 Chandrasekhar. S. Liquid Crystals. 2nd ed. New York: Cambridge University, 1992
- 8 谢毓章. 液晶物理学. 北京: 科学出版社, 1981
- 9 Gennes. P. G. de Scaling Concepts in Polymer. Ithaca: Cornell University Press, 1979
- 10 Ball. P. Made to Measure: New Materials for 21 Century. Princeton: Princeton University Press, 1997

第4章 点缺陷

撰稿人 李 齐

点缺陷，是指那些对晶体结构的干扰仅波及几个原子间距范围的晶体缺陷。晶体中最小的，也是最基本的点缺陷是点阵空位和填隙原子，它们分别相应于空缺一个原子的点阵位置和占据点阵间隙位置的原子，它们的存在破坏了晶体结构的完整性，是一种结构缺陷。晶体中的另一类基本点缺陷是杂质原子，包括代位杂质和填隙杂质，它们改变了晶体的化学成分，因而被称作化学点缺陷。

点缺陷在决定晶体的许多物理性质方面起着重要的作用，特别在控制晶体中的物质输运过程及其派生出的一些性质方面尤为突出。此外，它们通过散射导电电子影响金属的电导率，通过对声子的散射影响晶体的导热性，在半导体晶体中作为施主或受主影响电学性质，而在离子晶体中，由于在带隙中造成缺陷能级而影响其光学等性质，而离子晶体的离子导电现象则更是直接来源于点缺陷的运动。点缺陷理论正是为了解释离子晶体导电性的实验事实而提出来的。已有多本专著系统地论述了点缺陷的组态、热力学性质、动力学性质以及与各种物理性质间的关系^[1,2]。

4.1 空位与填隙原子

4.1.1 热平衡态的点缺陷浓度

晶态固体中的原子总是在其平衡位置附近一刻不停地作微小的振动。由于热振动的非线性，任一原子的热振动都与周围原子的热振动状态密切相关，使热振动能量存在涨落。当一原子一旦具有足够大的动能时就可能脱离正常位置，跳到邻近的原子间隙中去，形成填隙原子，并在原位置留下一个空位。

晶体中存在点缺陷后，内能将增加，但同时无序度增加也使熵增大了。在一含 N 个原子和 n 个空位的单原子晶体中，如果忽略空位间的交互作用，则其自由能可表为：

$$F(T, p) = F_0(T, p) + nU_v - T(nS_f + k_B \ln \Omega) \quad (4-1)$$

式中， F_0 为完整晶体的自由能， U_v 为空位的形成能， S_f 为空位周围原子振动态改变引起的振动熵，而 $k_B \ln \Omega$ 则为无序度增加而致的组态熵，

$\Omega = (N+n)! / N! n!$ 为该系统的所有可能组态数。利用斯特林近似, (即当 $a \gg 0$ 时, $\ln a! \approx a \ln a - a$), 晶体自由能表式可写作:

$$F(T, p) = F_0(T, p) + n(U_v - S_f T) + k_B T \left(N \ln \frac{N}{N+n} + n \ln \frac{n}{N+n} \right) \quad (4-2)$$

热平衡时, 满足关系 $\partial F / \partial n = 0$, 而有:

$$\frac{n}{N+n} = c_v = \exp\left(\frac{S_f}{k_B}\right) \exp\left(-\frac{U_v}{k_B T}\right) = A_v \exp\left(-\frac{U_v}{k_B T}\right) \quad (4-3)$$

可见, 任一温度下晶体中总会存在一定浓度的空位以使系统自由能降低。

填隙原子与此类似, 只是由于可能的间隙位置数可以大于正常格点数, 使无序度更大, 因而式中需增加一个与此相关的数值因子 Z_i , 使其平衡浓度变为:

$$c_i = Z_i A_i \exp\left(-\frac{U_i}{k_B T}\right) \quad (4-4)$$

4.1.2 点缺陷的形成能

易知, 这一类处于热平衡态下的点缺陷, 其浓度与温度及其形成能密切相关。首先, 点缺陷的平衡浓度随温度升高而指数式地增大, 这意味着高温下的点缺陷浓度可比室温下高出很多。其次, 由于点缺陷的形成能 (约为 1 eV) 远大于 $k_B T$ 的值 (高温下也仅约为 0.1 eV), 因而形成能的大小不仅影响着热缺陷的总浓度, 而且也决定了晶体中的主要平衡点缺陷是空位还是填隙原子。

空位的形成能被定义为从晶内正常点阵位置上取出一个原子并放到晶体表面上所需要的能量。为了不改变晶体的表面能, 如图 4-1 所示, 取出的原子应置于晶体表面的台阶处以不改变晶体的表面积。与此类似如图 4-2 所示, 填隙原子的形成能定义为从晶体表面台阶处取走一个原子并挤进晶内间

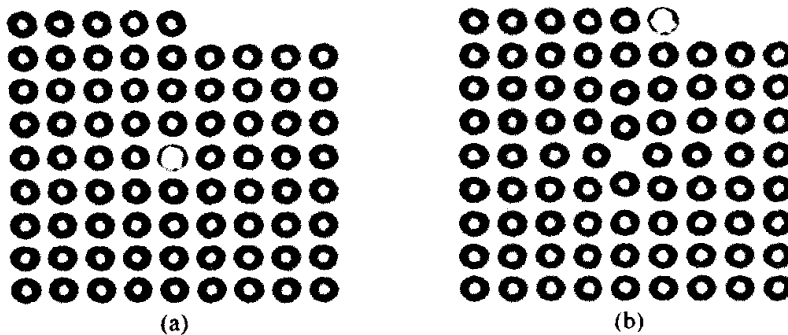


图 4-1 空位的形成能

- (a) 含表面台阶的完整晶体, 其中较浅色原子将被移走;
(b) 移走原子至表面台阶处留下的空位, 近邻原子向空位处弛豫

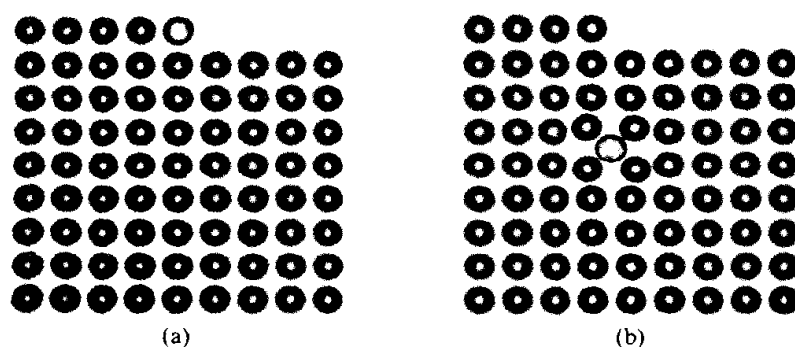


图 4-2 填隙原子的形成能

(a) 含表面台阶的完整晶体；(b) 表面台阶处的原子进入晶内
间隙位置而形成填隙原子，近邻原子被向外挤出

隙位置所需的能量。较精确地计算点缺陷的形成能，需要全面考察点缺陷的引入对电子状态的影响，以及点缺陷周围的点阵畸变等因素，因而是相当复杂的。尤其是对离子晶体中点缺陷形成能的计算中，不仅要考虑复杂的极化效应，而且要考虑正负离子在不同方向上所作的位移。大型计算机的介入使这些问题得到了很好的解决，理论模型通常还是采用 N·F·莫特 (N.F.Mott) 和 M·J·列特尔顿 (M.J.Littleton) 早在 1938 年提出的壳层模型，即将晶体分为两个区域，一个区域为包含点缺陷在内的内区，区内原子间的相互作用可表为成对原子间相互作用能的叠加，并最后弛豫至相互作用力为零的位置；另一个区域在内区之外，称外区，其间的弹性畸变用连续弹性介质模型来处理。多种方法理论计算的结果表明，面心立方金属中空位的形成能大致在 1 eV 左右，而填隙原子的形成能约为空位形成能的三倍；碱卤化合物晶体中的正负空位对的形成能约为 2.5 eV。这些均与实验值比较接近。由此也不难理解，在具有密堆积结构的面心立方金属中，空位是主要点缺陷，但在堆积密度较小的离子晶体中，由于填隙原子形成能较低，因而在较高温度下填隙原子应为主要点缺陷。

4.1.3 点缺陷的迁移激活能

由于固体的不可流动性，晶体内原子的输运过程只能凭借原子在晶内格点间一步步地相继跳动来进行。然而，理想完整晶体的所有格点均为正常原子所占满，即使是格点原子互换也不易发生。实际晶体中存在的点缺陷却给原子迁移提供了可能性。空位与近邻原子的互换、填隙原子在近邻间隙位置间的跳动、格点原子被近邻填隙原子取代而移至新的间隙位置、对分填隙组态和挤列组态〔指两个原子挤占一个正常点阵位置，和 $(n+1)$ 个原子挤占 n 个点阵位置〕与格点原子间的接力式迁移，以及诸如此类的许多点缺陷迁移方式，都将造成原子在晶内的长程移动。因此，可以说点缺陷是晶内原子

输运过程的媒介物。

点缺陷的任何一种移动方式都同时伴有相邻原子在格点间的运动，这一过程又总是在由周围原子构成的运动势场中进行，越过势垒，才能到达新的平衡位置。以空位机制为例，当移动原子到达相邻平衡位置中间时，所引起的点阵畸变显然是最大的，因而被称作鞍点组态。鞍点组态与正常空位组态之间的势能差称作空位的迁移激活能，以 g_v 表示。若近邻原子的热振动频率为 ν ，其能量超过 g_v 的几率正比于 $\exp(-g_v/k_B T)$ ，因而单位时间内可能翻越势垒的次数可表作：

$$P = A\nu \exp(-g_v/k_B T) \quad (4-5)$$

若考虑到晶内一个空位周围存在 Z 个等价的原子位置，则上式应改写为：

$$P = AZ\nu \exp(-g_v/k_B T) \quad (4-6)$$

P 也称作空位的跳跃几率。式中 A 为决定于原子迁移激活熵 S_v 的一个因子，通常为 1 的量级，故以下略去。对于金属 Cu 而言，指数前的常系数约为 10^{15} ， g_v 约为 1eV，因而在近熔点（约 1350℃）时：

$$P_m = 3 \times 10^{10}/s$$

室温（约 300℃）时：

$$P_r = 10^{-6}/s$$

显然，空位的跳跃几率十分敏感于温度，室温下几乎要 11 天才能跳动一次，如果考虑到熔点附近与室温下空位的平衡浓度相差 10^{12} 倍，室温下仅靠热缺陷迁移对晶体中相关性质的影响完全可忽略不计。

填隙原子的迁移激活能甚低，对 Cu 而言， $g_i = 0.1\text{eV}$ ，因而室温下可能有高达 $10^{11}/s$ 的跳跃几率，短时间就能移动宏观大小的距离。

事实上，点缺陷的迁移激活能的理论计算是相当复杂的。随着点缺陷在晶格中的移动，它周围的原子组态（甚至还有电子态）随之而不断变化，势场也将不断变化，因而迁移激活能的计算要比形成能的计算复杂得多。近年来用壳层模型进行计算已取得了较大的成功，甚至对碱卤化合物和结构较简单的氧化物中的各种点缺陷的迁移激活能的计算结果，也与实验值有较好的一致性。

4.2 点缺陷的产生

如前所述，一个纯净的和严格化学配比的晶体，在一定温度和压力下，由于热力学平衡条件的要求，将存在一定浓度的点缺陷。对于一般的纯净金属而言，熔点附近的空位浓度可以达到 $10^{-3} \sim 10^{-4}$ ，但在室温下却只有 10^{-12} 甚至更低的量级；对于纯净的共价半导体晶体，如纯 Si 和 Ge 中的缺

陷浓度比熔点相近的纯金属还要小得多,即使在熔点温度下,其最大空位浓度也仅为 10^{-8} 。然而,在实际晶体中,可以存在较热缺陷浓度高许多的非平衡态点缺陷,特别是在较低温度下平衡态点缺陷极少时,以各种方式引入的额外点缺陷显得更为重要,它们对材料的许多物性起着决定性作用。

以下是几种最常见的引入非平衡点缺陷的物理过程。

(1) 淬火 由前知,晶体中点缺陷的热平衡浓度随温度下降而指数式地减小。如果极缓慢地冷却晶体,则高温下平衡而低温下过量的点缺陷将可能通过合并湮灭(如空位与填隙原子的复合)或消失于晶内其他缺陷(如位错、晶界等点缺陷尾间)和晶体表面处等过程而减少,始终保持相应温度下的热平衡浓度。如果使晶体迅速冷却,即进行淬火处理,那么高温下形成的高浓度点缺陷将被“冻结”在晶内,形成过饱和点缺陷。令实际点缺陷浓度为 c ,该温度下的平衡浓度为 c_e ,则点缺陷的过饱和度可写为:

$$\eta = \frac{c - c_e}{c_e}$$

当淬火温度甚低于初始温度时,

$$\eta = \frac{c}{c_e} = \exp \left[-\frac{U}{k_B} \left(\frac{1}{T_1} - \frac{1}{T_2} \right) \right] \quad (4-7)$$

晶体中的点缺陷由于存在复合和消失过程,其实际过饱和度与淬火速率、合并湮灭速率及尾间密度等因素有关。对于金属而言,完全可以耐受 $10^4 \sim 10^5 \text{ } ^\circ\text{C/s}$ 的高淬火速率;但对非金属晶体而言,快速淬火所产生的温度梯度足以引起晶体的明显变形,甚至断裂,因而一般是不能耐受这种热冲击的。

(2) 辐照 核能的广泛应用和空间技术的迅猛发展使相关材料的辐照效应研究引起了人们的极大关注。辐照在金属与非金属晶体中的效应是明显不同的。在金属晶体中,只有将原子由其正常位置打出来的粒子才能产生点缺陷,只能激发电子的辐照则不能。而在非金属晶体中,由于电子激发态可以局域化且能保持相当长的时间,因而电离辐照就能使晶体严重损伤,产生大量点缺陷。例如被 X 射线辐照的 NaCl 晶体中, Cl^- 离子可多次电离损失两个电子变为带一个正电荷的反常离子 Cl^+ , 它为近邻正离子排斥而跑离原占格点,形成一个空位和一个填隙离子。

必须指出的是,不同类型的辐照粒子在产生离位原子从而形成点缺陷方面是相差甚大的。产生离位轰击的电子需要兆电子伏量级的能量,而中子与质子只需数百电子伏,而如果轰击粒子的质量与被轰击原子很接近,则只需数十电子伏便已足够。

点阵中的原子在被轰击以后的行为是决定辐照效应具体结果的关键所

在。如果接受的能量小于离位阈能 E_d ，将不产生离位原子，但其能量将分配给附近的一群原子，使晶格振动加强，温度升高，这种局部热点称作热峰。如果接受的能量大于 E_d ，则该原子被打离正常的点阵位置形成一个空位和一个填隙原子，这一成对的点缺陷被称作弗兰克尔（Я.И.Френкель）缺陷。通常电子辐照只能产生一对弗兰克尔缺陷，而 α 粒子的辐照在产生一对弗兰克尔缺陷后，尚余的动能已不足以产生点缺陷，但还将产生热峰，而如图 4-3 (b) 所示那样。如果辐照能量甚大，如快中子辐照，首次离位的原子能量也足够大，它和晶内其他原子相撞可以继续产生离位原子，即次生离位原子，发生所谓的级联过程，而随着能量不断损耗，离位原子的自由程逐渐变小，直至诸次生离位原子的间距到达原子间距量级时，将在较大范围内产生一大群无序状态的原子，形成离位峰，同时原路径上留下大量弗兰克尔缺陷对，而如图 4-3 (a) 所示。

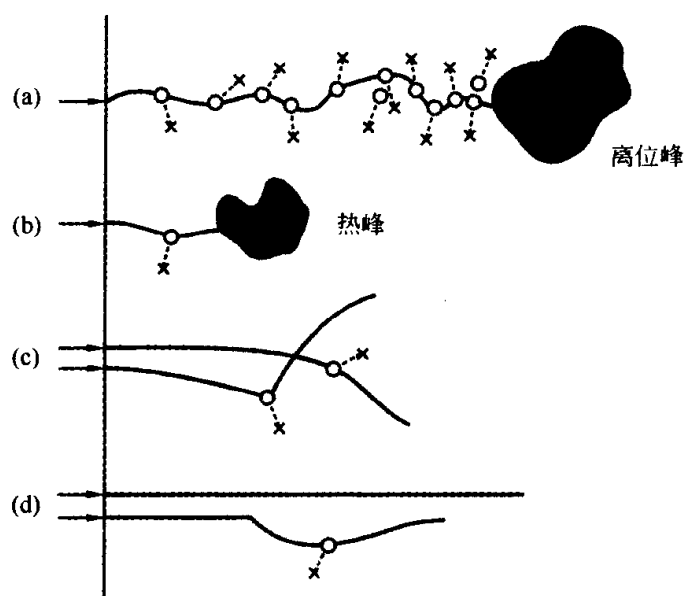


图 4-3 辐照引起的点缺陷

(a) 快中子辐照；(b) α 粒子辐照；(c) 电子辐照；(d) 光子辐照

(3) 离子注入 这是用高能离子轰击材料将其嵌入近表面区域的一种工艺。离子注入晶体中可以产生大量点缺陷。注入组分离子，产生空位和填隙离子；注入杂质原子，则产生代位或填隙杂质。在半导体器件工艺中，离子注入是引入掺杂层的有效途径。在制备某些合金材料时，不溶的合金元素只有借助离子注入技术才能实现合金化。此外，高能离子注入还能产生位错环和各种类型的面缺陷，甚至非晶层。

(4) 非化学配比 许多氧化物晶体，特别是过渡金属氧化物和变价金属

氧化物晶体，常允许其组分对化学配比的较大偏离。例如金红石晶体(TiO_2)在真空炉内还原可得非化学配比晶体，一种是在较高氧分压下处理得到大量的氧空位，一种是在低氧分压下处理则是出现大量的钛填隙。由于电导率随导电机理及导电粒子的浓度而变化，因而人们成功地利用这种现象在半导体工艺中控制化合物半导体 PbS 的导电类型。

在较高的温度和较低的非化学配比度条件下，简单点缺陷是重要的。随着非化学配比度的增高，简单点缺陷将变得不重要，它们将聚集成复合点缺陷和点缺陷群，甚至形成位错、晶体学切变面和第二相等多种类型的广延缺陷。

(5) 塑性形变 塑性形变的物理本质是晶体中位错的大量滑移。位错滑移运动中的交截过程和其他位错的非保守运动，都可能产生大量空位和填隙原子。如果温度足够低，不能发生明显的固态扩散过程的话，这些点缺陷则处于非热平衡态。

4.3 离子晶体中的点缺陷与色心

4.3.1 离子晶体中的点缺陷

离子晶体中的点缺陷与金属和共价半导体等单元素晶体不同，其差异主要来源于它的多原子性、缺陷带电及晶体保持电中性的要求。

离子晶体至少有两种带反号电荷的离子，它们各自构成离子亚点阵而有其特征的点缺陷，因而可有较多种类的简单点缺陷。

离子缺陷都是带电的。填隙离子带有本身的电荷；离子空位的存在，由于使正常电荷分布中缺少了相应的电荷而相当于空位带有相反符号的有效电荷，如正离子空位带负的有效电荷，负离子空位带有正的有效电荷。这种带电点缺陷能够捕获与其有效电荷反号的电子缺陷而改变其带电状态，例如正离子空位捕获空穴、负离子空位捕获电子而形成色心。此外，由于点缺陷有效电荷的库仑场作用，其近邻的正、负离子除分别产生极化效应外，还将向相反的方向调整其位置，以降低能量，从而形成较为复杂的组态。

一个完整的离子晶体其净电荷为零，我们说它是电中性的。如果引入点缺陷造成了过量的某种电荷，则在热力学上不利，必须同时引入包含等量而反号电荷的点缺陷，使之得到完全的电荷补偿。从另一角度看，晶体中开始总是优先产生形成能最低的点缺陷，但同时它将在晶体中建立起一个阻遏其继续产生的库仑场，而这却有利于带异号电荷点缺陷的产生，最终保持晶体的电中性。于是，在离子晶体中，具有反号有效电荷的点缺陷常常成对（或成组）出现，形成相互分开却又处于某种弱束缚状态下的点缺陷对^[3]。如图 4-4 所示，由一对正、负离子空位构成的点缺陷对称肖脱基 (W. Schottky)

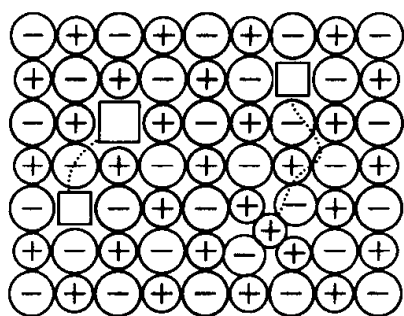


图 4-4 离子晶体中的点缺陷对
左部为肖脱基缺陷，右部
为弗兰克尔缺陷

缺陷，由一个离子空位与一个同类填隙离子构成的点缺陷对称弗兰克尔缺陷。

如果晶体中只存在低浓度的肖脱基缺陷，且忽略点缺陷间的库仑互作用时，其平衡浓度为：

$$c_s = \exp(-U_s/2k_B T) \quad (4-8)$$

式中， U_s 为肖脱基缺陷的形成能，近似为正、负离子空位形成能之和，指数中的 $1/2$ 是考虑晶体中同时产生两类同浓度点缺陷而引入的因子。

同样，如果只存在正离子或负离子的弗兰克尔缺陷，则其平衡浓度分别为：

$$c_{F^+} = Z^+ \exp\left(-\frac{U_{F^+}}{2k_B T}\right) \quad (4-9)$$

$$c_{F^-} = Z^- \exp\left(-\frac{U_{F^-}}{2k_B T}\right) \quad (4-10)$$

实践表明，除去纯银化合物，如 AgCl 、 AlBr ，由于 Ag^+ 足够小，易于形成填隙离子，因而以正离子弗兰克尔缺陷为主， CaF_2 结构的 CaF_2 和 ThO_2 等以负离子弗兰克尔缺陷为主外，其他碱卤化合物和许多简单氧化物，如 BaO 、 BeO 、 CaO 、 FeO 、 MgO 、 MnO 等均以肖脱基缺陷为主。高价金属氧化物和复合氧化物晶体则有较复杂的点缺陷组态。

4.3.2 离子导电和超离子导电现象

如前所述，离子晶体中的点缺陷都是带有有效电荷的电荷中心。在无外电场作用时，点缺陷作不规则的布朗运动，不产生宏观的电荷流动。若施于晶体一外加电场，则点缺陷的运动将择优于在电场下的有利方向进行，正离子空位朝向正极移动，负离子空位朝向负极移动。这种运动是叠加在无规热运动之上的一种规则运动，它等价于晶内离子带着本身电荷作相应运动，所造成的导电性称离子导电性。

晶体中的多种带电点缺陷在外电场作用下的运动都将对离子导电性有贡献。如图 4-5 所示，在弱场条件下，即运动势场的增（或

减）量 $\frac{1}{2} q_i E_a \ll k_B T$ 时，第 i 类带电缺陷的净流密度应为单位体积内的该缺陷数与其净

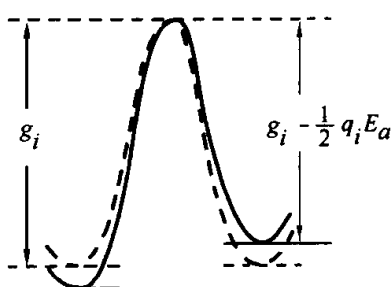


图 4-5 外电场作用下点缺陷运动势场的变化

速率之积：

$$J = c_i NV = \frac{N\nu_i a^2 q_i E}{k_B T} \exp\left(-\frac{g_i + U_i}{k_B T}\right) \quad (4-11)$$

相应的电流密度 $J_e = Jg_i$ ，因而其离子导电率为：

$$\sigma_i = \frac{N\nu_i a^2 q_i^2}{k_B T} \exp\left(-\frac{g_i + U_i}{k_B T}\right) = \sigma_{oi} \exp\left(-\frac{g_i + U_i}{k_B T}\right) \quad (4-12)$$

式中， ν_i 为点缺陷的振动频率； q_i 为所带电量， g_i 及 U_i 分别为点缺陷的迁移激活能和形成能。

实际晶体中的离子导电现象常常表现出比较复杂的特征，这是由于参与导电的点缺陷类型、浓度，以及温度等条件的差异所决定的。人们通常利用这种变化，进行系统的实验研究，并借助电子计算机技术的分析，求得包括杂质在内的点缺陷的各种参数值^[4]。

必须指出的是，绝大多数离子晶体的电导率 σ 随温度 T 的变化率都很高，室温下通常是极好的绝缘体。然而也存在一些离子导电材料，其扩散激活能非常低，较点缺陷低一至两个数量级，因而即使在室温下也有甚高的电导率，这类离子晶体被称作超离子导电材料。导致超离子导电性的结构无序大致可以分为两类。一类是某一组元离子的亚点阵存在着高度无序，包含着大量可被占据而并未被占的位置，这种松散结构相当于存在大量点缺陷（它们不是热平衡点缺陷，而是结构本身就有的），晶内离子可在这些位置间较自由地运动。AgI、RbAg₄I₅、CuI 等晶体属于此类，其中的结构无序发生在 Ag⁺ 和 Cu⁺ 的正离子亚点阵中。例如， α -AgI 晶体中，两个 Ag⁺ 离子在体心立方单胞中可占据十二个四面体间隙位置而随意跑动。具有萤石结构的 CaF₂、SrF₂、SrCl₂、PbF₂ 等晶体中，离位负离子不是固定地占据体心间隙位置，而是在其正常位置近旁大幅度的来回跑动。另一类结构无序是晶体结构本身存在一系列较松散的层或管道，使某些离子极易在其间运动。具有框架结构和钨青铜型结构的一些晶体，由于氧八面体构架成一些具有大沟道的结构，正离子，特别是碱金属离子可以在这些沟道中较自由地移动。

近 20 年来，人们已经发现了越来越多的超离子导电材料，在固态化学传感器和固态电解质方面有着广泛的应用。

4.3.3 色心^[1]

一个理想完整的离子晶体其禁带宽度甚大，例如 NaCl 晶体可达 7eV，单纯的热激发（0.1eV 量级）只能获得极少的自由电子，而在光照条件下只有紫外波段才有本征吸收，对整个可见光范围都是无吸收的，因而纯的离子晶体通常既是很好的绝缘体，也是无色透明的。

在实际晶体中，由于杂质和各种点缺陷的存在，它们附近的电子能级将

会改变, 在禁带中出现特征的杂质能级或缺陷能级。这就是如图 4-6 所示的那样, 正离子空位在价带以上产生一个已为电子所占据的能级 A , 负离子空位则在导带以下产生一个空着的电子能级 B 。它们虽然影响晶体的吸收谱, 但 A 能级至导带及价带至 B 能级的跃迁所需能量仅稍小于本征吸收, 因而纯粹点缺陷相关的吸收只对紫外区的本征吸收有影响, 晶体仍是无色的。

如果在晶体中引入了电子或空穴, 那么, 由于它们与点缺陷之间的静电交互作用, 将分别被带有正、负有效电荷的点缺陷所俘获, 形成多种俘获电子中心和俘获空穴中心, 同时产生新的吸收带。由于部分中心的吸收带位于可见光范围内, 使晶体呈现不同的颜色, 因此这类中心被称作色心。部分吸收带位于近紫外区, 虽不使晶体着色, 可也是吸收光的基因, 故统称为色心。

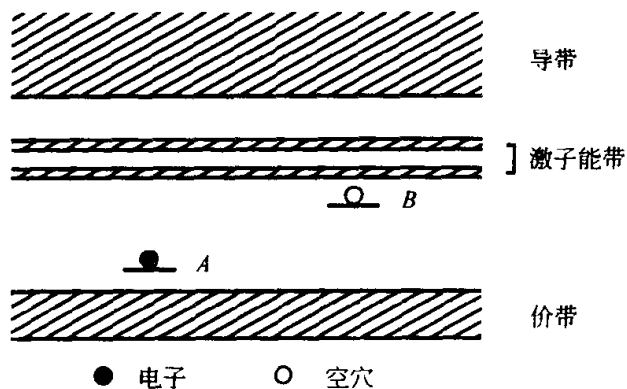


图 4-6 离子晶体的能带结构和缺陷能级

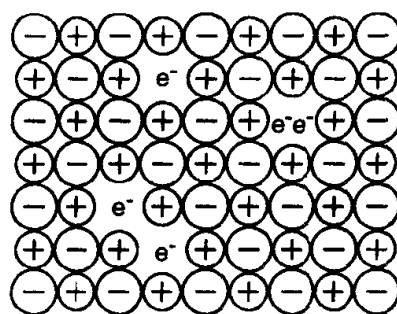


图 4-7 碱卤化合物晶体中的 F 心、F' 心和 F_2 心

最简单的俘获电子中心是 F 心, 如图 4-7 所示, 它是由负离子空位俘获一个电子所构成。光照时, F 心能级至导带间的电子跃迁因吸收光量子而形成 F 心吸收带。F 带吸收峰的能量取决于 F 心能级的位置而与产生 F 心的原因及过程无关, 因而 F 带光吸收是含 F 心晶体的最突出特征。对于用 X 射线或紫外光照射而着色的晶体, 如果以波长在 F 带内的光照射它, 则 F 心将吸收光量子而释放出被俘获的电子, 留下孤立的负离子空位, 从而使晶体退色。但是对于在碱金属蒸气中加热引入逾量碱金属原子而被着色的晶体, 由于碱金属原子离化所产生的自由电子被负离子空位俘获形成 F 心, 晶体内部已建立了新的电平衡, 即使用 F 带光照, F 心释放的电子将在晶内游离, 最终仍被失去电子的 F 心俘获, 因而不会发生退色现象。

F 心本身是电中性的, 一般说来, 它在电场作用下不发生移动。但是通过热激发或 F 带光照可以使部分 F 心离化, 释放出的电子和留下的负离子

空位可在电场作用下分别向正、负电极方向移动，因而既可表现出 F 心的宏观移动，也可造成附加的导电性，即光电导性。

在纯的离子晶体中，由 F 心为基本单位组成的电子中心可以有好多种^[5]。诸如一个负离子俘获两个电子形成的 F' 心，由两个以上 F 心紧靠在一起而构成的电子中心，被称作 M 心(F_2)、R 心(F_3)、N 心(F_4)……，它们在 F 带的低能量一侧形成一系列分列的吸收峰。一般说来，当晶内 F 心的浓度很高时，可以出现这一类多 F 心的复合中心，在适当的处理后，甚至可以形成许多原子构成的胶态粒子。

在含杂质的离子晶体中，杂质离子与 F 心间的交互作用常会形成有杂质参与的电子中心。如果 F 心近邻的碱金属离子被另一碱金属杂质离子所取代时，则形成 F_A 心，由于杂质离子的加入， F_A 心的对称性较 F 心为低，因而具有光学偏振效应。如果 F 心近邻的两个碱金属离子被碱金属杂质所取代，则形成 F_B 心。 F_A 、 F_B 心的基态 $F_A(I)$ 、 $F_B(I)$ 心与 F 心的性质差别不大，但是它们的激发态 $F_A(II)$ 、 $F_B(II)$ 心发生了组态的重大变化，如图 4-8 所示，一卤素离子占据杂质旁(F_A)或杂质间(F_B)的间隙位置，空位一分为二分布于它两侧而呈哑铃状，同时其俘获的电子也占据它两翼的势阱。由于这种结构上的差异，它们的光吸收及发射性质均发生了很大的变化。

如果将碱金属卤化物晶体置于卤素蒸气中加热，或者用 X 射线照射以后，则其光吸收谱的近紫外区将出现一系列吸收带，这表明晶内出现了与电子中心完全不同的另一类中心，即俘获空穴中心。如果我们将空穴看作电子的反型体，那么对应于各种电子中心，理应存在多种相应的空穴中心。然而在包括卤化物和氧化物在内的离子晶体中，空穴中心的确很多，但均不是电子中心的反型体，关键在于被俘获的空穴总是局域化于一个准分子态的区域内。例如，卤素亚点阵中一对相邻卤素离子俘获一个空穴构成的 V_K 心，实际上是一个卤素分子的离子（图 4-9 左部），而一系列卤素离子中插入一个卤

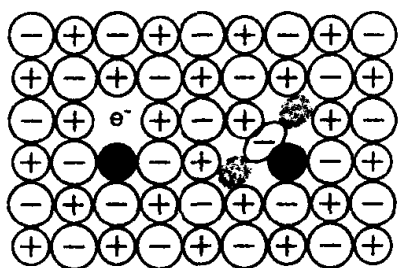


图 4-8 碱卤化合物晶体中有杂质参与的
电子中心 F_A 心
左部为基态 $F_A(I)$ 心；
右部为激发态 $F_A(II)$ 心

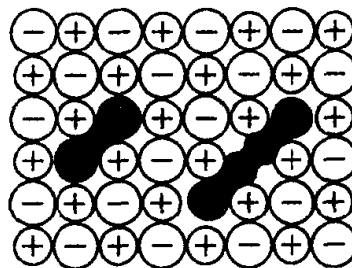


图 4-9 碱卤化合物晶体中的空穴中心
左部为 V_K 心，右部为 H 心

素原子而形成的 H 心，也可看成一个卤素分子离子占据一个正常卤素离子位置的挤列式填隙组态（图 4-9 右部）。

空穴中心还有其他许多类型，如 V_K 心近邻存在正离子空位构成的 V 心，H 心的近邻存在碱金属杂质离子构成的 H_A 心……。由于它们的吸收带能量相差无几，相互重叠而难于分辨，常需借助高分辨电子顺磁共振、核磁共振等技术，或者利用其各向异性性能进行研究。

简单氧化物晶体中的色心类型与碱金属卤化物甚为相似，但是，由于它们由双价离子组成，因而色心类型应较后者更多样化。其中俘获电子中心主要有两类， F^+ 心由氧空位俘获一个电子构成，F 心由氧空位俘获两个电子构成。后者重新建立了晶体的电中性，因而与碱金属卤化物中的 F 心是相当类似的。

4.4 杂质原子

晶体中的另一类点缺陷是杂质原子，它们是存在于晶体中而与晶体组元不同的外来原子。在非常稀的合金或固溶体中，少数组元也可看成杂质原子，但习惯称其为溶质原子，以区别于多数组元的基质原子。

杂质原子取决于其在晶体中所占位置而分为代位杂质和填隙杂质两类，前者占据组元原子所正常占据的位置，后者则是占据晶内的间隙位置。于是，前一类晶体称代位固溶体，后一类晶体称填隙固溶体。

在浓合金中，溶质原子间的交互作用会导致有序排列，通常这种有序排列仅仅包含了溶质原子的短程空间相关性。另一方面，这种交互作用也可能导致无序-有序相变，形成有序固溶体。

杂质原子在晶体中的作用主要决定于它的大小、价态与基质原子之间的差异。

对于代位杂质而言，杂质原子与基质原子的尺寸差异在晶格中引入了一个畸变中心。如果弹性畸变很小，则内能增量很小，但组态熵增加较大，因而这类代位杂质的存在热力学上是有利的。反之引入较大弹性畸变的杂质，形成代位的可能性就很小，溶解度很小。例如，在 $Nd^{3+} \cdot YAG$ 激光晶体中，由于 Nd^{3+} 半径 (0.104nm) 较 Y^{3+} (0.095nm) 大得多，不易取代 Y^{3+} ，否则将引起较大弹性畸变，降低晶体的完整性而影响其激光性能。人们采取双掺的方法，于掺入激活离子 Nd^{3+} 的同时掺入半径较 Y^{3+} 为小的 Lu^{3+} ，虽未增加 Nd^{3+} 的溶解度，但由于“尺寸补偿”效应而明显提高了晶体的完整性，使激光性能大为提高。

元素半导体晶体在绝对零度时的导带是空的，其性质像绝缘体，当温度足够高时，由于热扰动破坏了共价键，产生相同数量的电子和空穴，出现导

电性，此为本征导电性。空位、自填隙原子等点缺陷的存在可使带隙中出现相应的局域缺陷能级，对导电性有一定贡献，但纯 Si 和 Ge 中的热缺陷浓度比熔点相近的纯金属中要小得多，在不高的温度下对导电性的影响不可能很大。然而许多半导体材料是具有一定的低温电导率的，这表明它们包含着以代位或填隙方式溶解于晶体中的杂质。人们利用这一特性用掺杂方法制成了具有优异性能的非本征半导体材料；掺入Ⅲ族杂质(如 B)，在价带上方引入受主能级制成 p 型半导体；掺入Ⅴ族杂质(如 P)，在导带下方引入施主能级制成 n 型半导体。

化合物半导体中的缺陷结构和杂质效应与元素半导体中的情况大为不同。首先是非化学配比组成直接引入大量的点缺陷，特别是在Ⅱ-Ⅵ族化合物中对电导率的影响尤为明显。其次是施主能级和受主能级可以在两种组元的亚点阵中引入，例如在Ⅲ-Ⅴ族化合物半导体中，锌和镉都是代位到三价组元亚点阵中形成受主，硫、硒和碲则在五价组元亚点阵中形成施主，而四价杂质可优先代位两个亚点阵中的一个，分别起到受主和施主的作用。

在离子晶体的低温电导现象中，与基质离子同量电荷的杂质离子对电导率的贡献是很小的。但当存在异价杂质离子时，由于它破坏了晶体的电中性，将会产生带有有效补偿电荷的点缺陷。高价杂质正离子可引起正离子空位，低价杂质正离子则可产生自填隙正离子或在负离子亚点阵中产生负离子空位。同时，由于补偿点缺陷与杂质离子的有效电荷等量而反号，当温度不太高时，两者之间的库仑互作用将使它们结合成具有较强结合能的复合点缺陷。最突出的例子是在碱卤化合物晶体中掺入二价金属卤化物时，如图 4-10 所示，为了补偿 M^{2+} 的多余电荷，将在正离子亚点阵中产生一个正离子空位。一般说来，在较低温度下， M^{2+} 杂质离子将与空位形成最近邻复合点缺陷组态；而在较高温度下，则形成次近邻组态。同样的复合点缺陷在简单氧化物中掺入三价金属离子时也极易发生。这种效应使掺有异价杂质的碱卤化合物和氧化物晶体的电解电导率得到了明显增强。

异价杂质的引入引起杂质-点缺陷，尤其是杂质-空位复合体的现象，在辐照损伤了的半导体 Si 中则是另一种物理过程。在这里，异价杂质，如 P、Al、Ga、As 等作为代位杂质在贡献出相应的载流子以外，不再形成电荷补偿点缺陷。但是，辐照引入了大量空位，杂质与空位间的互作用使之成为可动空位

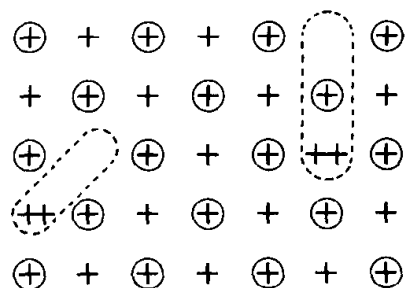


图 4-10 NaCl 结构的晶体中的正离子亚点阵中，由于二价正离子杂质加入而形成的最近邻(左部)和次近邻(右部)杂质——空位对

“+”示纸面内的正离子，

“⊕”示纸面上或下的正离子，

“++”示二价杂质正离子

的陷阱中心，最终形成空位-杂质复合体。

参 考 文 献

- 1 Crawford J. H. Jr, Slifkin L. M. Point Defects in Solid. Vols.1, 2. New York: Plenum Press, 1972
- 2 Leibfried G. Breuer N. Point Defects in Metals. Vols.1, 2. Berlin: Springer, 1978; 1980
- 3 Chiang Y.M, Birnie D.P. III, Kingery W.D. Physical Ceramics. New York: John Wiley & Sons, Inc. , 1997.146
- 4 Beumont J. B. Jacobs P W M. J.Chem. Phys. 1966, 45: 1496
- 5 Pick H. Z.Phys.1960, 159: 69

第 5 章 位错与向错

撰稿人 李 齐

位错和向错是两种典型的线缺陷，严重的原子错排仅出现于沿位错线或向错线近旁甚小的区域内，远离处只存在弹性畸变。

纯粹的弹性体内部不存在内应力，而实际上许多物体即使不受外力作用也可能存在内应力。可以这样设想，如果在连续介质内部沿任意面 S 剖开，加外应力使剖面的两岸作相对位移，去除重叠部分、空隙处填以同种材料后完整地胶合起来，然后撤去外应力，则此物体必将存在内应力。对于任意的相对位移，剖面是弹性场中的奇面，除非相对位移被加上如下的限制条件：

$$\delta U(\mathbf{r}) = \delta \mathbf{g} + \mathbf{w} \times \mathbf{r} \tag{5-1}$$

被称作 G. 韦因伽登 (G. Weingarten) 定理。其中 $\delta \mathbf{g}$ 代表刚体式的平移，而 \mathbf{w} 与刚体式的旋转相当，原点取在剖面上， \mathbf{r} 为矢径。这样的情形下，剖面上应力与应变均具有连续性而不留下任何痕迹。然而，剖面的周界处无论剖面一边的相对位移有多小，其应力也是无穷大，这是与初衷相背的。为避免周界处的应力发散，一般沿周界挖一空心管道，包含此空心管道的甚小区域就是留在材料内的位错线或向错线。其中，式 (5-1) 右侧仅有平移 $\delta \mathbf{g}$ 的构成位错线，仅有旋转 \mathbf{w} 的构成向错线。V. 沃特拉 (V. Volterra) 在 20 世纪初就提出了线缺陷的概念和模型，而当时还远未知晓材料中确实存在此类缺陷。所提

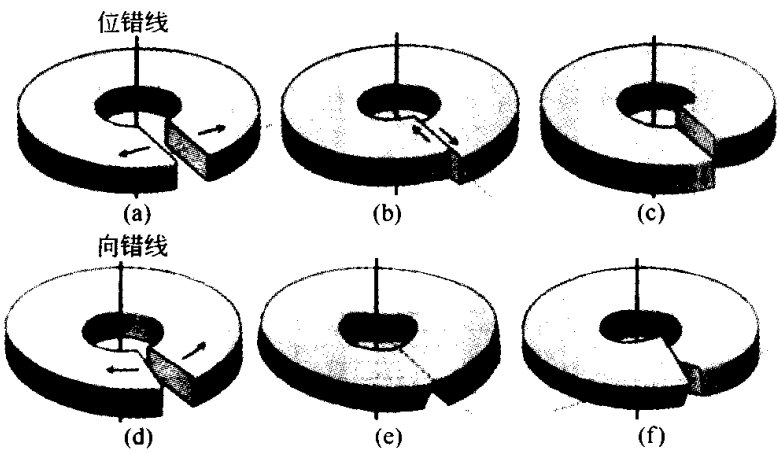


图 5-1 Volterra 的线缺陷模型

(a)、(b) 刃型位错；(c) 螺型位错；(d) 楔型向错；(e)、(f) 扭型向错

出的模型如图 5-1 所示。

首先在圆盘状连续弹性介质中挖去中心管道，并在其一侧作一径向剖面，然后以箭头所示方式让剖面两岸作相对平移或相对旋转，舍弃重叠部分或以相同材料填补空隙部分后完全粘合，则得沿轴线的线缺陷。其中与轴垂直平移的两种情况其几何特征相同，定义为刃型位错[图 5-1(a)、(b)]；与轴平行的平移得螺型位错[图 5-1(c)]。它们为位错线的两种基本类型。绕圆盘轴旋转造成的是楔型向错[图 5-1(d)]；绕与盘轴垂直的轴旋转所得的则是扭型向错[图 5-1(e)、(f)]。它们为向错线的两种基本类型。

位错是晶体中最为常见的缺陷之一，它对晶体材料的各种性质都有程度不同的影响，很早就被人们关注和研究，有了比较成熟的理论和大量的实验研究结果^[1~7]，因而是本章讨论的重点。由于晶体中的旋转对称性最多为六重对称，即使形成楔角最小的向错也得使剖面两岸作 60° 旋转，这在大块晶体中是不可能的，然而某些特殊情况下出现向错的情况已不断被人们发现，但是，在所有液晶材料中向错都是很常见的线缺陷，这些将在本章最后一节简单介绍。

5.1 晶体中位错的几何特征

以上所说在连续弹性介质中产生位错的过程，看来似乎是人为的，但它实质上是晶体中实际形成位错过程的一种模拟。例如，晶体中大量的过饱和点缺陷可以聚集成盘，空位盘相当于局部取走一层多余介质，填隙盘相当于局部填进一层介质，空位盘的崩塌和填隙盘撑开两侧晶面则相当于剖面两岸的相对位移，最终形成的是位错环（如图 5-2）。其他形成位错的过程也都有类似的形式。必须指出的是，晶体中的位错与连续介质模型中的位错有着重要的差别：介质中的位错相对平移可以是任意的，晶体中的平移则受平移对称性的限制只能是点阵平移矢量；介质中位错核心是空心管道，晶体中一般

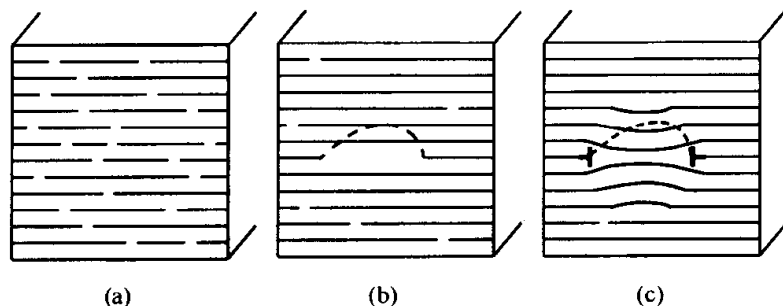


图 5-2 位错环的形成

(a) 均匀分布的过饱和和空位；(b) 空位聚集成盘；(c) 空位盘崩塌形成位错环

不是空心的而是分立的原子的特殊错排组态(见图 5-3); 介质中形成位错的剖面可以是任意形状, 而晶体中形成位错时, 常常与某些特定晶面相关。

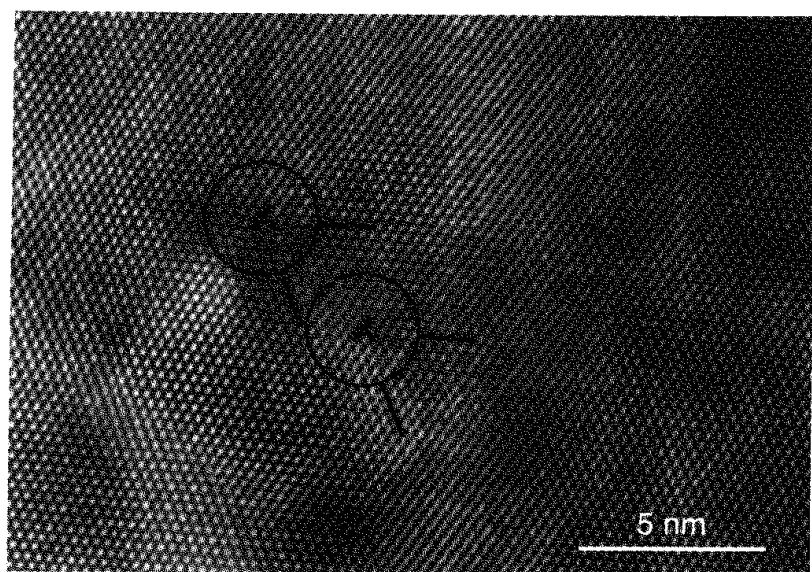


图 5-3 GaN 晶体中刃型位错的高分辨电子显微像

(引自朱健民, 李齐等, 1998, 未发表)

尽管在位错形成的许多实际过程中常有含义明确的剖面, 但最终在弹性场中并不会留下剖面的任何痕迹, 因而描述位错几何特征的要素仅有位错线的位置和相对位移矢量两者。作为标志位错特征的剖面两侧的相对位移矢量称作柏格斯 (J.M.Burgers) 矢量。如果注意到相对平移方向的不确定性, 那么在确定位错的柏格斯矢量时, 必须分外小心。通常的做法是: 首先任意选定位错线的走向, 然后循位错线方向按右手螺旋法则选定剖面法线方向, 并规定法线矢量由剖面的负岸指向剖面的正岸, 于是, 剖面正岸相对于负岸所作的平移矢量定义为相对平移矢量 δg 。这样, 位错柏格斯矢量的确定就完全取决于使用何种惯例了。按弗兰克 (F.C.Frank) 惯例, 柏格斯矢量定义为:

$$\mathbf{b} = -\delta \mathbf{g} \quad (5-2)$$

而按瑞德 (W.T.Read) 惯例, 结果相反。较多的书中取前者, 本书亦不例外。

对于如图 5-4 (a)、(c)所示的简单立方晶体中的刃型位错和螺型位错, 其相应的柏格斯矢量可以用柏格斯回路法来确定。具体做法是: 在含位错的晶体中, 取远离位错线的一阵点作为起点, 绕着位错线按右手螺旋法则连接相邻的原子作一远离位错线的闭合回路; 再在完整晶体中如图 5-4(b)、(d)所示以任一阵点为起点, 以类同的方法作回路, 此回路将不闭合。则后者由回路终点指向起点的差距矢量就定义为该位错的柏格斯矢量 \mathbf{b} 。易知, 这相当于图中的 \mathbf{FS} 矢量, 差别在于 \mathbf{FS} 在含位错晶体中, 该矢量隐含弹性畸变, 而 \mathbf{b} 在理

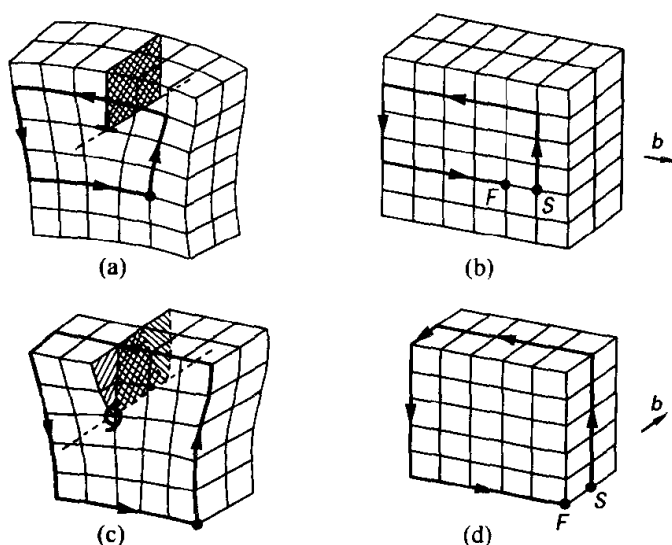


图 5-4 简单立方晶体中的位错及其柏格斯矢量的柏格斯回路作图法示意图
(a) 刃型位错；(c) 螺型位错及其闭合回路；(b)和(d)为理想晶体中的相应回路，回路起点为 S 点，终点为 F 点， \overrightarrow{FS} 为位错的柏格斯矢量

想晶体中，有明确而惟一的含义，同时这样所确定的 b 与弗兰克惯例相符。

必须指出的是，对于同一根位错线而言，柏格斯回路的上述作法并未对回路具体路线作任何限制，因此，可以断定，所有回路路线得出的结果完全相同，表明一根位错线只有一个柏格斯矢量。获此结果并不意外，实际上在产生位错时剖面两岸的相对位移是“刚性”的已蕴含了这一结果。显然，位错非直线性也并不影响这一结论。

如果位错线有分岔，或者换个说法，即数根位错线交汇于一点，那么，基于同样的理由，可以用简单的方法证明：若取交汇于一点的所有位错线的方向全部指向（或全部背向）结点，则应有：

$$\sum_i b_i = 0 \quad (5-3)$$

图 5-4(a)、(c)已分别给出了晶体中刃型和螺型两种基本位错类型的三维示意图，这样我们就能够更直观地对它们进行描述。对于刃型位错而言，其几何特征是它的柏格斯矢量垂直于位错线。正由于此，可以发现在所有包含位错线的晶面中，只有同时包含其柏格斯矢量的晶面是完整的，该晶面称刃型位错的滑移面；其余晶面都将终止于位错线上，其中垂直于滑移面的半原子面称攀移面。由此图像出发，用“ \perp ”符号代表刃型位错，长划代表滑移面，垂直的短划代表半原子面。对于螺型位错而言，其柏格斯矢量平行于位错线，而原子面则是构成了绕位错线的一个连续的螺旋面。由于螺旋面可以是右螺旋的，也可以是左螺旋的，按照弗兰克惯例，它们的柏格斯矢量前者与位错

线平行，而后者反平行，我们分别称其为正、负螺型位错，并用符号“ \oplus ”和“ \ominus ”分别表示它们。

在对晶体中位错的静态几何性质进行了充分讨论之后，我们还必须对它的运动几何性质进行简单的说明。位错概念在本章引言中是由介质中内应力源泉的考虑中引入的，而在晶体中，位错概念则是从诠释实际晶体易于塑性形变的内在原因而引入的。简单的理论计算表明，理想完整晶体中要使晶面间作整体的相对滑动，从而实现晶体的塑性形变，大约需要高达切变模量的 $1/10$ ，即 $G/10$ 的外切应力，这在实践过程中是不太可能的。于是人们设想了图5-4中的位错模型，这样的位错从现有的位置移动到相邻的类同位置，只需少数原子作一些位置调整，位错连续作这样的移动并扫过整个晶体，则完成整个晶面间的相对滑动，最终实现塑性形变。这一过程只需甚小的外切应力，在塑性较好的金属材料中，实测的临界切应力与对含位错晶体的理论计算值是相当接近的，仅为 $10^{-4} \sim 10^{-5} G$ 上下。

晶体中位错的运动有两种基本形式，一种是滑移，一种是攀移。

位错的滑移运动是局限于滑移面内的运动。如前所述，滑移面是位错线与其柏格斯矢量所决定的面。对刃型位错而言，滑移面是惟一的，通常是晶体的密排面；对螺型位错而言，包含位错线的所有晶面都同时包含柏格斯矢量，因而都可能是滑移面，在实际晶体材料中，由于晶体的各向异性，不同晶面对位错滑移的阻力是大不相同的，只有少数晶面，一般是密排或较密排晶面，才是真实的滑移面。

位错移出滑移面的运动称攀移运动。螺型位错的任何运动都是滑移运动，不存在攀移；刃型位错则可作攀移运动。图5-4(a)中的刃型位错向上部或向下部的运动都是攀移，如果移动一个晶面间距，那么前者将要有一整列原子跑掉，后者将要增添一整列原子，因而攀移运动是一种非保守运动（相应地，滑移运动是保守运动），同时伴有原子的长程扩散，一般是较为缓慢的过程。

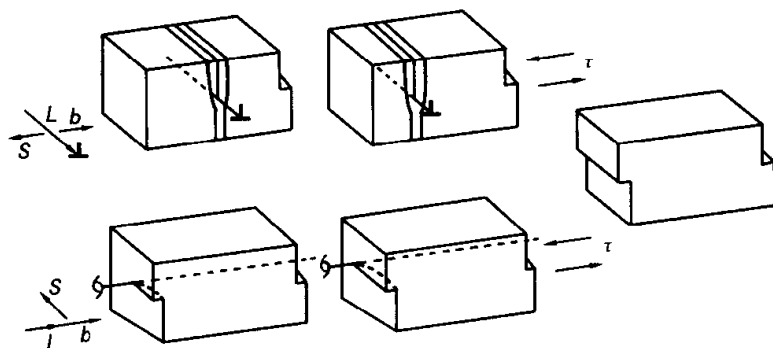


图 5-5 刃型和螺型位错滑移运动造成晶体的宏观形变

值得注意的是，同一滑移面内，柏格斯矢量相同的刃型位错和螺型位错，如图 5-5 那样，尽管其走向 (L) 是相互垂直的，两者滑移运动的方向 (S) 也是相互垂直的，但是它们所引起的滑移面两侧的相对位移只决定于位错的柏格斯矢量 (b)，因而是相同的，滑移运动所造成的晶体形变也是完全相同的。

5.2 位错的弹性性质^[5]

不管位错线周围有多么严重的局部畸变，以致不能看作是严格的虎克位移，但位错线核心区以外的所有晶面仍然是匹配的，位错所致晶格畸变可以用线性弹性理论来处理。

直线螺型位错的应力场是特别简单的。如图 5-1(c) 所示，平行于位错线的相对位移是在外径为 r_1 内径为 r_0 的圆筒中由单纯的切变产生的。对于各向同性介质而言，切应变在位错线周围是均匀分布的，平行于轴线且具有轴对称特征，因而切应变与 θ 无关，而有：

$$\epsilon_{\theta Z} = b/2\pi r \quad (5-4)$$

切应力：

$$\sigma_{\theta Z} = Gb/2\pi r \quad (5-5)$$

它们均随 r 增大而减小。

对于有限长度的介质圆筒，其内外壁上的应力为零的边界条件仍然能满足，但两端面却不能。 $\sigma_{\theta Z}$ 的作用使两端面上产生一个力偶矩，要松弛这一外力偶，介质将产生一个附加形变，相当于使中空柱体沿轴线产生扭转，其单位长度的扭转角：

$$\alpha = b/\pi r_1^2 \quad (5-6)$$

显然，除非 r_1 很小，扭转是不易察觉的。某些晶须（如 $\alpha\text{-Al}_2\text{O}_3$ 、 NaCl 、 Pd 等）中的确观测到了这种扭转，据 α 值算得的 b 值确实为点阵参数的整倍数。

直线刃型位错的情况较为复杂，但由于位移的 Z 分量为零，另两个分量不随 Z 而变，因而是一个弹性力学中易解的平面形变问题。所求得的诸应力分量，在圆柱坐标系中为：

$$\begin{cases} \sigma_{rr} = \frac{D}{r} \sin\theta \\ \sigma_{\theta\theta} = \frac{D}{r} \sin\theta \\ \sigma_{r\theta} = -\frac{D}{r} \cos\theta \\ \sigma_{zz} = \nu(\sigma_{rr} + \sigma_{\theta\theta}) \end{cases} \quad (5-7)$$

式中 $D = \frac{Gb}{2\pi(1-\nu)}$ 。在直角坐标系中为：

$$\begin{cases} \sigma_{xx} = D \frac{y(3x^2 + y^2)}{(x^2 + y^2)^2} \\ \sigma_{yy} = D \frac{y(y^2 - x^2)}{(x^2 + y^2)^2} \\ \sigma_{xy} = D \frac{x(y^2 - x^2)}{(x^2 + y^2)^2} \\ \sigma_{zz} = \nu(\sigma_{xx} + \sigma_{yy}) \end{cases} \quad (5-8)$$

图 5-6 为几个应力分量的等值曲线图。当然这里的解还不能完全满足表面应力为零的边界条件，不过一般说来附加修正项的效应并不重要，可不予考虑。邦德 (W.L.Bond) 等和英顿博姆 (В.Л.Инденбом) 等用红外偏光显微法测得的相应应力场分布确与上述结果大致相符。在图 5-7 中展示了 YAG 晶体中刃型位错应力场在正交偏光下的双折射图像，可以看出至少显示了四种不同柏格斯矢量的位错。

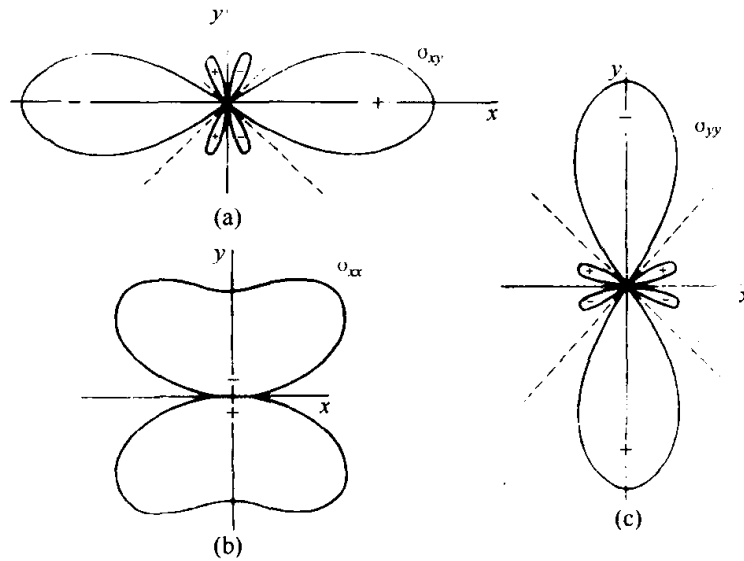


图 5-6 柏格斯矢量 b 沿 $+x$ 轴向，位错线 L 沿 $+z$ 轴向的直线刃型位错的等值应力场曲线

(a) σ_{xy} ; (b) σ_{xx} ; (c) σ_{yy}

在已知位错应力场的情况下，根据应力场下弹性畸变能密度：

$$dw = \frac{1}{2} \epsilon \sigma dV \quad (5-9)$$

其中 $dV = 2\pi r dr$ ，将之对整个晶体积分，即可求得位错的弹性畸变能。对于单位长度的位错线而言，螺型与刃型位错分别为：

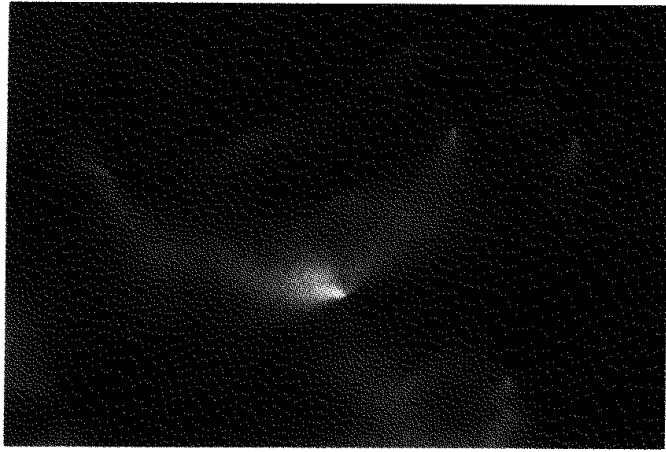


图 5-7 YAG 晶体中刃型位错在偏光显微镜中所得的
应力双折射像

注意至少有四种不同的柏格斯矢量 (引自李齐, 1975, 未发表)

$$W_s = \frac{Gb^2}{4\pi} \ln\left(\frac{r_1}{r_0}\right) \quad (5-10)$$

$$W_e = \frac{Gb^2}{4\pi(1-\nu)} \ln\left(\frac{r_1}{r_0}\right) \quad (5-11)$$

通常, 每原子长度位错线的弹性能量约为 10eV。下一节中我们还将对位错核心区的结构和能量进行讨论, 结果表明核心能仅为弹性能的 10% ~ 20%。因而通常就以弹性能的大小大致代表位错线的总能量。而如果假定 r_1 为实际晶体中“嵌镶块”或位错网的尺度, 即单个位错线应力场所展布的范围, 约为 10^{-4}cm , 则单位长度位错线的能量可近似表为:

$$W = Gb^2 \quad (5-12)$$

位错线能量与柏格斯矢量的平方成正比, 这一关系启示我们, 晶体中最常见的位错应是柏格斯矢量为最短点阵平移矢量的位错; 若有可能, 柏格斯矢量较大的位错线可以分解为两个或两个以上小柏格斯矢量的位错。此外, 由于位错线的总能量正比于其长度, 因而位错线有缩短其长度的趋势, 从而造成了位错的线张力, 通常以 $T_d = \frac{1}{2}Gb^2$ 作为其粗略估计值。

5.3 位错核心结构

在讨论连续弹性介质中位错的弹性性质时, 必须挖去其核心部分以避免弹性场的发散。晶体的情况则不同, 它由分立的原子所组成, 位错核心区内只有个别原子的位移及个别原子间的力才有物理意义, 而这些总是有限大小的, 不会出现弹性场发散的困难而无需空心管道。然而这也带给我们一些问

题：晶体中位错的核心结构如何？它又如何影响晶体的物理性质？

对位错核心区内原子确切位置的计算相当困难，对少数特殊结构晶体中特种位错类型的计算机数值计算也很难作为普遍讨论的依据。这里我们采用较定性的方法来说明一些问题。其中引入了位错宽度 2ζ 这一参量。位错宽度的概念是基于位错两侧原子沿滑移面的相对位移从零到 b 不是突变而是渐变，其主要部分局限于位错核心区内的认识，通常以相对位移量从 $b/4 \sim 3b/4$ 的区域界定位错宽度^[7]。

以刃型位错为例，其滑移面上下两侧的原子面内原子位移最显著，其一侧原子列受到压缩，另一侧原子列则是伸张。如图 5-8 所示，如果这种应变较均匀地散布在每个原子列上，则总的弹性畸变能将会较小，这一因素必将使位错宽度增大；然而这时滑移面两侧原子间的错排（指相对位移甚大）范围增大，总的错排能也将会增大，这一因素必将促使位错宽度减小。易知，当这两个效应相反的因素达到平衡时，位错宽度将被确定。

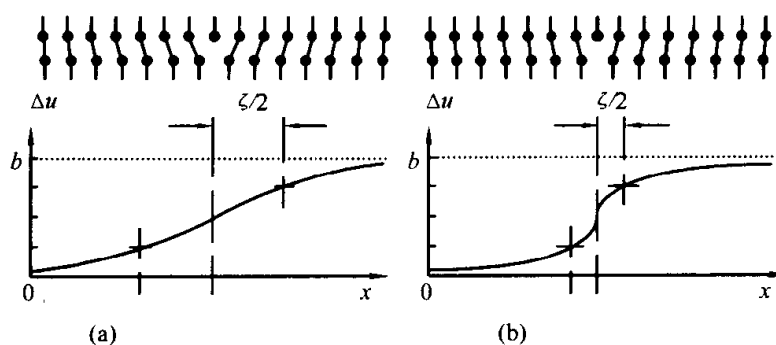


图 5-8 简单立方晶体中宽位错(a)及窄位错(b)的原子位置、错配度 Δu 及其相关宽度 $\zeta/2$

不难想像，共价键的晶体中，键的强方向性使错排能的因素较弹性畸变能有大得多的影响，因而位错较窄；而金属晶体中，金属键的无方向性使错排对位错能量的影响相对较弱，因而位错较宽。从下面对位错滑移点阵阻力的讨论，我们可以看出位错宽度对晶体塑性性质的重要影响。

位错易动是位错理论的基础，但并非运动无阻力。位错的弹性能量不随位错位置而变，因而不造成对位错滑移运动的阻力，但位错核心的错排能可随位置而变，从而引致对位错滑移的阻力。设想简单立方点阵中的一直线刃型位错，如图 5-9 所示，开始时处于对称的平衡位置。当位错相对于点阵向右侧滑移时，如果以 ab 记其滑移量，则在此期间将经由不对称位置到达 $a = 1/2$ 的另一对称位置，再经由不对称位置而到达 $a = 1$ 的平衡位置，并以此规律运动下去。显然，位错的错排能应是滑移量的周期函数，其周期为 b 。

派耳斯与纳巴罗 (R. Peierls, F. R. N. Nabarro) 提出了一个关于刃型位错

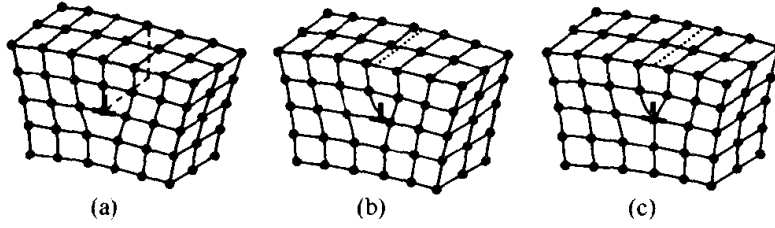


图 5-9 在晶体中滑移的位错核心组态的变化

(a) $a=0$; (b) $a=1/4$; (c) $a=1/2$

的较粗糙的点阵模型，具体计算了滑移面两侧原子相对位移的解析表达式：

$$U_x = -\frac{b}{2\pi} \arctan \frac{x}{\zeta} \quad (5-13)$$

其中 ζ 为位错的半宽度，其值为：

$$\zeta = \frac{a}{2(1-\nu)} \quad (5-14)$$

以此为出发点，计算得位错错排能随其在点阵中位置的变化：

$$W_{AB} = \frac{Gb^2}{4\pi(1-\nu)} [1 + 2e^{-\frac{4\pi\zeta}{b}} \cos 4\pi\alpha] \quad (5-15)$$

第一项与 α 无关，即 $W_0 = \frac{Gb^2}{4\pi(1-\nu)}$ ，可以作为错排能的近似值，较位错弹性能 Gb^2 小得多。第二项绝对值虽然较小，但这是位错位置的周期函数，当位错沿滑移面移动时，将通过一系列势能谷、峰的位置，其峰谷差值就是单位长度位错线滑移运动的激活能：

$$\Delta W_{AB} = \frac{Gb^2}{\pi(1-\nu)} \exp\left(-\frac{4\pi\zeta}{b}\right) \quad (5-16)$$

也称作派耳斯能，它是受位错半宽度 ζ 所制约的。位错在翻越势垒时所受克服的阻力为：

$$F_M = \frac{2Gb}{(1-\nu)} e^{-\frac{4\pi\zeta}{b}} \quad (5-17)$$

也称作派耳斯力。在 $a=b$ 的点阵中，若 $\nu=0.3$ ，则 $F_M = 3.6 \times 10^{-4}G$ ，较理论屈服应力小得多，较近于实测的 $10^{-5}G$ 。可以看出， ζ 可以引起 ΔW_{AB} 和 F_M 甚大的甚至是数量级的差异，此外，位错柏格斯矢量的大小对此也有重大影响。一般金属的位错较宽，共价晶体和含较强共价成分的晶体中位错较窄，这就是前者塑性好，后者只在高温时才有一定塑性的物理根源。许多具有复杂结构的氧化物晶体，如石榴石、铈酸锂等甚至在高温也很难滑移，则是由于它们点阵参数大，位错具有大柏格斯矢量的缘故。

位错错排能随位错核心位置的周期性变化，还使位错线倾向于躺在滑移

面内错排能最低的位置，通常这是晶内的原子密排方向。在实际晶体中，由于种种原因，位错线并不一定沿密排方向成直线形，这时，为了降低其总能量，如图 5-10 所示，位错线上可出现一系列弯结，只于弯结处处于能峰位置，以保持位错线的大部分处于能谷位置。由于弯结的存在，位错线的整体运动可以通过弯结沿位错线的运动来完成，而这时弯结是在能峰上的移动，并不明显改变位错的总能量，因而位错借助于弯结而滑移可进一步减小点阵阻力，从而更接近实际晶体的实测值。

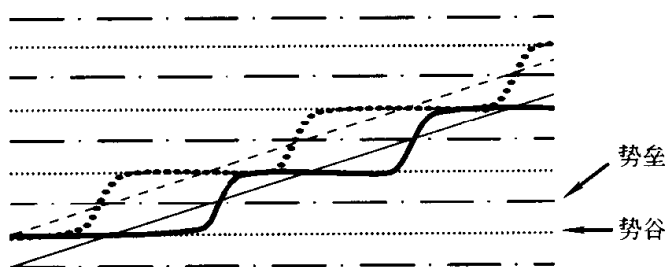


图 5-10 滑移面内与派耳斯-纳巴罗势垒斜交的位错线上的弯结（粗实线）

（注意弯结的侧向运动造成位错线的法向运动）

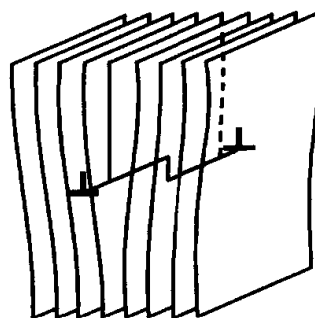


图 5-11 攀移面内的割阶

当位错线偏离滑移面时，如图 5-11 所示，位错线上将出现一系列割阶，位错的攀移运动也可以借助于割阶沿位错线的运动来实现，大大降低攀移运动所需的驱动力。

最后，我们还须注意的是非金属晶体中的位错带电问题。共价半导体中位错处原子配位数变化而出现的悬键，相当于一个受主，位错核心处的一整列悬键显然将极大地影响半导体材料的电学性质。而在离子晶体中，刃型位错的半原子面若终止于同类离子列，则位错将带有高密度的同类电荷，也将对晶体的电学和光学等性质产生重要的影响。

5.4 位错与其他缺陷之间的交互作用

位错与位错间，位错与其他晶体缺陷，如空位、填隙原子、溶质原子、自由表面及面缺陷间，可以通过弹性场或静电场发生交互作用，使其对晶体物理性质的影响变得更为复杂，因而是我们诠释许多物理现象所必须了解的。

位错与缺陷弹性交互作用的问题可以看成是缺陷应力场中的位错受力问题。当含位错晶体受到足够大外应力作用时，位错将发生滑移或攀移运动，使晶体发生形变，我们说外应力对晶体作了功，而位错则相当于受到了一个力的作用。如果将此所作之功除以位错运动的距离即可求得该力的大小。设想柏格斯矢量为 \mathbf{b} 的位错线元 $d\mathbf{l}$ 受到滑移面内沿 \mathbf{b} 方向的切应力 τ 的作用而

向前运动了 ds 的距离, 则晶体的切位移为 $\left(\frac{dsdl}{A}\right)b$, 该面 A 受力 $A\tau$, 故所作之功为

$$dW = A\tau \left(\frac{dsdl}{A}\right)b = \tau b dsdl \quad (5-18)$$

如前所述, 这也等于位错线元 dl 滑移 ds 距离中位错所受力所作之功, 因此作用于单位长度位错线上的滑移力为

$$F_s = dW/dsdl = \tau b \quad (5-19)$$

它不是位错处原子实际所受的作用力, 而是一种组态力, 它沿整个位错线都与线方向垂直, 且指向位错滑移面的未滑移方向。

同样, 应力场作用于单位长度位错线上的攀移力

$$F_c = \sigma b \quad (5-20)$$

决定于作用于位错攀移面上的正应力分量 σ 。

下面我们分别简单讨论位错间、位错与溶质原子间, 位错与自由表面间的交互作用。

5.4.1 位错间的交互作用

平行位错间的交互作用最为简单, 且各处相同。两平行螺型位错间的相互作用力为

$$F_r = \pm \frac{Gb_1b_2}{2\pi r} \quad (5-21)$$

这是一个中心力型的作用力, 同号位错取正值为斥力, 反号位错取负值为引力。当位错可自由滑移时, 或相互排斥至无穷远, 或相互吸引以致合并消失, 不存在平衡位置。

两平行刃型位错, 当柏格斯矢量同平行于 x 轴, 则所受作用力的滑移力 F_x 和攀移力 F_y 分量分别为:

$$\begin{cases} F_x = \frac{Gb_1b_2}{2\pi(1-\nu)r} \cos\theta \cos 2\theta \\ F_y = -\frac{Gb_1b_2}{2\pi(1-\nu)r} \sin\theta (1 + \cos 2\theta) \end{cases} \quad (5-22)$$

两分量均随 θ 有较复杂的变化。对于同号位错, 滑移力 F_x 仅在 $\pi/2$ 和 $3\pi/2$ 处为稳定平衡位置, 攀移力 F_y 无稳定平衡位置。如果允许位错自由运动, 那么它们将通过滑移和攀移运动, 最终变为沿 x 面上上下下平行排列的组态。对于反号位错, 则相互吸引而合并, $b_1 = b_2$ 时则互毁。如果两位错的柏格斯矢量相等但相互垂直, 则其平衡位置在与其滑移面成 45° 的晶面内。

两根非平行位错间的作用要复杂得多。由于位错各部分受力不同, 允许运动时, 空间交截而并不相交的两位错将在接近处附近发生扭曲, 直至位错

线张力与所受作用力达到平衡为止；相交的两位错在交结点附近扭曲后将趋于平行而发生位错反应，形成新的位错线段，一个结点变为两个三岔结点（见图 5-12）。

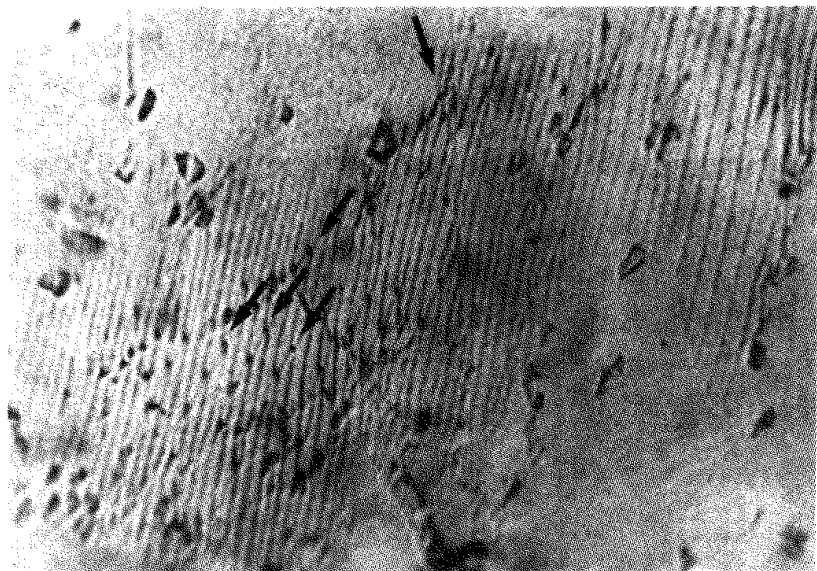


图 5-12 Mo 单晶体中与一系列平行刃型位错相交的位错，
其交点经反应变为一段新位错
(引自李齐，冯端，1964，未发表)

5.4.2 位错与溶质原子间的交互作用

位错与溶质原子之间的交互作用主要来源于溶质原子与基质原子间体积差引发的弹性畸变与位错间的弹性交互作用；在离子晶体中异价溶质原子与带电位错之间还存在电性交互作用，但通常较前者要小得多。处理弹性交互作用的简单模型是：设想在连续弹性介质中的球形空洞中填入半径略有不同的球体。如果这一过程在位错应力场中进行，则位错应力场将做功，其负值就是两者的交互作用能。参与做功的正应力分量在球体表面造成的流体静压力的平均值为 $\frac{1}{3}(\sigma_{xx} + \sigma_{yy} + \sigma_{zz})$ ，它所引发的球体积变化为 $4\pi r^3 \epsilon$ 。由于螺型位错无正应力分量可不予考虑，假定刃型位错位于原点，滑移面角坐标为零，溶质原子位于 (R, α) 处，则两者的交互作用能为：

$$\Delta U = \frac{4}{3} \frac{1+\nu}{1-\nu} \frac{Gb\epsilon r^3 \sin \alpha}{R} \quad (5-23)$$

易知，当溶质原子大于基质原子，即 $\epsilon > 0$ 时，在刃型位错的压缩区 ($\pi > \alpha > 0$)，交互作用能为正值而被排斥，在膨胀区为负值而被吸引，当其可能运动时，终将被吸引至膨胀区内。为了进一步降低交互作用能，溶质原子将聚集到位错线核心区附近形成所谓的科垂耳 (A.H.Cottrell) 气团，其浓度分布可

表成：

$$c = c_0 \exp(-\Delta U/k_B T) \quad (5-24)$$

式中， c_0 为溶质原子的平均浓度。如果位错近旁局部区域的浓度超过了溶解限，还将有沉淀物析出，我们说位错被缀饰了。在实践中，人为地将某些杂质向晶内扩散以缀饰位错，是对透明晶体中位错进行观测的一种有效方式（见图 5-13）。

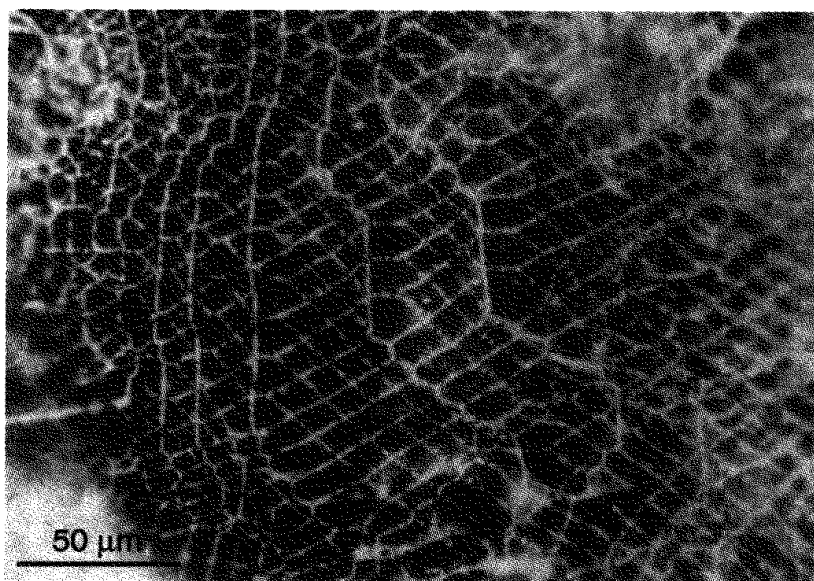


图 5-13 KCl 晶体中的被缀饰的位错网络

（引自 Amelinckx S. Acta Metall. 1958, 6: 34）

必须指出的是，实际晶体，尤其是结构较复杂的晶体，由于弹性各向异性，螺型位错也存在正应力分量，此外，溶质原子的应力场也可以是非球形对称的，因而螺型位错与溶质原子之间也会有明显的弹性交互作用，甚至形成气团。此外，由于位错应力场对弹性偶极子的取向效应，填隙原子倾向于处于低能量位置上而形成斯诺克（J.Snoek）气团。

5.4.3 位错与自由表面的交互作用

一个靠近自由表面的位错，其弹性畸变场是不完整的，而当它越发接近表面时，其应力场不断地被表面所切掉，弹性畸变能也不断地下降，这从定性上可以推断存在着一个将位错拉向表面的力。因此，位错与自由表面的交互作用问题类似于电介质中点电荷与表面的交互作用，不妨仿照静电学中的镜像法来处理^[3]，自由表面对位错的作用力亦称像力。

像力的大小可以这样近似地估计：设想在表面的另一侧晶内位错的镜像位置上存在一反号位错，如果这时是平行于表面的螺型位错，那么两位错应力场在表面处的作用相等而反号相消，完全满足自由表面应力为零的边界条

件，因此位错与自由表面的交互作用便转化为一对异号位错间作用力的问题。如果位错不与表面平行，或不是纯螺型位错，那么尽管边界条件不能完全满足，但其主要部分是满足的，因而仍可近似地用镜像法来处理。对于与表面平行的位错而言，表面作用于单位长度位错线上的像力应为：

$$F_i = -\frac{Gb^2}{4\pi kl} \quad (5-25)$$

式中， l 为位错与表面的距离， k 是如前所述决定于位错类型的常数。易知，这是一个与 l 大小成反比的吸引力。

像力的存在导致了一个重要结果，这就是接近表面的位错有与自由表面正交的倾向。人们对熔体生长或水溶液中培育的晶体中位错组态的观察结果证实，位错的确倾向于垂直于生长界面，且随界面形态的变化而改变走向。人们利用这一性质，采用凸锥形界面法（见图 5-14）使位错向周围发散而消失于侧表面，大大降低了晶体中的位错密度，然后改用平界面法使生长条件稳定而获得高完整性高均匀性的晶体材料。

位错与表面的交互作用还有其另外的一面，那就是露头于表面的位错对晶体生长和晶体解体过程的作用。垂直地露头于表面的螺型位错在其周围表

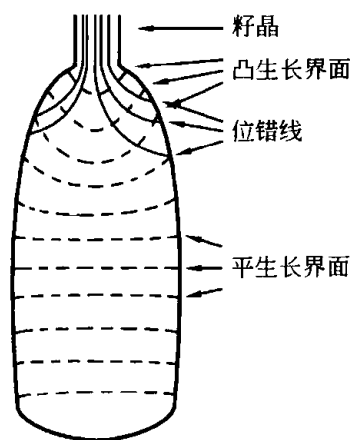


图 5-14 熔体生长晶体中凸锥形界面法消除籽晶遗传位错线的示意图

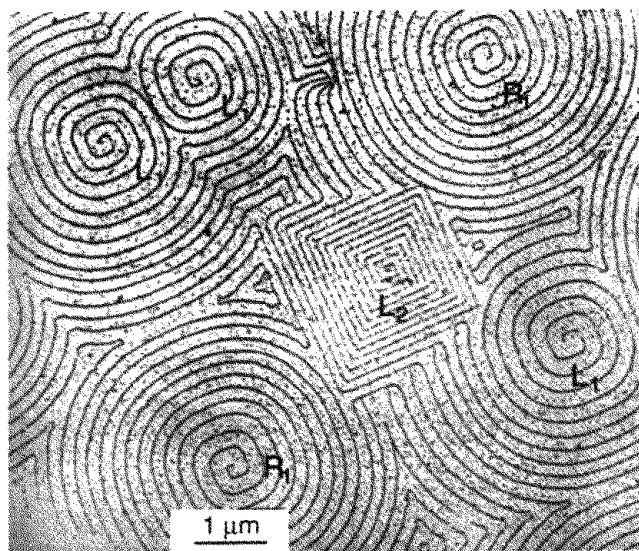


图 5-15 以 Au 粒子缀饰显示出的 NaCl 晶体表面的生长蜷线
L 和 R 分别相应于有左螺旋及右螺旋台阶的两类反号螺型位错， L_2 的台阶高度为 L_1 及 R_1 的两倍。表明其柏格斯矢量的表面法向分量的差别^[8]

面造成了一个高度为 b 的螺旋线形台阶，提供了晶体生长过程中一个永不消逝的台阶，从而使其在甚低的过饱和度或过冷度下就能有足够大的生长速度，这就是实际的晶体生长过程中极常见的螺型位错生长机制，以及许多晶体表面上出现螺旋线形生长台阶^[8]的物理原因（见图 5-15）。其实，即使是刃型位错或混合型位错，只要它们的柏格斯矢量在表面法向有足够大的分量，也将促进晶体生长。对于晶体的解体，如化学侵蚀过程，位错露头处也同样提供了优先解体的位置，因而也成为显示位错的侵蚀法的依据之一。

5.5 位错的产生和增殖

5.5.1 位错的产生

在完整晶体中产生位错，需要在晶体中沿某原子面（一般是密排原子面）的某一区域使其两岸作一相对位移。如果这种相对位移是滑移，那么实现这一过程所需的切应力常需接近理论屈服强度的大小。对于大多数面心立方金属而言，这大概是 $G/30$ ，对于体心立方金属和许多简单化合物（如 NaCl 、 Al_2O_3 等）则是 $G/10$ 左右，因而在完整晶体中哪怕是成核小的滑移位错环也几乎是不可能的。

实际晶体中，位错通常是在晶体生长过程中，或是在非均匀形变过程中产生的。

在晶体生长时，要求大量的结构单元——原子或分子完全无误地排列成完整晶体是不太容易的。当原子或分子堆砌偶尔出现错误时，缺陷，有时是位错就可能形成了。单晶硅中出现的呈旋涡状分布的大量微缺陷曾经困扰人们很长时间，经过较长期的实验研究，人们发现了多种类型的小尺度晶体缺陷，其中位错环占有较重要的位置。

非均匀塑性形变可能是晶体中引入位错的最有效途径了。塑性形变发生时，部分位错跑出晶体造成宏观形变，部分位错相互纠结而陷埋在晶体内部。这种位错的数量和分布常受一些偶然因素的影响而难于作一般性的分析，但如果从不均匀塑性形变的协调性问题出发，就可能较确切地估计这类协调位错的数量和分布。

棱柱挤压和塑性弯曲产生的协调位错是已被透彻地研究过的两类典型例子。

当一平头柱状硬质压头压进晶体表面时，压头下方产生强烈形变，这时可以通过产生一定数量的间隙型棱柱位错环的方式来松弛其畸变。如果压印深度为 Δx ，且压进方向平行于晶体中位错的柏格斯矢量时，协调位错环的数量显然为：

$$N = \Delta x / b \quad (5-26)$$

这种棱柱挤压不仅在研磨、抛光等表面机械加工过程中发生，而且可以在晶体生长、热处理和相变等过程中发生于晶体内部。这时，晶内存在的包

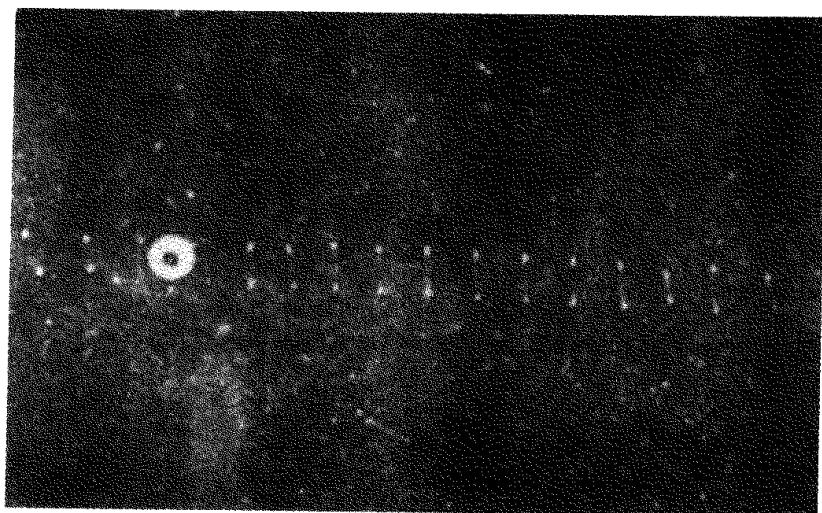


图 5-16 AgCl 晶体中小玻璃球挤压所产生的棱柱位错环组态
(引自 Jones D A, Mitchell J W. Phil. Mag. 1958, 3: 1)

裹体或第二相颗粒由于其热膨胀系数与基质晶体有差异，原本相互协调的情况为冷却引起的比容差异所破坏，在晶内造成挤压而不断地产生棱柱位错环。事实上，晶内挤压产生的棱柱位错群组态可以具有多种形态，AgCl晶体中是同轴成串且大小相同的棱柱位错环

(图 5-16)，钇镓石榴石和钇铝石榴石中的棱柱位错环则是分别躺在以包裹体为中心的几个攀移面内 (图 5-17)。棱柱位错群形态上的差异来源于晶体中位错运动阻力的差异，AgCl 中滑移阻力甚小，挤压形成的位错环通过棱柱滑移而排成串；高温生长的石榴石中大量过饱和点缺陷使攀移阻力变小，挤压形成的位错通过攀移而扩张出去。

完整晶体经受弯曲形变后，将会通过在晶内产生一定数量的刃型位错，以补偿由于弯曲而造成的出现于矢径不同的表面上的尺度差异，而如图 5-18 所示。如果弯曲晶体的

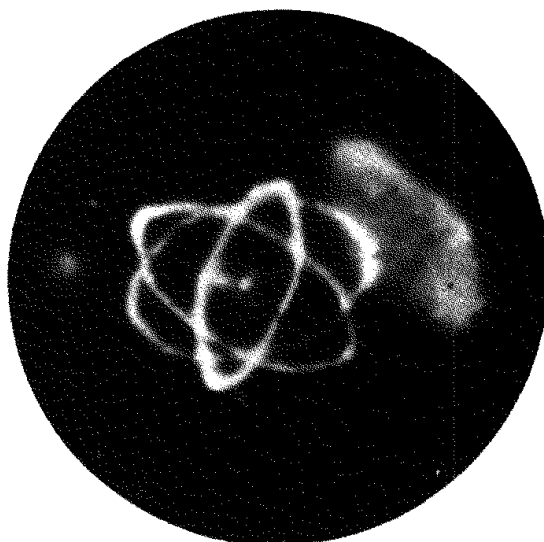


图 5-17 钇铝石榴石晶体中包裹体挤压产生的棱柱位错环组态

(引自邓佩珍, 张守都, 王炳浩等.
硅酸盐学报, 1979, 7: 183)

曲率半径为 r ，且滑移面平行于晶片底面，那么协调位错的密度为：

$$\rho_s = \frac{\theta d/b}{r \theta d} = \frac{1}{rb} \quad (5-27)$$

若滑移面与底面夹角为 α ，则上式修正为：

$$\rho_s = \frac{1}{rb \cos \alpha} \quad (5-28)$$

晶体生长过程中，由于不均匀温度梯度造成的热应变，和由于不均匀溶质浓度梯度造成的组分应变，都是晶体中局部区域塑性弯曲的物理根源，因而将产生一定密度和分布状态的协调位错。简单的计算表明，协调位错密度正比于温度梯度，即

$$\rho_s = \frac{\alpha}{a} \frac{dT_K}{dr} \quad (5-29)$$

式中， α 为热膨胀系数， a 为点阵参数。但却是与浓度梯度的变化率相关，而有

$$\rho_c = \frac{1}{a^2} \frac{da}{dc} \delta \left(\frac{dc}{dx} \right) \quad (5-30)$$

式中， $\frac{1}{a} \frac{da}{dc}$ 为单位溶质浓度引起的点阵参数相对变化。只当浓度梯度有强

烈的变化时才能引起可以观测到的位错密度，这样大变化的浓度梯度在大量溶质原子由自由表面或晶界扩散到晶体中时，可以在晶体表面或晶界附近得到。例如采用硼或磷扩散制备硅的 p-n 结时，由于掺杂原子半径比硅大得多，常可在样品中观测到由此产生的位错网络，有时其密度甚至超过原生晶体中的位错密度，对器件影响不能忽视。后来用所谓“完整晶体工艺”，将磷与砷按一定比例掺杂，抵消尺寸因子的差别，从而获得低位错密度，制成具有低噪声、高击穿电压的器件。

5.5.2 位错的增殖

原生晶体中的位错数量通常较少，即使通过滑移或攀移全部跑出晶体，也只能引起微量的塑性形变。晶

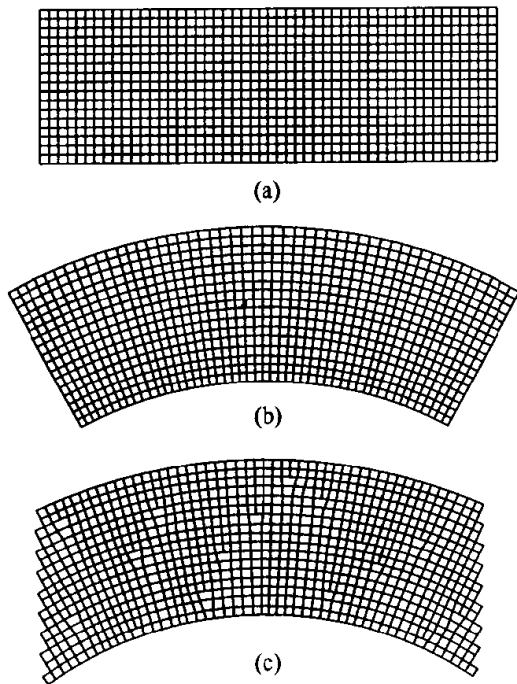


图 5-18 弯曲晶片中的协调位错

(a) 未弯曲晶体；(b) 弯曲而未弛豫；

(c) 弛豫产生协调位错

体的宏观塑性形变必须有大量位错扫过晶体，而且观测表明，塑性形变发生后的晶内位错不是少了，而是有几个数量级的增加，这些都表明位错不仅能运动，而且能大量增殖。

在塑性较好的晶体中，增殖常通过滑移方式进行。最常见的滑移增殖机制有弗兰克-瑞德 (F.C.Frank-W.T.Read) 源机制和双交叉滑移机制。设想一段两端点被林位错、微小的第二相粒子等钉扎住的可滑移位错线 (如图 5-19

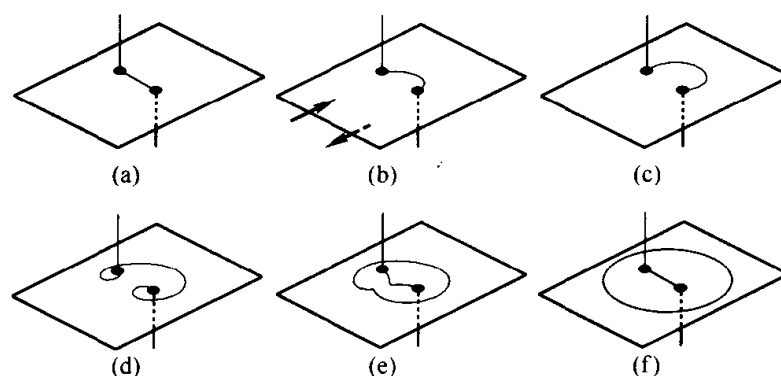


图 5-19 位错的弗兰克-瑞德滑移增殖源增殖过程示意图

所示), 当在其滑移面上的分切应力足够大时, 位错段将发生弯曲。随着切应力增大, 位错曲率增大, 因而其上的线张力也愈大。后者的作用是使弯曲位错有恢复原直线状的倾向, 一旦应力减小或消除, 位错将复原而无增殖。但当切应力大到足以使位错段弯曲成半圆形时, 其曲率达到最大, 容易看出, 如果位错半环继续扩大时, 其曲率反而变小, 因而可在不增加切应力的条件下扩张。如图所示, 位错将在两端点附近绕向后方并相遇。相遇的两部分位错线段具有相同的柏格斯矢量, 但其走向相反, 反号位错通过合并而相消。于是位错的一部分构成了闭合的位错环而继续扩张, 另一部分则回到原始位置而完成一次增殖过程。只要切应力保持足够大, 这一增殖过程就能不断重复进行。这就是位错增殖的弗兰克-瑞德源机制。易知, 该源操作的临界切应力应为:

$$\tau_c = \frac{Gb}{L} \quad (5-31)$$

式中, L 为位错线段长度。图 5-20 展示了硅单晶中观测到的弗兰克-瑞德源图像, 钉扎位错段的林位错清晰可见。单点弗兰克-瑞德源仅有一固定点, 增殖在滑移面内形成绕固定点的平面蜷线。

双交叉滑移源中的滑移位错具有螺型特征, 它没有固定的滑移面, 在滑移过程中可因局域的切应力变化而改变滑移面, 又因局域切应力减弱而回到原滑移面内而发生双交叉滑移。但这种局域应力的作用仅使一段位错发生双

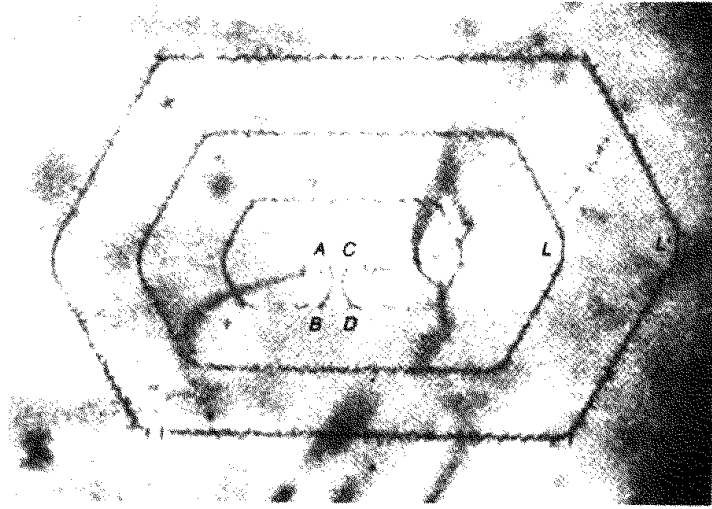


图 5-20 Si 单晶体中的弗兰克-瑞德源，位错线以 Cu 沉淀
缀饰后，以红外显微镜观测所得

(引自 Dash W C. Dislocations and Mechanical Properties of
Crystals. New York: John Wiley & Sons, Inc., 1957, 57)

交叉滑移，因而在双交叉滑移发生由次滑移面至主滑移面转化时，出现了相对固定的两点，于是像弗兰克-瑞德源那样的开始增殖。

在许多复杂结构的化合物晶体中，位错的滑移阻力较大，因而通常可能通过攀移过程实现位错的增殖。一种典型的攀移增殖源称作巴丁-赫润 (J. Bardeen-C. Herring) 源，其操作过程与弗兰克-瑞德源相似，但发生于攀移面内。另一种常见的攀移增殖方式是蜷线位错的形成和发展。在高浓度过饱和和点缺陷渗透力的作用下，接近螺型取向而两端点固定的位错段，将不能保持直线形，而是如图 5-21 那样首先发生弯曲。这种攀移在两端以相反的方向进行，因而形成以原位错段为轴线的蜷线状位错，并随着过饱和度的增加而增加蜷线的大小和圈数，甚至在端部不断发出棱柱位错环。这种增殖方式在

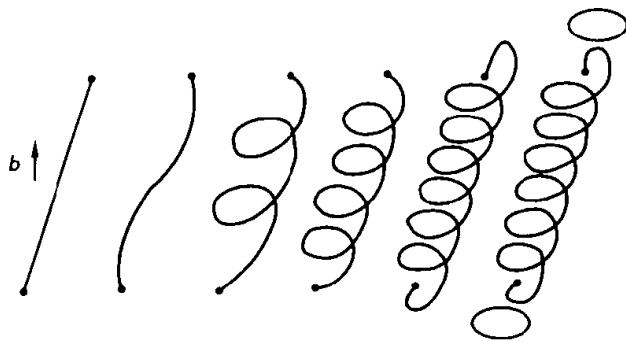


图 5-21 蜷线位错增殖源的增殖过程示意图

许多化合物晶体(见图 5-22), 特别是在完整性高、位错密度低的高熔点氧化物晶体中大量发生, 其结果是破坏了晶体完整性和均匀性而可能恶化晶体材料的使用性能, 需努力避免。

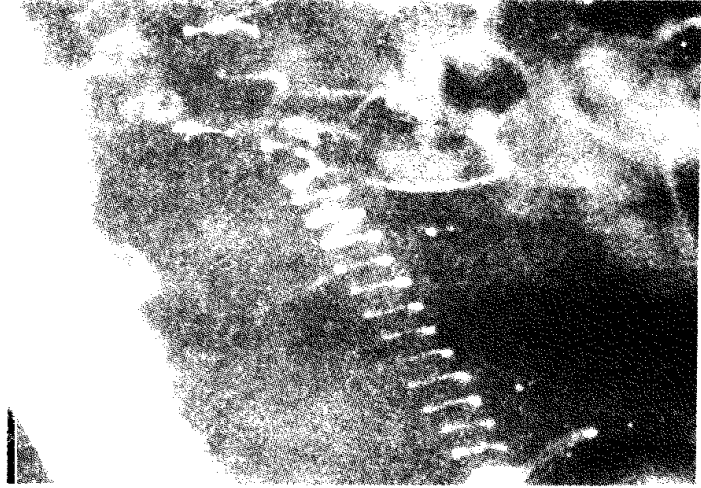


图 5-22 CaF_2 晶体中的蜷线位错和位错环

(引自 Bontinck W, Amelinckx S. Phil. Mag. 1957, 2: 94, 561)

5.6 向错^[4,9,10]

在本章的简短引言中, 式(5-1) 阐明了介质中引入线缺陷的条件, 其中剖面两岸作一刚体式的旋转时, 将在剖面的周界处形成向错线。同时在图 5-1(d)~(f)中也演示了产生楔型向错和扭型向错的操作过程。

对于连续弹性介质而言, 向错的引入(挖去中心管道的情况下)引起的应力场(内外壁半径分别为 r_0 与 r_1)为:

$$\begin{cases} \sigma_{rr} = A \left(\ln \frac{r}{r_1} - \frac{r_0^2}{r^2} \ln \frac{r_0}{r_1} \right) \\ \sigma_{\theta\theta} = A \left(\ln \frac{r}{r_1} - \frac{r_0^2}{r^2} \ln \frac{r_0}{r_1} + 1 \right) \\ \sigma_{zz} = \nu(\sigma_{rr} + \sigma_{\theta\theta}) \\ A = G\Omega/2\pi(1-\nu) \end{cases} \quad (5-32)$$

可以看出应力是随 r 增大的增加的。单位长度向错线的弹性能:

$$W = \frac{G\Omega^2}{16\pi(1-\nu)}(r_1^2 - r_0^2) \left[1 - \frac{4r_0^2 r_1^2}{(r_1^2 - r_0^2)} \left(\ln \frac{r_1}{r_0} \right)^2 \right] \quad (5-33)$$

其量级相当于柏格斯矢量为 $\frac{1}{2} r_1 \Omega$ 的刃型位错, 能量之大使之以简单的形式很难出现。

没有结构的连续介质事实上是不存在的，所有物质均由原子构成，特别是晶体材料的原子排列还具有周期性和对称性。这就决定了晶体中的向错必须为晶体中的旋转对称性所约束。众所周知，晶体中的旋转对称元素限于旋转等于 $\Omega = \pm 2\pi/n$ ，而 $n = 1, 2, 3, 4, 6$ 。因此要求剖面不留痕迹的全向错的最小向错角也有 60° ，其应力场和畸变能都非常大，以致难于存在。图 5-23 中我们给出了晶体中 $\Omega = 90^\circ$ 的几种假想向错的示意图，可见其畸变之

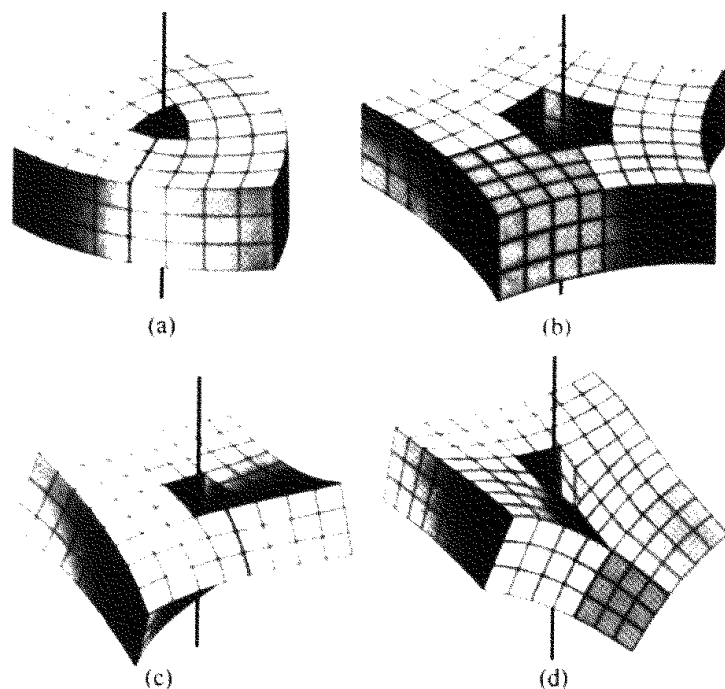


图 5-23 立方晶体中 $\Omega = 90^\circ$ 的几种假想向错的示意图

(a) $+90^\circ$ 楔型向错；(b) -90° 楔型向错；(c)、(d) 90° 扭型向错

大定是实际晶体所不能容忍的。一种可能的特殊情况是非常靠近的一对平行的正反向错，其长程畸变被相互抵消，具有较低的能量（见图 5-24）它相当于柏格斯矢量为 $b = 2d \tan\left(\frac{1}{2}\Omega\right)$ 的刃型位错。中断在晶体中的对称倾侧型亚晶

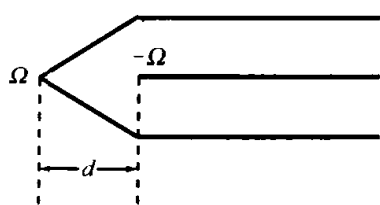


图 5-24 一对平行的正反向错靠得很近时，相当于一根刃型位错线

界及面心立方晶体中的伪五角形复合孪晶也可以看成是晶体中向错的例子，其中前者 $\Omega = b/d$ ，后者 $\Omega = \pi/24$ 。最近，在具有较复杂结构、纳米级大小的高压相 TiO_2 球形颗粒中也发现了伪五角形孪晶态的向错（见图 5-25）。由前知，向错线的应力随 r 增大而增加，因而在纳米级，有时在微米级大小的晶态颗粒中出现较小向错角的五重孪晶态向错是容易实现的。

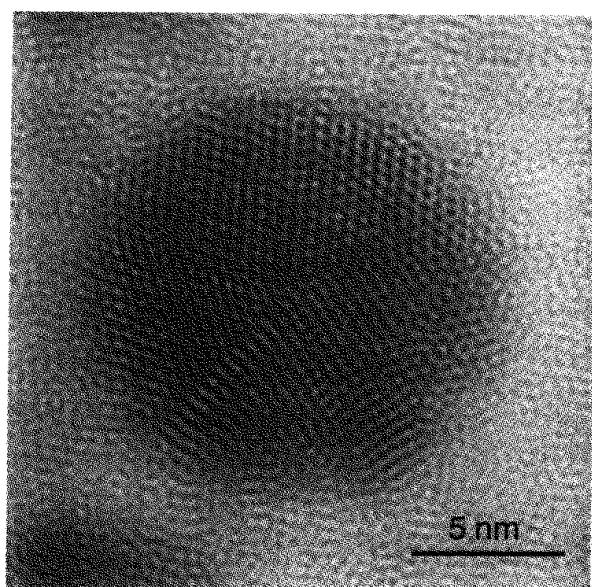


图 5-25 高压相 TiO_2 纳米颗粒中五重孪晶的高分辨电子显微像
(引自马国斌, 南京大学博士后出站报告, 2001)

向错在普通晶体中很难于存在, 然而, 在大量具有薄层周期性结构的病毒外壳中很早就被发现, 20 世纪 70 年代时, R·O·艾里克森 (R.O.Erickson) 就制作了如图 5-26 所示的病毒蛋白质外壳的模型。更为有趣的是, 近年来人工制备的纳米碳管中, 两端部封闭的碳管结构竟完全与之一模一样, 而球状 C_{60} 的足球状结构也与此异曲同工。可以证明, 任何能够形变成球面 (拓扑学上等效于球面) 的表面必须具有总旋转角为 720° 的向错。在这种表面中, 其基本对称元素是六角形的, 为使其闭合需要插入 12 个五角形, 每一个五角形均代表一个 60° 向错。

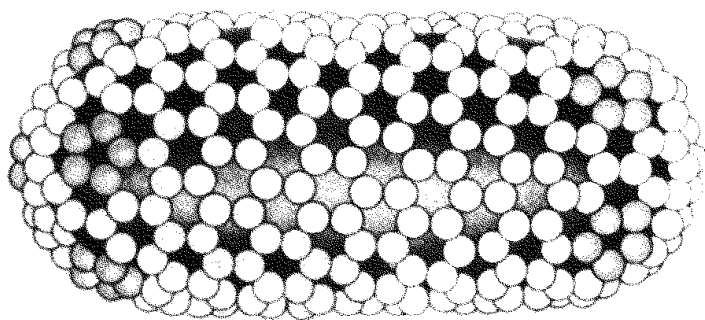


图 5-26 病毒蛋白质外壳的结构模型^[9]

事实上, 在液晶和一些有序介质中向错是极为常见的一类缺陷组态。因而已从理论和实验上进行了大量而卓有成效的研究工作。从一个点阵出发,

如果每一阵点上都附加一个取向因子, 则在各阵点取向不相关联时, 点阵的对称性并不改变。但是如果阵点取向相互有关联, 则对称性显然会发生变化, 例如铁磁性或螺旋磁性介质中, 阵点上的自旋取向就是相互关联的。这时如果阵点的位置是不相关联的, 那么就将引入长程的自旋角相关性表示介质的对称性, 这就给向错的出现提供了物理基础。液晶的情况也与此类似, 其有序性在于杆状分子的取向相关性, 所存差异在于杆状两端无差异, 无正负取向之分, 而自旋有其本身的取向。液晶之所以能大量出现向错, 主要因为向错产生的畸变可以通过粘滞性的流动得到部分消弛。

向错这种线缺陷的产生与位错有一定的类似性, 都是剖面的周界, 因而与位错类似, 向错线也不能终止在晶体内部, 它可以出露在晶体表面、终止于晶体内部的其他向错或某些缺陷上, 或者可以在晶体内部形成向错环。与用柏格斯回路来确定位错的相关平移柏格斯矢量 b 类似, 用类似的弗兰克-纳巴罗回路法可以确定向错的特征量——强度 S , 它代表旋转量和符号。此外, 理论研究表明, 向错与位错之间存在着相互作用, 它既可以产生位错, 也可以吸纳位错, 同时还可以借助于位错而使自身得以运动。

参 考 文 献

- 1 Friedel J. Dislocations. Oxford: Pergamon Press, 1964
- 2 Nabarro F.R.N. Theory of Crystal Dislocations. Oxford: Oxford University Press, 1967
- 3 Hirth J. P. Lothe J. Theory of Dislocations. 2nd ed. New York: McGraw-Hill, 1982
- 4 Nabarro F.R.N. Dislocations in Solids, Vol.1~6. Amsterdam: North-Holland, 1979~1984
- 5 冯端等. 金属物理学. 第一卷. 结构与缺陷. 北京: 科学出版社, 1987.266
- 6 杨顺华. 晶体位错理论基础. 第一卷. 北京: 科学出版社, 1988; 杨顺华, 丁棣华. 晶体位错理论基础. 第二卷. 北京: 科学出版社, 1998
- 7 Hull D, Bacon D J. Introduction to Dislocations. 3rd ed. New York: Pergamon Press, 1984. D. 赫尔, D J 培根著. 位错导论. 丁树深, 李齐译. 北京: 科学出版社, 1990
- 8 Verma A. R. Crystal Growth and Dislocations. London: Butterworth. 1953
- 9 Harris W.F. Scientific American. 1977, 237: 66
- 10 Kléman M. Points, Lines and Walls. New York: John Wiley & Sons, Inc., 1983

第 6 章 表面与界面

撰稿人 李 齐

理想完整晶体在于其结构完全的周期性和对称性。零维的点缺陷和一维的位错以外，最重要的缺陷就是通常所说的面缺陷了。有限大小的晶体，其表面是平移对称性终止处，是一类面缺陷。由于晶体材料与外界相互作用是通过表面来实现的，因而表面的结构特征无论从基础理论或技术应用的角度看，都是至关重要的。另一类面缺陷出现于有限晶体的内部，称作界面。它们种类繁多，有的是和理想点阵结构发生偏离，有的牵涉到化学组分甚或相的差异，有的则是对应于电磁结构的有序跃变，因而它们对于晶体材料的各种性质可以产生更为广泛和重要的影响。

近年来，表面与界面起突出作用的新型材料，如薄膜与多层膜、超晶格、超细微粒与纳米材料等的发展正如日中天，既发现了一系列新的物理现象和效应，又展示了应用上的巨大潜力。

6.1 晶体的表面

6.1.1 晶体外形与表面微观形貌

处于晶体自由表面上的原子键合状态与晶内不同，由于键合能具有负值，近邻数的减少，将使系统能量增高，这就是表面能的物理来源。通常定义增加单位面积的表面所引起的自由能增量 γ 为表面能系数。由于降低自由能的要求，系统将倾向于缩小其表面积，这相当于表面上任一面元的周界上都会受到一个力的作用，每单位长度上所受的力 f 被称作表面张力。在各向同性的情况下， γ 和 f 在量纲和数值上都是相同的。

在平衡状态下生长晶体时，在体自由能保持恒定的情况下，自由能极小条件可以归结为表面能极小。若 γ 是各向同性的，自由能极小可以约化为表面积的极小，其平衡形态便是球形。这就是液滴在忽略其他外力，如重力的情况下所取的外形。近年来对纳米颗粒所作的研究表明，一些处于结构不稳定状态下的纳米颗粒也是球形或近球形形态(见图 6-1)。

然而，由于晶体结构的各向异性，不同晶面上原子的密度、配位数及键合角等的不同，将使其作为晶体表面时的表面能系数也将有所不同。这表明，表面能系数随表面取向 n 而有差异，其数值等于 γ_n ，即该晶面表面能

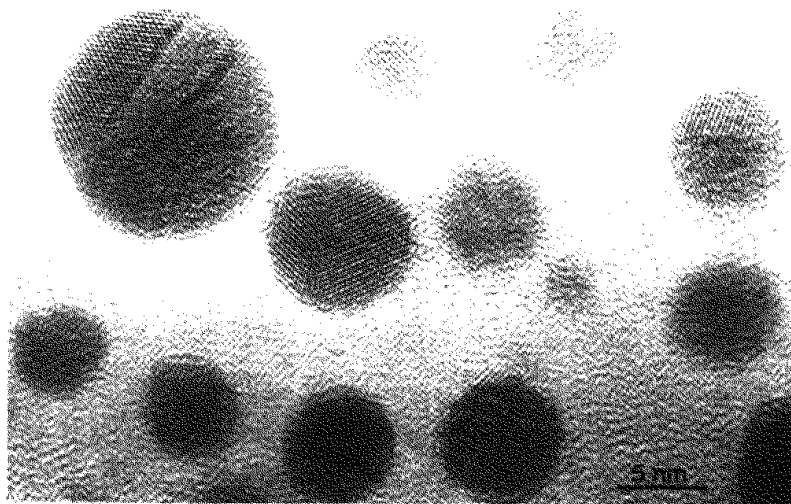


图 6-1 α - PbO_2 型高压相 TiO_2 球形纳米颗粒的高分辨电子显微像
(引自马国斌, 南京大学博士后出站报告, 2001)

系数在其法线 n 方向上的投影。据此可以做出表面能系数极图^[1], 即由一

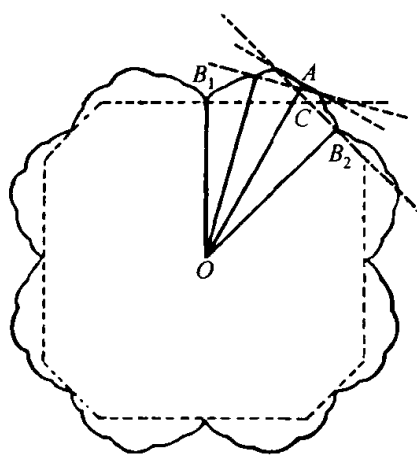


图 6-2 表面能系数极图与晶体
平衡外形的二维剖面图
点线为垂直于矢径的晶面,
实线为表面能系数极图,
虚线为晶体的平衡形状

原点出发引出矢径, 令其长度正比于 γ_n , 所有矢径终点的轨迹将为一曲面 (如图 6-2 中的实曲线)。而由此极图容易导出晶体的平衡外形: 在极图上的每一点作垂直于矢径的平面, 去掉这些平面相重叠的区域, 剩下的体积最小的多面体即为晶体的平衡外形, 这实际上就是该极图中体积最小的内接多面体 (如图 6-2 中的虚线多边形)。

在不受约束条件下生长出的许多晶体的确呈现出多面体的外形, 显露的晶面一般也是表面能较低的密排低指数面。然而, 实际的晶体生长过程并非在平衡态下进行, 如杂质的参与改变了表面能系数, 位错等缺陷生长机制的操作增大了某些晶面的生长速率, 重力作用造成的对流给晶体生长附加了不等

同的环境, 如此等等, 都将使实际晶体的外形与理论结果发生偏差, 甚至改变晶体的惯态面。

如前所述, 只有在严格控制生长条件使其在近乎平衡态下生长时, 晶体表面才能出现完好的密排低指数面, 在低温下这类表面在原子尺度上是光滑的。一般非平衡态下生长的晶体, 其表面稍许偏离低指数面的部分称邻位

面，远离低指数面者则称粗糙面。通常晶体总是存在为数不少的低或较低指数的晶面，因而只要生长条件合适，邻位面总是占表面的大部分，格外受人们关注。

邻位面可以用平台-台阶-扭折模型^[2]来描述。如图 6-3 所示，其主要部分的平台是密排低指数晶面，表面对低指数面的偏离则可以用平行于密排方向的台阶，以及沿台阶出现扭折来实现。偏离的大小和偏离的方向可以通过

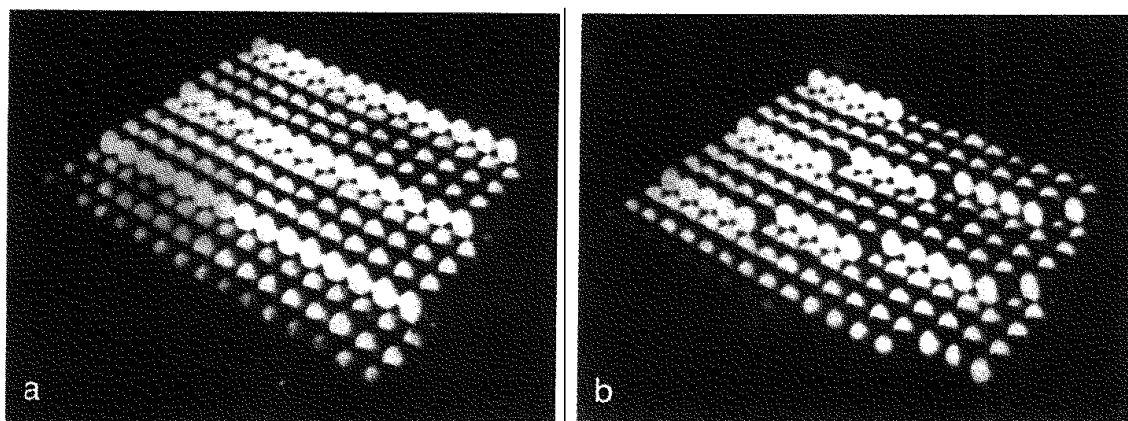


图 6-3 面心立方晶体接近 (100) 面的邻位面的平台-台阶-扭折模型图
(a) (100) 面平台上的台阶；(b) 台阶上的扭折

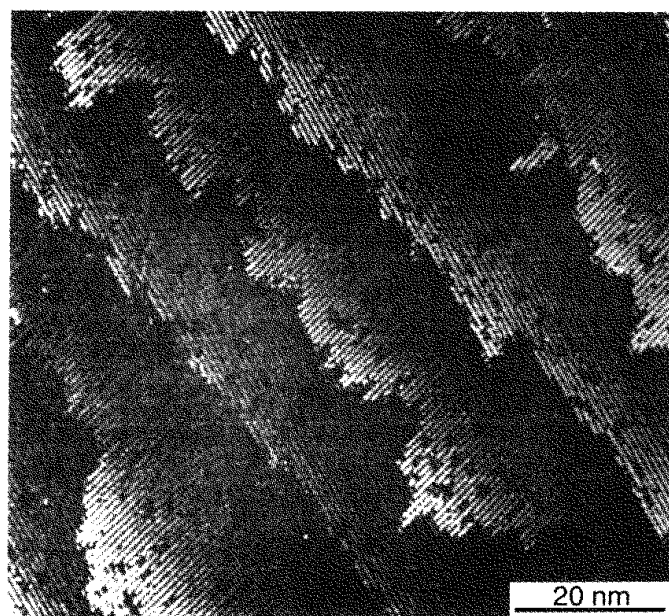


图 6-4 Si 单晶(100)面邻位面的 STM 像,显示了其平台、台阶及扭折的典型形貌。注意其表面的(2×1)重构组态,白线为配对原子列
[引自 Hamer R J, Kohler U K, Demuth J E. J. Vac. Sci. Tech. 1990, A(8):195]

改变台阶或扭折的密度来调节。20 世纪 50 年代末,场离子显微镜的直接观测已证实了平台和台阶的存在,而近年来的扫描隧道显微镜及原子力显微镜的观测则更全面地证实了这一模型的正确性(见图 6-4)。

邻位面的台阶和扭折是晶体生长时原子优先沉积的位置,原子沉积导致扭折沿台阶的运动,从而使台阶向前推进,相继的台阶不断地扫过晶体表面,最终引致了晶体表面的法向生长。必须指出的是,台阶和扭折也是表面反应优先发生的场所。萨摩杰(G.A.Somorjai)等人用分子束技术研究 Pt 表面催化效应,就发现位于平台上和台阶处的 Pt 原子催化效应迥然不同,后者的平均催化活性要大得多。此外在化学侵蚀和解理断裂等过程中它们也起着重要的作用,诸如表面吸附和偏析、电介质材料的表面电导和介电击穿等物化性质也与晶体表面的微观形貌有密切关系,还有大量工作有待深入。

以上讨论表面形貌时,我们忽略了表面缺陷的存在。事实上,由于热力学平衡点缺陷的存在,奇异面和邻位面的平台部分可能存在附加原子和表面空位,台阶上也可出现单原子大小的附加正、负扭折;晶体中的各种广延缺陷,如位错和各种界面都可能露头于晶体表面,形成特殊的台阶。这种台阶在晶体生长或解体的过程中永不消逝,从而在这两种物理化学过程中发挥重要作用。

6.1.2 表面弛豫与重构

在以上的讨论中,还隐含着—个简化了的基本假定,即晶体内部的结构无改变地延续到表面层,直至为表面截断为止。这一假定的缺憾之处直至 20 世纪 70 年代以后低能电子衍射术、高分辨电子显微术等表面晶体学研究手段完善之后,才开始得到修正。观测结果表明,实际的表面结构与理想的体内结构的差异主要表现为两种形式,一是表面弛豫,一为表面重构。

表面弛豫现象表现在晶体结构基本相同,但点阵参数略有差异,特别是表现在表面与其下少数几个原子层间距的变化上,即法向弛豫。如果考虑到表面原子层受力情况的明显不对称性,产生法向弛豫,尤其发生压缩弛豫是不难预料的。金属晶体结构简单,弛豫现象相对较为明显,因而已进行了较多的研究。近年来的理论模型认为,金属表面电子改变空间分布以降低能量,结果造成表面电偶极子层,并通过它与表面原子层的互作用而产生压缩弛豫。实测也表明,这种倾向在密排的表面上虽然并不明显,但随着非密排程度的提高,法向收缩效应的确更为突出。这样的模型也解释了常见的下述振荡式弛豫现象:表面第一、二层间原子间距受压缩,第二、三原子层间膨胀,第三、四层间又受压缩[如金的(110)面表面递次为 -1.25% 、 0.7% 和 -0.25%]。个别金属表现出表面的法向膨胀,其原因尚不清楚。

晶体表面结构变化的另一种方法是表面重构,通常表现为表面超结构的

出现。由于表面弛豫仅是表面层晶格不大的畸变（通常为 1% 左右，少数如 Mo、W、Al、Ag 等能到 3%），而表面重构能使表面结构发生质的变化，因而在许多情况下，表面重构在降低表面能方面比表面弛豫要有效得多。所谓超结构是指晶体平移对称周期，即晶胞基矢成倍扩大的结构状态，合金的无序-有序转变和复合氧化物固溶体的失稳分解都是造成超结构的物理过程。作为表面超结构，则是表面层二维晶胞基矢的整数倍扩大。最常见的表面重构有两种类型，一种是缺列型重构，一种是重组型重构。图 6-5 表示出了简单立方点阵中这两类重构的示意图。

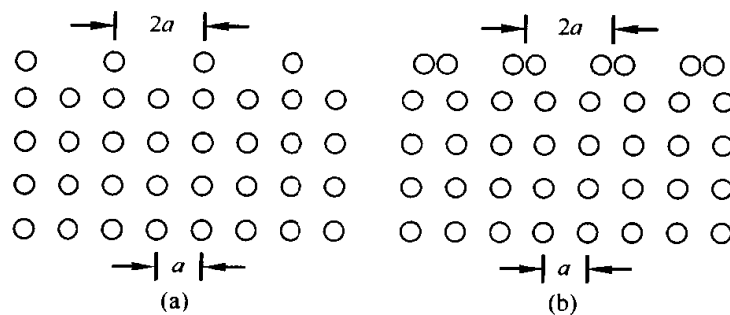


图 6-5 简单立方晶体中两种表面重构引起的原子分布示意图

(a) 缺列型重构；(b) 重组型重构

缺列型重构是表面周期性地缺失原子列造成的超结构。洁净的面心立方金属铌、铂、金、钯等 $\{110\}$ 表面上的 (1×2) 型超结构是最典型的缺列型重构的例子，这时晶体 $\{110\}$ 表面上的原子列每间隔一列即缺失一列。图 6-6 是金 (110) 面 (1×2) 重构沿 $[1\bar{1}0]$ 带轴摄取的高分辨电子显微像^[3]。显然，重构后， (110) 表面由平面转变为由很狭窄的 $(1\bar{1}1)$ 、 $(\bar{1}11)$ 小面相间排列而成的折面。理论计算表明，如果只考虑最近邻的交互作用，则两种表面形态的表面能量无明显差别，但若考虑较长程的互作用，后者则具有较低的能量。

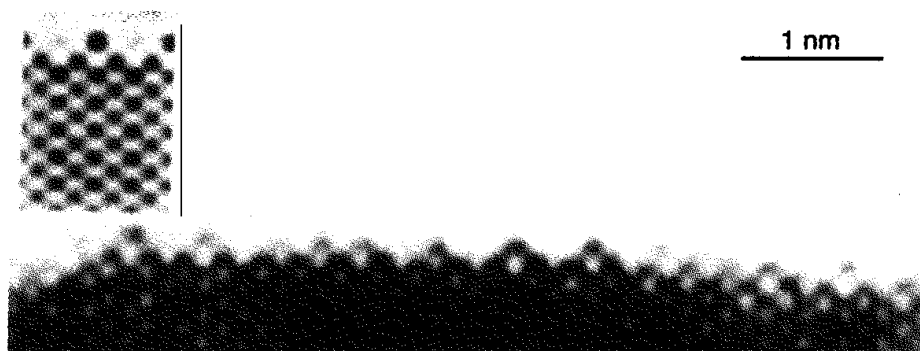


图 6-6 金 (110) 表面 (1×2) 重构的高分辨电子显微像，沿 $[1\bar{1}0]$ 带轴拍摄

而这一因素的作用对于元素周期表越下方的金属而言更为突出，表面能有更大的减小。这就是周期表底部的铀、金、铂等不多的金属晶体得以观测到这类重构的原因所在。

重组型重构并不减少表面的原子数，但却显著地改变表面的原子排列方式。通常，重组型重构发生在共价键晶体，或有较强共价成分的混合键晶体中。共价键具有强的方向性，表面原子断开的键，即悬键处于非常不稳定的状态，因而将造成表面晶格的强烈畸变，最终重排成具有较少悬键的新表面结构，显然这种新结构是具有较大周期的超结构。以硅的 $\{100\}$ 表面为例，其表面原子有两个悬键，未重排时，每个悬键均有一个不成对电子而极不稳定；重排后，表面硅原子一对对相互靠近配对，配对原子间基本转变为电子成对而无悬键的组态，从而大大降低了表面能^[4]。这类配对现象在其他低指数表面，如 $\{111\}$ 和 $\{110\}$ 也会发生，同时产生重构。观测发现，在硅晶体中， $\{100\}$ 面上 (1×2) 重构是最常见的重组型重构，但配对方式的不同也可造成其他类型的重构，如 (2×2) 重构(见图 6-7)；而在其他取向的表面上，由于初始原子排列方式的不同，其重构方式也有不同，如常见于 $\{111\}$ 表面上的是 (7×7) 型重构；此外，表面上台阶处额外的悬键也会造成其他形态的重构结构。

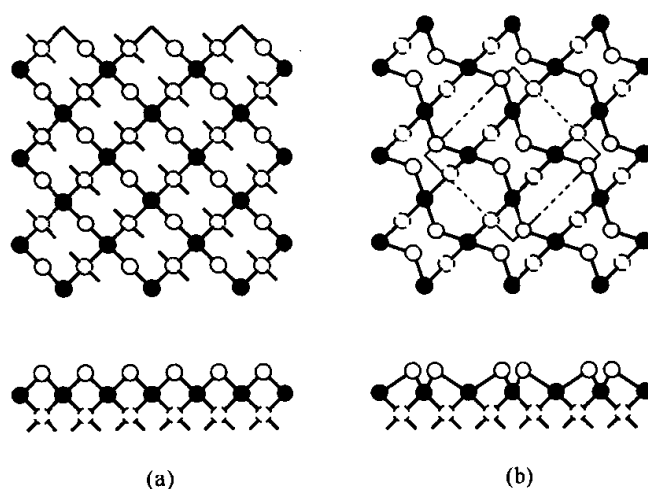


图 6-7 硅(100)表面 (2×2) 重构前(a)及重构后(b)表面原子分布及键合状态示意图
上为顶视图，下为侧视图

必须指出的是，重组型重构常会同时伴有表面弛豫而进一步降低能量，只是就对表面结构变化的影响程度而言，表面弛豫比重新配对要小得多。综合以上两种重构方式的系统研究结果，可以认为，当原子键不具有明显方向性时，表面重构较为少见，即使重构也以缺列型重构为主，当原子键具有明

显的方向性，如共价键时，则洁净的低指数表面上的重组型重构是极为常见的。近年来发现多种具有较强共价成分的晶体中也存在重组型重构，最典型的例子就是 SrTiO_3 晶体。

6.1.3 技术晶体的表面

前面讨论的表面均属洁净表面，实际应用的技术晶体的表面并非如此。差别主要来源于以下几种情况。

① 表面吸附。这是固体表面以外气相中的原子或分子沾集于其表面上的现象。气相中的原子或分子可以被吸附于表面，而被吸附的原子或分子也可以重新蒸发而脱离表面。当吸附和脱附间达到动态平衡时，表面吸附坐位被吸附分子占据的分数，即吸附覆盖率 θ 满足：

$$\frac{\theta}{1-\theta} = p\kappa \quad (6-1)$$

式中， p 为气相压强， $\kappa = \frac{\alpha \exp(U_0/k_B T)}{\gamma(2\pi m k_B T)^{1/2}}$ 是与吸附分子质量、动态过程振动频率及温度等相关的物理量。可以明确看出的是，当压强甚小时， θ 与 p 成正比；当压强大时， $\theta \rightarrow 1$ ，吸附趋于饱和。这是郎缪尔 (I. Langmuir) 提出的吸附模型导出的结果，其前提是晶体表面上只有一种吸附坐位，而被吸附的分子间没有相互作用。近年来的低能电子衍射结果发现，不同气体在同一种金属表面上的吸附层往往具有相同的二维周期性结构，而最近对镍 (110) 表面氢吸附结构的 STM 三维形貌观测表明，吸附周期会随着 θ 值的变化而改变。这种表面吸附层的结构组态在许多复杂氧化物表面也同样存在，高分辨电子显微像表明，LTL 型沸石表面上溅射 Pt 时，Pt 原子团以沸石晶体周期大致相同的距离规则地排列。这些观测结果说明吸附层内存在着明显的横向交互作用。

吸附原子与晶体表面的交互作用，吸附原子的横向交互作用，表面重构现象以及许多热力学量随温度的变化等观测结果，还预示着有可能在晶体表面发生丰富多彩的相变。

② 表面偏析。这是固溶体中溶质原子在表面层富集的现象。由于表面偏析与表面吸附两者均是异类原子与晶体表面间相互作用的结果，异类原子前者来源于体内，后者来源于体外，它们遵从相似的热力学规律，因而表面偏析现象中表面溶质原子的占有率与吸附有类似的表式：

$$\frac{\theta}{1-\theta} = x \exp(-\Delta G/k_B T) \quad (6-2)$$

式中， x 为溶质原子的摩尔浓度， ΔG 为溶质原子在表面和体内的吉布斯自由能之差。显然，表面层溶质原子浓度最高，朝向晶内浓度迅速减小直至与体内相同。

③ 表面化学反应。当晶体表面含有化学活性甚强的组分时，可以产生化学吸附（区别于前述以范德瓦尔斯力吸附时的物理吸附，这时吸附能一般低于 0.2eV ），以致发生化学反应。这时晶体表面原子与被吸附原子构成了统一的电子体系，这种体系的电子结构将有助于理解化学催化作用的本质。其实，表面化学反应的现象是极为普遍的。许多金属材料和非金属材料表面形成氧化物或氢氧化物，当它们形成均匀而致密的结构时可以保护材料表面，而形成不均匀的、疏松的结构时，则材料表面极易被腐蚀。

前面的讨论尽管论及表面弛豫与表面重构，但前者仅是弹性畸变，后者表面结构保持了自身相对的完整性，并没有将晶内广延缺陷在表面的露头以及其他原因引入表面的晶体缺陷考虑进去，而这些因素对表面结构的基本特征及表面结构的完整性都有着重要的影响，并进而影响材料的使用性质。

固体材料在其制作成可用的元器件时通常是需要经过切割、表面磨制和抛光的。硬质磨料在表面的刮削和压印，可以在非晶态材料表面造成划痕和微裂纹，在塑性较好的晶态材料表面形成足够厚（ $>0.1\mu\text{m}$ ）的含有高密度位错和强烈晶格畸变的表面损伤层，在脆性晶体材料表面引起解理性的或非解理性的微裂纹。而在许多铁电、铁弹晶体表面还能引入大量深入晶内的微畴。这些表面问题，对技术材料的使用性质都有或多或少的，有时是很严重的影响。例如电介质材料表面电导和电击穿现象中，表面结构缺陷的影响是不容忽视的；激光晶体的表面缺陷不仅可能降低其使用效率和寿命，甚至可能造成器件的炸裂；对许多存在相互接触而又承受机械力作用的表面（如金属机械）之间，由于表面存在峰与谷的不平整状态，实际接触面积较表观的要小得多，因而在负载的作用下将容易磨损或损坏；而对于近年来迅速发展起来的薄膜与多层膜等新型材料，在其制作过程中，对衬底材料表面的处理则要求极严，器件质量问题很多由此而来，成为整个制备工艺中很重要但却易被忽视的一环。

6.1.4 纳米粒子的表面

随着纳米科学技术从 20 世纪 80 年代开始崛起，纳米粒子的结构、表面结构以及纳米粒子的特殊性质引起了科学界的极大关注。特别是当粒子直径为 10nm 左右时，其表面原子数与总原子数之比已达 50% ，因而随着粒子尺寸的减小，表面的重要性越来越大。这里我们简单介绍与纳米粒子表面相关的一些研究结果。

具有弯曲表面的材料，其表面应力正比于其表面曲率。由于纳米粒子表面曲率甚大，因而将有甚大的表面应力作用于其上，使其处于受高压压缩（如表面应力为负值则为膨胀）状态。例如对半径为 10nm 的水滴而言，其

压力有 14MPa。对于固体纳米粒子而言，如果形状为球形，且假定表面应力各向同性，其值为 σ ，那么，粒子内部的压力应为 $\Delta p = 2\sigma/r$ 。由于该式非常类似于由边长为 l 的立方体推出的结果，而并非与曲率相关，因而该式也应适用于具有任意形状的小面化晶体颗粒。当然后者不同小面有不同的表面能，因而情况要复杂得多。如果因此而发生点阵参数的变化，那么这种变化也将是各向异性的。最近，马国斌等人对 TiO_2 纳米粒子的观测结果表明，20nm 以内的近球形纳米粒子在晶态时均为高压相，即 $\alpha\text{-PbO}_2$ 型相，正是各向异性表面应力作用的结果，而其中出现大量含五重孪晶的粒子，也正是高压相中两个 $\{111\}$ 晶面间夹角为 71.4° ，构成五重孪晶较面心立方晶体更为有利的缘故。

粒子尺寸减小的另一重要效应是晶体熔点的降低。由于表面原子有较多的断键，因而当粒子变小时，其表面单位面积的自由能将会增加，结构稳定性将会降低，使其可以在较低的温度下熔化。实验观测表明，纳米金粒子的熔点，当其尺寸小于 10nm 时甚至可以降低数百度。

此外，非常小的纳米粒子的结构不稳定性也早已被人们所关注，早期主要是在金属中，近期也在氧化物晶体中观测到。当其在高分辨电镜中观测时，诸如 Au、 TiO_2 等小纳米粒子非常快速地改变着它们的结构：从高度晶态化到近乎非晶态，从单晶到孪晶直至五重孪晶态，从高度完整到含极高密度的位错。通常结构变化极快，但相对稳定态则往往保留稍长时间。这种行为被称为准熔化态，仍是由于高的表面体积比所造成的，它大大降低了熔点，使纳米粒子在电镜中高强度电子束的激发下发生结构涨落。

6.2 平移界面

晶体内部的一部分相对于另一部分作一点阵平移矢量的相对平移，则将在其周界处形成位错，而距位错较远处将恢复完整，虽含有弹性形变，但不留下任何界面。如果相对平移只是点阵平移矢量的一部分，那么界面将不再完整，晶体能量将大幅度增高而不能存在，只有一些特定结构的晶体的某些特定非点阵平移，才能造成可以存在的低能量界面。这类界面称作平移界面，它与一特征的平移矢量 \mathbf{R} 相联系，在许多情况下，界面还必须是特定的晶面。密集结构及层状结构晶体中密排面上的堆垛层错，有序相晶体中的反相畴界以及一些化合物晶体中的晶体学切变面，属平移界面。

6.2.1 堆垛层错

从形式上看，任何一个晶体都可以看成是一层层原子按一定方式堆砌而成，密排面内原子间的键合较强，相邻密排面间原子的键合一般较弱。事实上，晶体的生长常是按密排面来堆垛，而晶体内部的相对滑移也是发生在密

排面之间。

面心立方与密集六方结构是两种最简单的密集结构，前者的密排面是 $\{111\}$ 面，后者是 (0001) 面。两种密排面具有相同的六方密排方式，只是六方密排面堆垛方式的不同造成了两种不同的密堆积结构。如图 6-8 所示，如果我们将第一层六方密排面的原子位置标为 A ，则其上可供堆垛的位置有两种，分别被标为 B 和 C 。如果只堆垛一层，则两者是等价的；如果继续向上堆垛，则将有两种不同的简单堆垛方式，分别构成不同的结构类型：按 $ABCABCA\cdots$ 方式堆垛则成面心立方结构，按 $ABABABA\cdots$ 方式堆垛则成密集六方结构。如果用弗兰克符号来表示，顺顺序堆垛，即 AB, BC, CA 时均用正三角形 \triangle 表示，逆顺序堆垛，即 BA, CB, AC 时均用倒三角形 ∇ 表示，则面心立方结构与密集六方结构可分别表示为 $\triangle\triangle\triangle\triangle\triangle\triangle\cdots$ 和 $\triangle\nabla\triangle\nabla\triangle\nabla\cdots$ ，见图 6-8(b)、(c)。堆垛方式与结构类型的对应关系就一目了然，而且不必考虑原子层堆垛时的参考原点了。

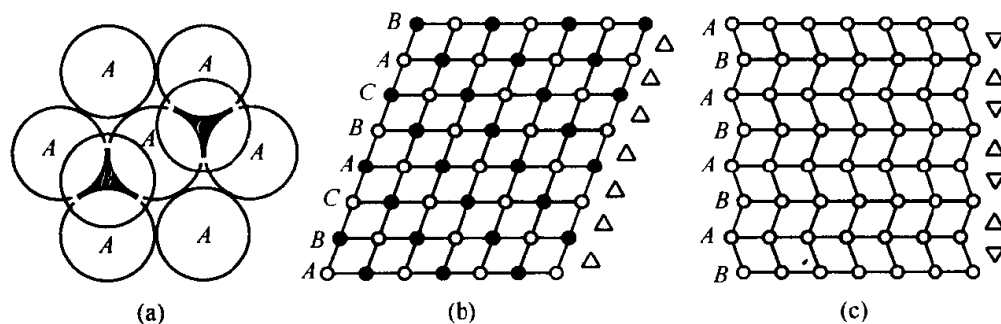


图 6-8

(a) 六方密排面 A 上的两种等价的堆垛位置 B 和 C ；(b) 面心立方结构的堆垛方式；(c) 密集六方结构的堆垛方式

堆垛层错（以下简称层错），就是指正常堆垛顺序中引入不正常顺序堆垛的原子面而产生的一类面缺陷。以面心立方结构为例，当正常层序中抽走一原子层，如图 6-9(a)所示，相应位置出现一个逆顺序堆垛层…… $ABCACABC\cdots$ ，即…… $\triangle\triangle\triangle\nabla\triangle\triangle\triangle\cdots$ ；如果正常层序中插入一原子层，如图 6-9(b)所示，相应位置出现两个逆顺序堆垛…… $ABCACBCAB\cdots$ ，即…… $\triangle\triangle\triangle\nabla\nabla\triangle\triangle\triangle\cdots$ 。前者称抽出型（或内禀）层错，后者称插入型（或外禀）层错，是层错的两种基本类型。显然，层错处的一薄层晶体由面心立方结构变为密集六方结构，同样在密集六方结构的晶体中层错处的一薄层晶体也变为面心立方结构。这种结构变化，并不改变层错处原子最近邻的关系（包括配位数、键长、键角），只改变次近邻关系，几乎不产生畸变，所引起的畸变能很小。但是，由于层错破坏了晶体中的正常周期场，使传导电子产生反常的衍射效应，这种电子能的增

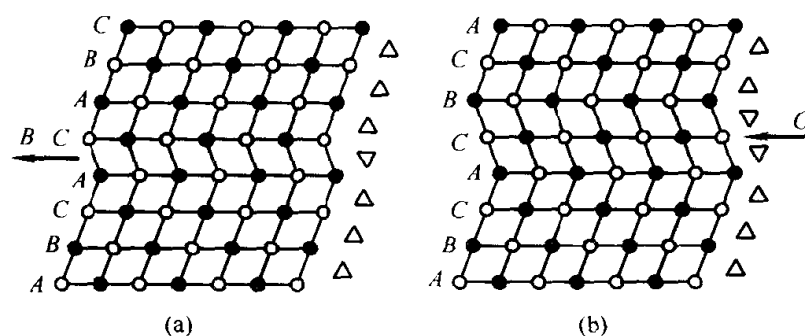


图 6-9 面心立方晶体中的抽出型层错(a)和插入型层错(b)

加构成了层错能的主要部分，总的说来，这是相当低的。因而，层错是一种低能量的界面。

层错可以通过多种物理过程形成。首先，在晶体生长中，以六方密堆积面的堆垛而生长晶体时，由于以正常和不正常顺序堆垛时的能量相差很小，偶然因素很容易造成错误堆垛而形成层错。其次，过饱和点缺陷在密排面上的聚集，再通过弛豫过程形成层错。空位聚集成盘状，通过崩塌式的弛豫形成的是抽出型层错；自填隙原子聚集成片，当然是形成插入型层错。易知层错两侧的晶体相对位移为 $1/3 \langle 111 \rangle$ ，这是一个非点阵平移，是描述这类层错的特征平移矢量。必须指出的是，形成这类层错的同时，层错的边界，即原点缺陷盘的边缘位置处将产生一根闭合的位错环，其柏格斯矢量 $1/3 \langle 111 \rangle$ 与位错环面垂直而为刃型位错，但柏格斯矢量是非点阵平移矢量，因而称不全位错。这种不全位错环不能沿其滑移柱面滑移，否则将使其扫过的面成为严重的原子错排面，被称作弗兰克不全位错环，见图 6-10(a)。

抽出型层错也可以在晶体塑性形变过程中产生，这时， $\{111\}$ 面两侧晶体相互间作一沿该面的 $1/6 \langle 112 \rangle$ 型的滑移，其相应的特征平移矢量此时为 $1/6 \langle 112 \rangle$ 。如果这种相对位移仅限于该晶面的一部分，则其周界是柏格斯矢量为 $1/6 \langle 112 \rangle$ 的不全位错环。这种不全位错的滑移发生在层错所在原子面内，仅使层错面积增减而不造成新的严重错排面，因而是可滑移位错，被称作肖克莱不全位错，见图 6-10(b)。

由前知，面心立方晶体中最常见的全位错是 $1/2 \langle 110 \rangle$ 型，如果某些晶体中的层错能较低时，以下的位错分解反应：

$$\frac{1}{2} [\bar{1}10] \rightarrow \frac{1}{6} [\bar{2}11] + \frac{1}{6} [\bar{1}2\bar{1}] \quad (6-3)$$

中，仅考虑位错能量时，由于 $b^2 > b_1^2 + b_2^2$ ，使能量降低而成为可能，当然在分解后的两不全位错之间将会有一层错。这种两根不全位错夹一片层错的组态称作扩展位错。图 6-11 中分别展示了不锈钢中扩展位错群的电子显微

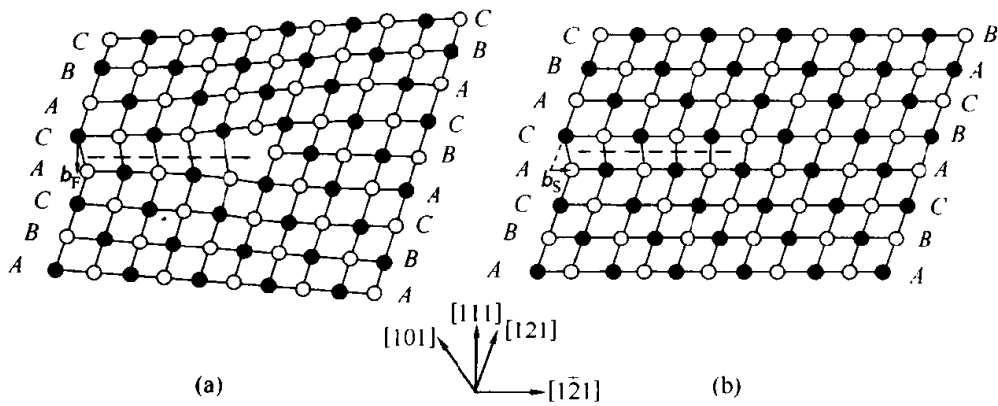


图 6-10 面心立方晶体中的弗兰克不全位错(a)和肖克莱不全位错(b),投影面为 $(10\bar{1})$

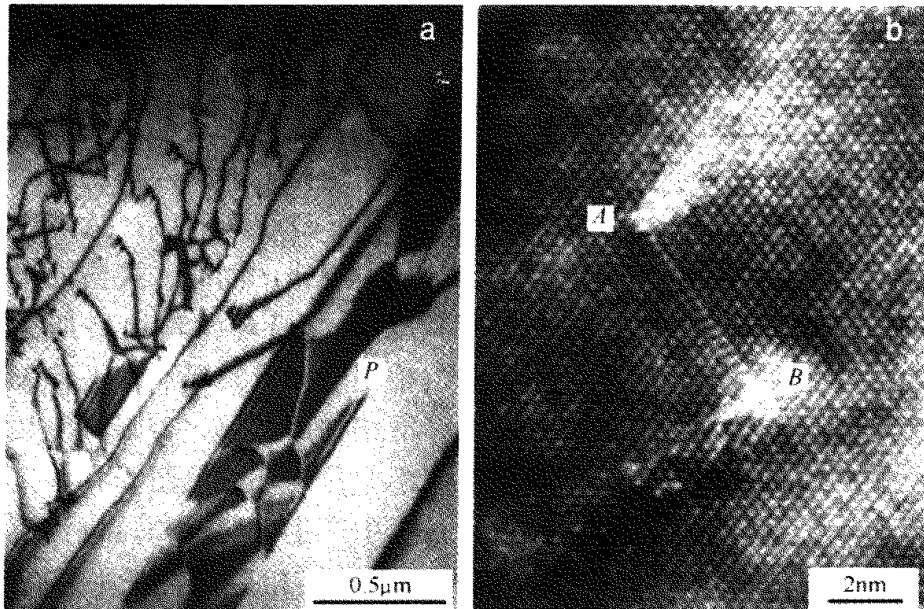


图 6-11 扩展位错的电子显微像

(a) 不锈钢中扩展位错的明场衍衬像；(b) 硅中扩展位错的高分辨点阵像

衍衬像^[5]和硅单晶中 60° 扩展位错的高分辨电子显微像^[6]。如果无相近的其他缺陷对它的干扰,则扩展位错的平衡宽度 d 决定于两不全位错间的排斥力与层错表面张力之间的竞争,而为:

$$d = \frac{G}{2\pi\gamma} \left(\frac{b_{1e}b_{2e}}{1-\nu} + b_{1s}b_{2s} \right) \quad (6-4)$$

由于扩展位错易于在电镜中观测到,其宽度可直接测量,因而材料的层错能极易测得。例如,图 6-11(b)所示硅单晶中的扩展位错,是由 60° 位错分解而得,层错边缘的不全位错分别为 $30^\circ(A)$ 和 $90^\circ(B)$ 不全位错。由层错宽度 d

$= 6.87 \pm 0.33 \text{ nm}$, 可算出层错能等于 $48.7 \pm 2.4 \text{ mJ/m}^2$ 。

近年来, 在许多具有层状结构的复杂氧化物晶体, 如各种高 T_c 氧化物超导体、电介质晶体中发现了多种形式的堆垛层错。这时堆垛层错相应于增加或减少了某些相应的原子层, 从而容纳对化学剂量配比的偏离。这类堆垛层错与前述有较大的不同, 不仅改变最近邻关系, 同时还改变配比, 因而为区别起见, 称之为堆垛错。图 6-12 是 $\text{TlBaCaCu}_2\text{O}_6$ 超导体中高密度堆垛错的高分辨电子显微像。这些复杂氧化物晶体常常存在多种配比的稳定相, 堆垛错相应的配比变化有可能局部地变成了另一配比的稳定相, 并不明显地增加晶体的错排能。在配比偏离很大时, 则可能造成不同配比相的交替生长, 有时甚至构成类似超晶格的规则排列。

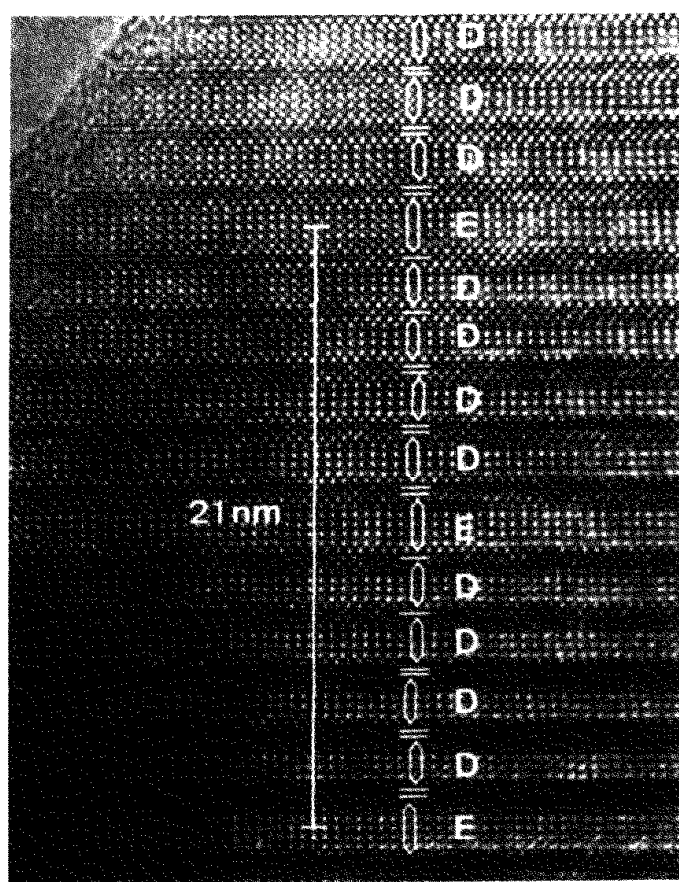


图 6-12 $\text{TlBaCaCu}_2\text{O}_6$ 超导体中高密度堆垛错有序排列形成的长周期结构

图中 D 和 E 分别代表有 3 和 4 层钙钛矿结构, 水平短线中的单线和双线

分别代表有一层或两层 Tl-O 层

[引自 Verwerft M, Van Tendeloo G, Amelinckx S. JEOL News. 1988, 26(E):33]

6.2.2 反相畴界

合金系统通常是无序固溶体, 但当合金原子比接近某些特定值, 例如

AB、 A_3B 、 A_2BC 等，并在相应的温度下热处理时，将转变成具有长程序的有序合金，即发生无序-有序相变。有序化的结果是使不同类原子有序地排列，无序相时的等同位置不再等同，因而晶体的对称性降低，形成了有序化的超点阵。以AB型的CuZn合金为例，如图6-13(a)所示，无序时Cu和Zn原子不规则地占据体心立方点阵的阵点；有序化后Cu和Zn则分别处于两个简单立方点阵的阵点上。两个亚点阵相互套叠，一个的阵点处于另一个的体心位置，组成一个复式格子，点阵类型由体心立方转变为简单立方。Cu和Zn由原来的无规律配对转变为全部的异类原子配对。

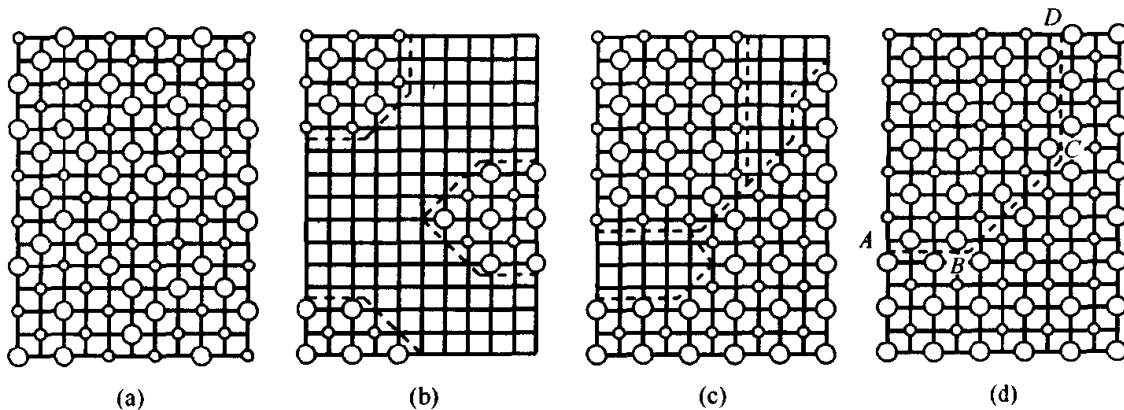


图 6-13 无序-有序相变时有序畴的形成过程示意图

- (a) 无序态；(b) 局部有序核的形成，空白区仍为无序态；(c) 有序核长大后，局部区域接触，反位相部分（中部）形成反相畴界，同位相部分合并；
(d) 全部有序化。注意 BC 部分为保守畴界，AB 和 CD 部分为非保守畴界

当这种无序-有序相变发生时，首先会形成许多有序核，见图6-13(b)，核长大最终全部会合完成有序化转变。不难想像，相邻有序核中不同类原子所占位置可以有两种不同的情况，一种是同类原子占据原无序相的等同位置，如CuZn合金中Cu占据原体心位置；一种是占据不等同位置，如一个核中Cu占据原体心位置，另一核中Cu占据顶点位置。当其长大会合时，正如图6-13(c)所示那样，前者合二为一，后者在整个界面上形成原子占位的位相差。于是晶内分成了许多区域，区内保持完整的有序结构，称有序畴；相邻区域间有一非点阵平移，但仍保持共格，只在界面处由正常的配对状态转变为非正常配对状态。这种界面称为反相畴界，也是一种低能量的面缺陷。从图6-14的高分辨电子显微像可以看出， Al_3Li 合金中已经形成了几个有序核。核内的超点阵清晰可见，其中图右部的两个核具有相同位相，但与左部的核则有不同的位相。图6-15中展示了 Ni_4Mo 合金中的反相畴界电子显微像，前者是其畴界的衍衬像，后者是高分辨电子显微像，可以明显看

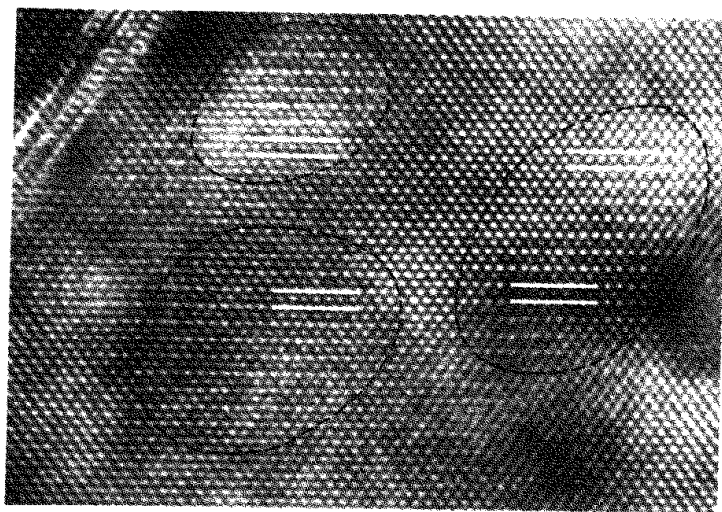


图 6-14 Al_3Li 合金中的有序核
核内超点阵清晰可见。注意到左右两部分有序核间位相的差异
(引自孙东立, 李齐等, 1992, 未发表)

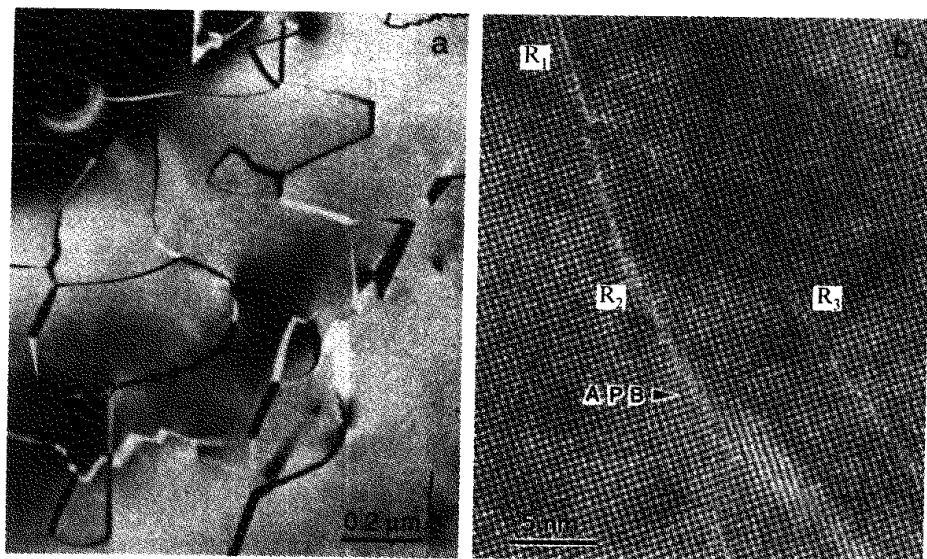


图 6-15 Ni_4Mo 合金中反相畴界的电子显微像

(a) 衍衬像 (引自 Amelinckx S. Dislocations in Particular Structures. In Nabarro F R N. ed. Dislocations in Solids, Vol. 2. Amsterdam: North-Holland, 1979, 246); (b) 高分辨像
(引自 Spruiell J E, Stanbury E E. J. Phys. Chem. Solids, 1965, 26: 811)

出相邻畴间的非点阵平移。

如前所述, 有序化造成的超点阵, 使无序相的某些晶体学方向的周期变大了。上述例子中无序相的最短点阵平移矢量为 $1/2\langle 111 \rangle$, 有序化后该方向的平移矢量则是 $\langle 111 \rangle$ 。因而, 原结构中的 $1/2\langle 111 \rangle$ 全位错, 变为超点

阵中的不全位错，当这样的超点阵位错扫过时，将在晶体中形成反相畴界。与此类似，如果反相畴界能量不高，有序相时的 $\langle 111 \rangle$ 全位错也可以分解为一种特殊形式的扩展位错，即由两个不全位错（相当于无序相时的全位错）夹一片反相畴界的组态。

反相畴界与层错同属平移界面，都是由于一个非点阵平移而造成的，但是这两者也有较大的不同之处。反相畴界的特征平移矢量是有序化超点阵中的非点阵平移矢量，如上述例子中的 $1/2\langle 111 \rangle$ ，却是无序相中的平移矢量，因而它是完整有序化相中连接被占位置间的矢量。而层错则不同，其相应的特征位移矢量通常都不是完整晶体中连接被占位置间的矢量。对于层错而言，层错面的任意改变，造成大量最近邻的变化和晶格畸变，能量迅速升高，因而通常只能取严格的密排原子面。反之，反相畴界面的改变，只改变最近邻的配对情况，不引起晶格畸变，因而能量变化不大，有些有序合金中反相畴界具有择优取向，更多的则呈空间曲面状。正因此，如图 6-13(d) 所示，即使是同一反相畴界面，其某些部分可以保持其化学配比而为保守界面，某些部分则偏离化学配比而为非保守界面。在许多情况下，固溶体对化学配比的偏离是以一定量的非保守反相畴界的方式来容纳的。一般说来，随组分偏离增大，反相畴界密度加大，当密度足够大时还会因较强的交互作用而有序地排列起来。

复杂氧化物晶体中的反相畴界也是很常见的。 $\text{Pb}_2\text{ScNbO}_6$ 和 $\text{Pb}_2\text{ScTaO}_6$ 等晶体是一类，其中 Sc 和 Nb 或 Ta 处于钙钛矿结构中的 B 位置，高温时无序分布，低温时有序化。高分辨电子显微术研究发现，相变初期首先出现仅数纳米大小的有序核，然后长大并交汇，反位相者形成反相畴界。 $\text{Ba}_2\text{NaNb}_5\text{O}_{15}$ 等晶体则是另一种情况，于 300°C 铁电相变时，四方相转变为正交相， $a_0 = \sqrt{2} a_t$ ， $b_0 = \sqrt{2} b_t$ ，单胞体积增大一倍。与此同时， Ba^{2+} 在 (001) 面内的微小位移使上下两相邻单胞中共顶点氧八面体作反方向的微小旋转，使 $C_0 = 2C_t$ ，单胞内两分离的氧八面体同样也作反向旋转，于是将可能出现反相畴界。

6.2.3 晶体学切变面

以前说过，点缺陷、点缺陷群以至点缺陷的有序化结构是晶体中容纳组分偏离的有效途径之一，但它所能容纳的偏离是相当有限的。人们早就发现，过渡金属氧化物 MoO_3 、 ReO_3 、 WO_3 、 TiO_2 及其复合物中，氧与金属原子数之比可在很大的范围内变化而容纳相当大的缺氧量。这些氧化物结构的基本单元是 RO_6 氧八面体，它们是共顶点连接。形成缺氧的非化学计量晶体时，最简单的途径是使晶体的两部分沿某一晶面相互切移，使界面处共顶点氧八面体部分地或全部转变为共边氧八面体组态。这将使某些原为两个

氧八面体共有的氧原子变为三个氧八面体所共有，从而省下氧原子。与此类似，某些共边氧八面体的组态通过类似的相互切移也可以变为共面氧八面体，从而省下氧原子。这一过程产生了一个平移界面，称作晶体学切变面。与此相关的相对平移矢量称晶体学切变矢量，它是非点阵平移矢量。

图 6-16 我们画出了 ReO_3 型结构中沿不同晶面作 $1/2[101]$ 切移所产生的晶体学切变面的平面图形。不难看出，当晶体学切变面为 $(P01)$ 时，界面处则由规则分布的 $2 \times P$ 个共边氧八面体组所组成，每组均较正常结构时少 $(P-1)$ 个氧原子。于是，随着 P 的增大，化学配比的偏离加大。图 6-17 则是 $\text{W}_n\text{O}_{3n-2}$ 晶体中 (102) 晶体学切变结构的高分辨电子显微像。高密度的晶体学切变面之间的交互作用使之排成了有序的层状切变结构。

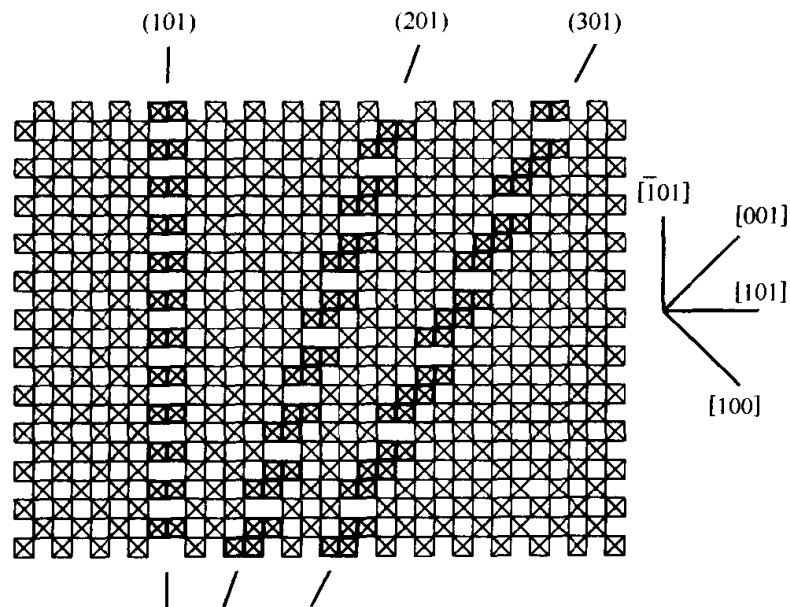


图 6-16 ReO_3 型结构中沿不同的 $(P01)$ 晶面作 $1/2[101]$

切移所产生的晶体学切变结构的 $[010]$ 视图

粗线勾画出的是 $2 \times P$ 个共边氧八面体群

在这类氧化物晶体中的另一类切变结构，则是由相交的两组晶体学切变面组成的块状切变结构，它可以容纳更大的缺氧量。有趣的是切变结构中有序排列的层状或块状组态中，还可以形成类似于位错、孪晶、层错和小角度晶界等缺陷组态。当然，它们不同于通常意义上的晶体缺陷，而是与组分的变化相关联，因而对容纳化学配比的偏离也有所贡献。

近年来对高 T_c 氧化物超导体及其他复杂氧化物晶体的研究中发现，由于它们也是以氧八面体作为重要的结构单元，因而也可能存在晶体学切变结构。

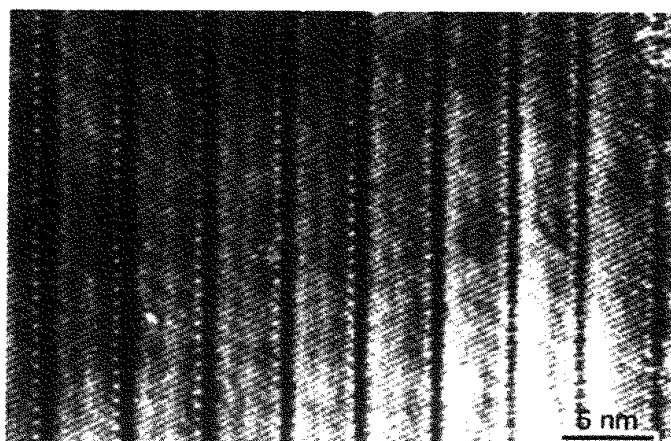


图 6-17 W_nO_{3n-2} 晶体中高密度 (102) 晶体学切变面有序排列的高分辨电子显微像

6.3 孪晶界面^[7]

6.3.1 反映孪晶界面

上节中已经知道，面心立方结构的晶体中的正常堆垛方式是六方密排面作……△△△△△△△△……的完全顺顺序堆垛（或与此等价，作……▽▽▽▽▽▽▽▽……完全逆顺序堆垛）。在正常顺序堆垛中出现一层或相继两层的逆顺序堆垛，则产生抽出型或插入型层错。如果从某一层起全部变为逆顺序堆垛，例如……△△△△▽▽▽▽……，那么这一原子面显然成为一个反映面，两侧晶体以此面成镜面对称（见图 6-18）。我们说这两部分晶体成孪晶关系，由于两者具有反映关系，称反映孪晶，该晶面称孪晶界面。

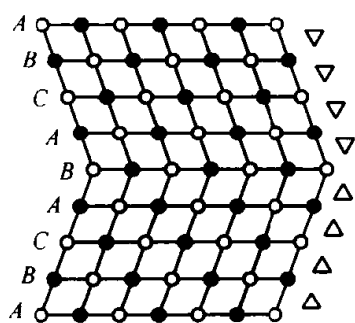


图 6-18 面心立方晶体中 {111} 面反映孪晶的 $\langle 110 \rangle$ 投影图

容易看出，沿着孪晶界面，孪晶的两部分完全密合，最近邻关系不发生改变，只有次近邻关系才有变化，引入的原子错排很小，这种孪晶界面称共格孪晶界面。考察图 6-9 中的层错图像，不管是抽出型抑或是插入型层错，都相当于有两个紧密相邻、仅一个面间距的反映孪晶界面。如果这两个孪晶界面所引起的电子扰动和弹性畸变很小，因而可以不考虑它们的相互作用时，孪晶界面的能量约为层错能之半。

从另一角度看，反映孪晶自孪晶界面开始的逆顺序堆垛，使其相对于完整晶体而言，每堆垛一层都沿晶面作一非点阵平移 R ，整个孪晶则是通过一个均匀的正比平移，即切变来获得，其切变量

为 $S = R/d$, d 为面间距。在晶体的塑性形变中, 除去上一章中指出的由位错的运动、增殖来完成外, 通过正比切变产生孪晶来实现塑性形变, 在许多晶体中是又一种重要途径。在实际过程中, 这种以正比切变方式形成的孪晶称机械孪晶。由于机械孪晶与切变相联系, 因而由弹性相容性条件所决定, 孪晶界面必须是切变时晶体中的不变面, 而切变产生的孪晶其点阵与母晶体点阵相同, 孪晶界面应包含连接两个孪晶晶体的对称元素, 因而孪晶界面 (即通常所说的组合面) 不能是任意的。面心立方为 $\{111\}$ 面, 体心立方为 $\{112\}$ 面, 而密集六方晶体则为 $\{10\bar{1}2\}$ 面。

以上这些孪晶界面并非相应晶体中固有的反映面, 但通过机械切变使该面两侧晶体构成反映对称关系, 遂成孪晶。这种晶体中原本并不存在, 而在

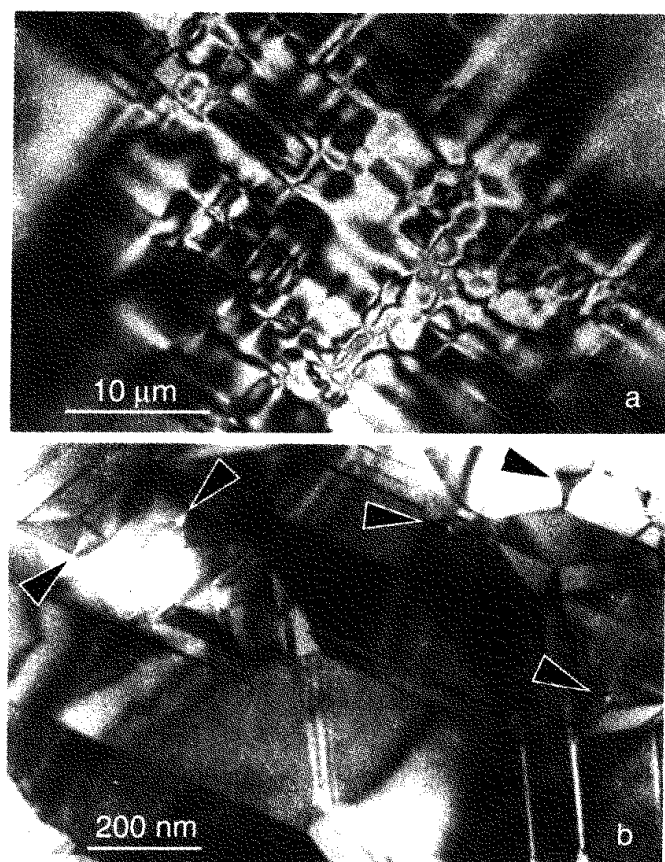


图 6-19 孪晶的光学和电子显微像

- (a) $\text{Ba}_2\text{NaNb}_5\text{O}_{15}$ 晶体中薄透镜片状反映孪晶的正交偏光光学显微像 ($[001]$ 观测方向)^[8]; (b) KNbW_2O_6 晶体中由 $\{110\}$ 和 $\{310\}$ 孪晶面所构成的复杂孪晶结构的电子显微衍衬像。注意箭头所指由大小三角星形孪晶多次套叠形成的特殊孪晶组态 (引自李齐, 1994, 未发表)

晶体学上又允许的对称操作，在某些物理过程中在晶内两部分晶块间出现，被称作孪晶化。于是就有生长孪晶、形变孪晶、退火孪晶、相变孪晶等多种成因的孪晶组态。在大量具有复杂结构的材料，尤其是氧化物电介质材料中，由于它们常常经历较多的相变，每一次相变都相应地有对称性的变化，如果对称性降低的相变中所损失的对称操作在相变后的晶体中的两部分之间重新建立起来，这就形成了孪晶。这其中，重新建立损失了的反映对称关系而形成的是反映孪晶， $\text{Ba}_2\text{NaNb}_5\text{O}_{15}$ 晶体^[8] 和 KNbW_2O_9 晶体中的反映孪晶就是这样的相变孪晶，见图 6-19(a)、(b)。

必须指出的是，面心立方结构的反映孪晶是最简单的。同为面心立方点阵，但结构不同的晶体，如金刚石结构和闪锌矿结构的晶体中情况却大为不同。如图 6-20 所示，金刚石结构晶体中的孪晶界面虽然仍为 $\{111\}$ 面，但却不与任何原子面重合，而处于图中点线位置，这时界面处不仅最近邻关系保持不变，而且与界面最近的原子及整个孪晶部分中三分之一的原子位置与孪晶前完全相同，因而是具有最低能量的反映孪晶界面。在闪锌矿结构的晶体中，由于由两类原子构成，因而以 $\{111\}$ 为孪晶界面可以有如图 6-21 的两种情况。图 6-21(a) 是反映孪晶，但界面处形成了同类原子配对，能量较高；另一种情况，即图 6-21(b) 中原子位置与前者相同，但原子类别相反，界面恢复异类原子配对，因而能量较低。但是后者不再是反映孪晶，而是旋转孪晶。

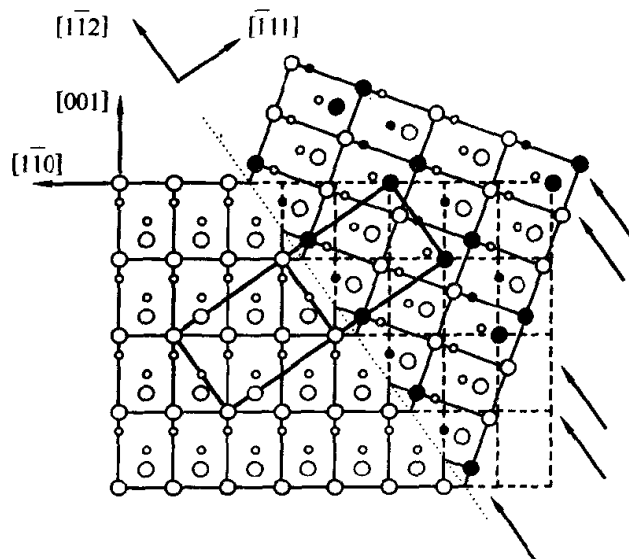


图 6-20 金刚石结构晶体中反映孪晶界面不与原子面重合
注意箭头所指晶面中原子（以实心圆表示）为孪晶化前后点阵中的重合位置原子，粗实线所围长方形为此二维点阵中的重合位置点阵单元

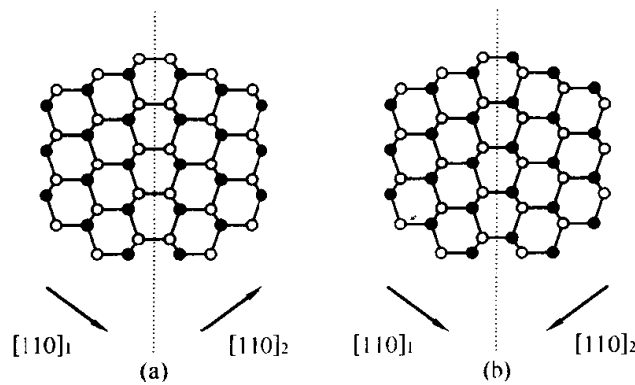


图 6-21 闪锌矿结构晶体中的两种 $\{111\}$ 孪晶界面

在一些有序合金材料中，由于孪晶两部分中的超点阵可以不构成完全的反映关系，而是在界面上附加一个超点阵的非点阵平移，从而形成一种特殊的平移孪晶界面。

6.3.2 旋转孪晶界面

旋转孪晶是孪晶两部分可以通过晶体学允许的旋转对称操作而相互重合的一类孪晶。上述闪锌矿中的旋转孪晶就是以界面法线($\langle 111 \rangle$)为孪晶旋转轴的 60° 旋转孪晶。由于它具有比反映孪晶界面低得多的界面能量，因而在闪锌矿结构的晶体中多为此类界面。不难想像，对于那些有两类以上原子组成，且不同配对状态能量有较大差异的晶体而言，反映孪晶往往于能量不利，旋转孪晶才是合适的孪晶组态。这在下面讨论铁电畴界时可以得到进一步证实。

6.3.3 铁电畴界

一个铁电晶体，在居里点以上是顺电相，晶体结构具有中心对称性；居里点以下为铁电相，不再具有中心对称性。在降温通过居里点时，位移型铁电晶体，如 BaTiO_3 、 KNbO_3 、 LiNbO_3 、 $\text{Ba}_2\text{NaNb}_5\text{O}_{15}$ 等含有氧八面体结构的氧化物晶体，通过单胞内正负离子的相对位移产生电偶极矩；无序-有序型铁电晶体，如 KDP、TGS 等含有 $\text{O}-\text{H}-\text{O}$ 键或 $\text{N}-\text{H}-\text{O}$ （或 $\text{N}-\text{H}$ ）键的化合物晶体，则通过无序的键取向到有序键取向的转变产生电偶极矩，使晶体发生自发极化现象。一般说来，晶内各部分自发极化的方向可以是不同的，我们将自发极化方向相同的区域称作铁电畴，其间的界面称铁电畴界。

铁电相变使晶体对称性降低。以三种重要结构类型的位移型铁电晶体为例，钙钛矿结构的晶体（ BaTiO_3 、 KNbO_3 等）由 $m3m$ 三方相转变为 $4mm$ 四方相，失去垂直于四次轴的反映面[见图 6-22(a)~(b)的转变]；钨青铜型晶体（BNN、SBN 等）由 $\frac{4}{m}mm$ 的四方相转变为 $4mm$ 的低对称性四方相，

同样失去垂直于四次轴的反映面；而如 LiNbO_3 、 LiTaO_3 等类钙钛矿型晶体，则由 $\bar{3}m$ 三方相转变为 $3m$ 三方相，也失去反映面。然而铁电相里，铁电畴之间可以通过晶体学允许的对称操作而使结构完全重合，因而从晶体学上讲，它们首先是孪晶关系，无一例外。

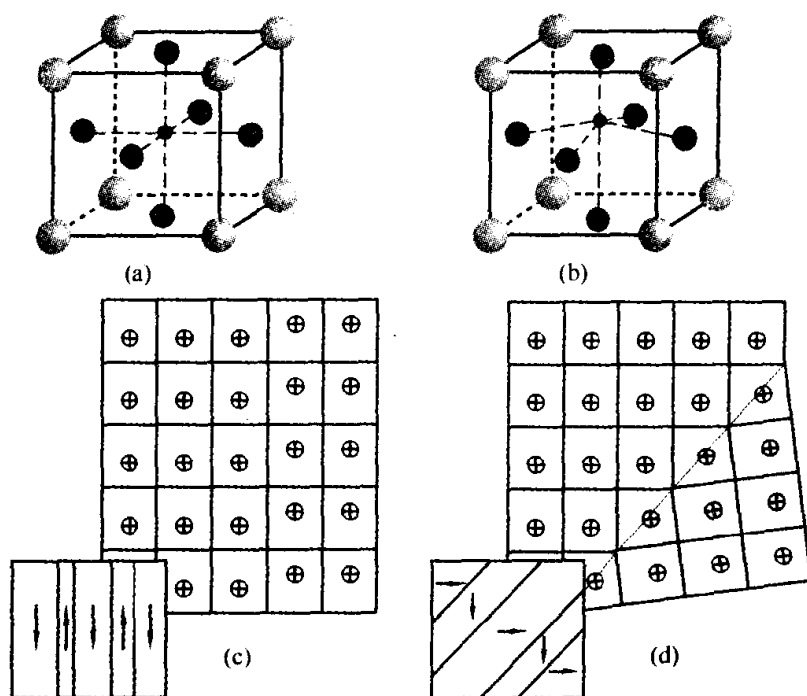


图 6-22 钙钛矿型结构 ABO_3 晶体的 (a) 立方顺电相及 (b) 四方铁电相的单胞原子分布，浅色球为氧，深色球为正离子；(c) 180° 畴及 (d) 90° 畴结构示意图

但是，在铁电畴界与一般的孪晶界面之间则存在一个重要的不同：由于畴内电偶极矩方向都一致，因而必然存在一个空间电荷分布，铁电畴界的取向和相邻铁电畴极化方向间的关系将决定其表面有无电荷积累，只有当极化矢量在畴界面处头尾相接，且法向分量相连续时，无电荷积累而有最低的静电能量。因而，在考虑铁电晶体中可能存在的低能量畴界面时，除去必须满足孪晶界面的晶体学关系以外，还必须考虑其为无电荷积累的低能量取向。例如，钙钛矿晶体四方铁电相中，如果以四次轴为法线的晶面作为畴界面，形成 180° 畴，则它是反映孪晶，但却是静电能最大的组态，实际存在于晶体中的是图6-22(c)所示的组态，其畴界面平行于极化方向，为 180° 旋转孪晶；如果相邻两铁电畴的极化矢量相互垂直，分别平行于立方相中不同的四次轴，形成 90° 畴，那么，显然它们必须形成如图6-22(d)所示的组态，它们也不是反映孪晶，而是以畴界面法线为转轴的 180° 旋转孪晶。图 6-23 给出

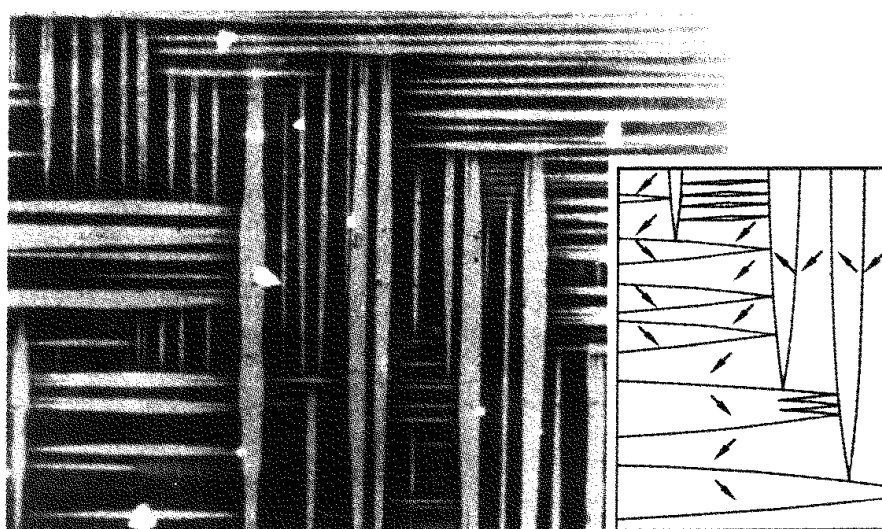


图 6-23 KNbO_3 晶体中片状 90° 铁电(铁弹)畴的偏光

光学显微像及其极化矢量取向的示意图

注意到由于相互垂直的片状畴具有反向的极化矢量,构成

图中组态时,不相互穿过而不出现高能量的畴界面

(引自陈峻,李齐,冯端,1984,未发表)

了 KNbO_3 晶体中 90° 畴的畴结构。

必须指出的是,虽然铁电畴间的关系毫无例外地是孪晶关系,但是铁电晶体中并非所有孪晶都构成铁电畴,只有那些孪晶化伴有极化方向改变的孪晶才同时是铁电畴。例如, LiNbO_3 晶体中以 $\{10\bar{1}2\}$ 晶面为孪晶界面的孪晶,

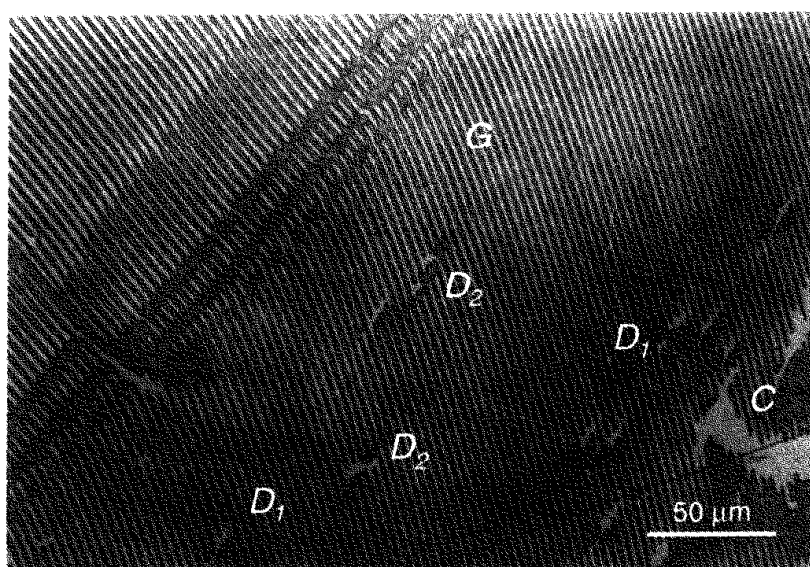


图 6-24 LiNbO_3 晶体中由生长层(G)、多种位错(D_1 、 D_2)

及组分过冷区(C)引发的各种特征形态的铁电畴组态^[8]

因极化方向改变而成 135° 铁电畴；但 $\text{Ba}_2\text{NaNb}_5\text{O}_{15}$ 晶体的铁电相变由四方顺电相转变为四方铁电相时，极化矢量沿惟一的四次轴方向，形成 180° 铁电畴，而发生于 300°C 时的相变由四方铁电相转变为正交铁电相，但极化方向不变，同时孪晶界面平行于它，因而孪晶不构成铁电畴关系，见图 6-19(a)。

观测表明，晶体的不均匀性和多种缺陷可以产生特征形态的铁电畴，甚至极大地改变其平衡态畴结构（见图 6-24）。例如，局部表面压印和晶内包裹体可使 LiNbO_3 晶体中产生针状微电畴，位错线可以诱发平行于位错的管状（LN 和 BNN）或薄透镜片状（KN）铁电畴^[8]。而通过旋转性生长层在掺杂的 LN 和 LT 等晶体中诱发聚片状铁电畴，已经成功地发现了多种可兹应用的非线性光学效应。

6.4 晶界

晶态材料通常是以多晶状态存在的，我们称不同晶粒之间以晶态相连的界面为晶粒间界，简称为晶界。较之前面所讨论的多种界面而言，晶界是有更大晶格畸变的界面，因而对于材料的性能（如力学性质、扩散等）有重大的影响，在晶体缺陷的研究中占有一定的地位。

6.4.1 小角度晶界

设想两块晶体学取向差较小的晶块相互接触时，如果它们保持各自的晶格完整性，那么将在界面处形成严重的原子错排状态，如图 6-25(a)，使界

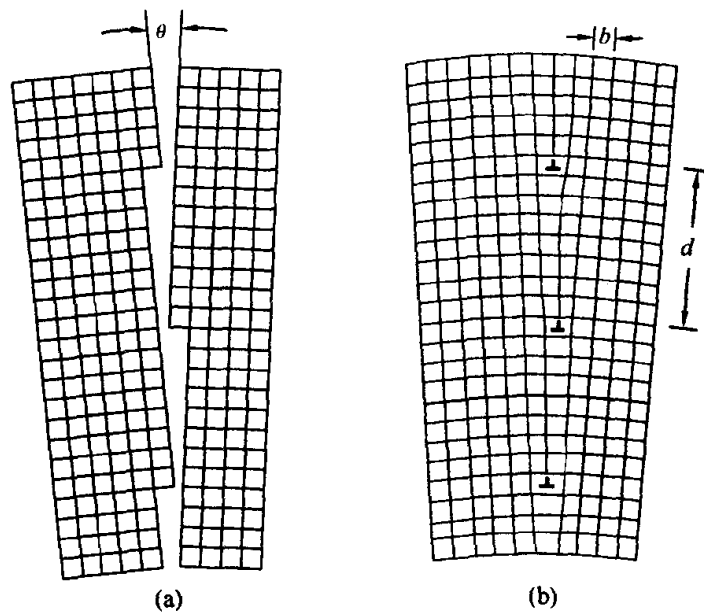


图 6-25 简单立方结构晶体中对称倾侧型亚晶界的位错模型二维图

(a) 弛豫前错排严重；(b) 弛豫后构成平行等间距的刃型位错列阵，仅位错附近局域错排

面能特别高。但是,若通过局域构成位错的方式,只需界面附近的原子位置作小量的调整,那么界面上位错以外的区域基本恢复完整,如图 6-25(b),大部分弹性畸变被弛豫,从而使界面能大为减小。这就是小角度晶界的位错模型,已为大量的实验观测结果完全确证。

由于确定晶界的几何学需要有五个旋转自由度,其中两晶粒间的相对旋转轴及旋转角共三个自由度,晶界面的法向矢量两个自由度(这里不计及两者的相对平移),因而小角度晶界的位错结构是多种多样的。具有简单位错结构的特殊晶界有两类,分别为倾侧型和扭转型晶界。

倾侧型晶界是相对旋转轴在晶界面内的一类晶界。如果晶界面平分相对旋转角 θ ,即两晶粒相对于完整点阵向相反方向各旋转相同的角度 $\theta/2$,则此晶界被称作对称倾侧型晶界。图 6-25 就是简单立方晶体中以(100)面为界面的对称倾侧型亚晶界。由图可知,这时的界面是由一组同号的刃型位错列阵所构成,它们相互平行且平行于相对旋转轴,柏格斯矢量垂直于晶界面,同时等间距排列。其间距为:

$$D = \frac{b}{2} \sin \frac{\theta}{2} \quad (6-5)$$

由于 θ 很小,可简化为:

$$D = b/\theta \quad (6-6)$$

对称倾侧型亚晶界在许多从熔体中生长的晶体中极为常见,而形变后经过充分退火的晶体中则更是其特征的位错组态,通常称为多边化(图 6-26)。

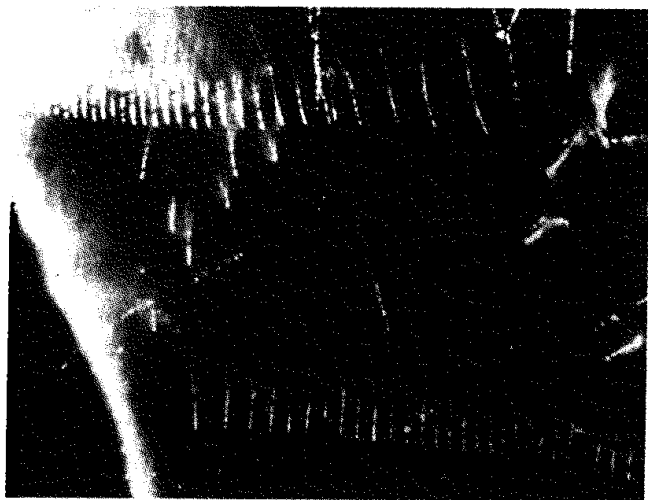


图 6-26 形变退火 KCl 晶体中多边化过程形成的对称倾侧型亚晶界。

位错经 Ag 缀饰处理而得以显示

(引自 Amelinckx S. Acta Met. 1958, 6: 34)

对于一个间距 h 、沿 z 轴平行排列于 x 面内的无穷刃型位错列阵,其应

力场:

$$\sigma_{xy} = \frac{Gb}{2\pi(1-\nu)} \cdot \frac{\pi^2 x}{h^2} \cdot 4 \exp\left(-\frac{2\pi x}{h}\right) \cos\left(\frac{2\pi y}{h}\right) \quad (6-7)$$

是随与界面距离加大而指数式递减的。当距离超过 h 数倍后就微不足道了。由此计算得的单位面积晶界能为:

$$E = E_0 \theta (A - \ln \theta) \quad (6-8)$$

式中的

$$A = 1 + \ln \frac{b}{2\pi r_0}, E_0 = \frac{Gb}{4\pi(1-\nu)} \quad (6-9)$$

易知, θ 较小时, 晶界能随 θ 以近似线性的方式增大。

如果倾侧型亚晶界的界面偏离对称情况, 称不对称倾侧型亚晶界, 这时将由柏格斯矢量分别为 $[100]$ 和 $[010]$ 的两组平行刃型位错组成, 其间距分别为:

$$D_1 = \frac{b}{\theta \sin \varphi}, D_2 = \frac{b}{\theta \cos \varphi} \quad (6-10)$$

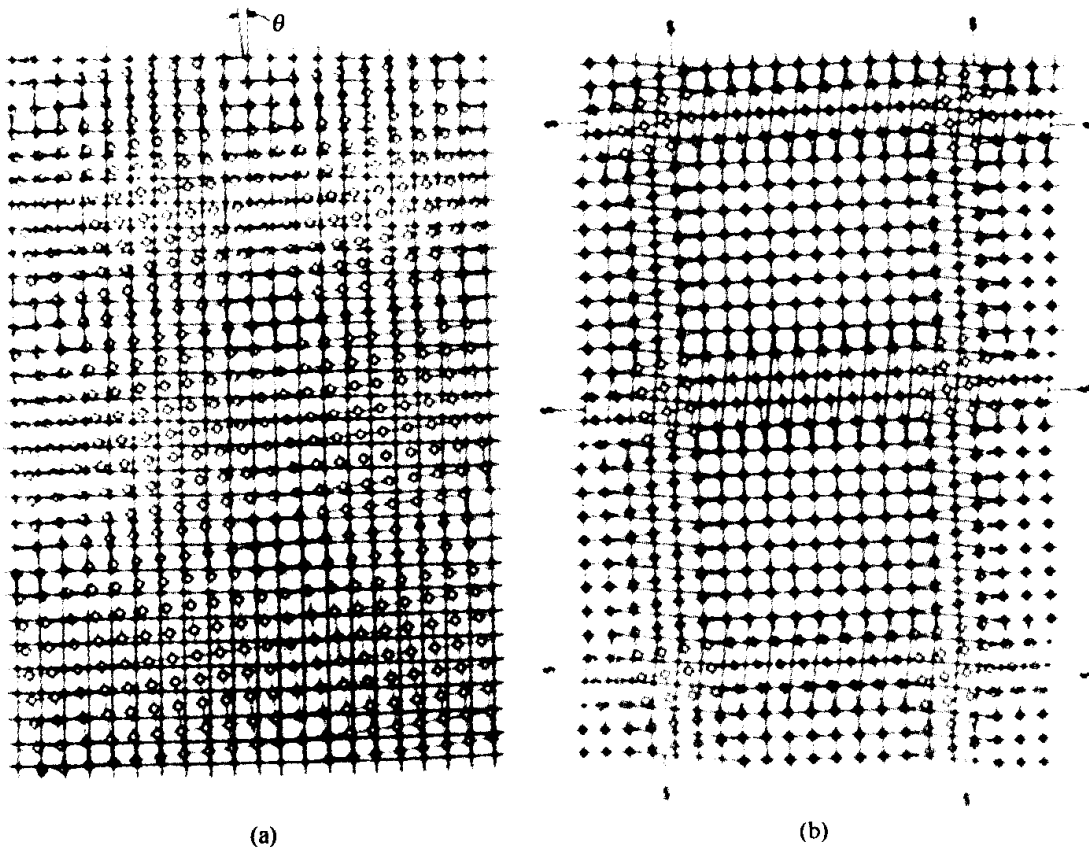


图 6-27 简单立方结构晶体中纯扭转亚晶界的位错模型二维图

(a)弛豫前,上(实心圆)下(空心圆)两原子层中原子基本处于错排状态;
(b)弛豫后,除构成方格状网络的螺型位错局域错排外,大部区域恢复完整

其中 φ 角为晶界面与 $[100]$ 方向间的夹角。

另一类特殊形态的小角度晶界是纯扭转型亚晶界，其相对旋转轴垂直于晶界面。如图 6-27 所示，简单立方晶体以 (001) 面为界面，以其法线为转轴作刚性旋转时，形成界面处严重错排，但弛豫后则形成了以两组相互垂直的螺型位错列阵构成的位错网络。网络的间距也满足：

$$D = b/\theta \quad (6-11)$$

这样的位错网络，其应力场分量在远离晶界处的近似式为：

$$\sigma_{yz} = \frac{Gb}{h} e^{-2\pi x/h} \left[\cos\left(\frac{2\pi y}{h}\right) - \cos\left(\frac{2\pi z}{h}\right) \right] \quad (6-12)$$

可知远程应力场甚小，可以构成稳定的界面。

一般情形的小角度晶界，其旋转轴和界面可以有任意的取向关系。其位错结构应不同于上述两种特殊晶界，通常它也是由两组（少数情况下由三组）位错构成的位错网络，但是晶界中的位错一般既不是纯刃型位错，也不是纯螺型位错。这时的位错网络也常常不是平行直线位错所构成，由于位错相交的结点处所发生的位错反应，会使结点转化为一段新的位错线，从而使晶界变成具有六边形网络的位错网络。图 6-28 透射电镜衍衬像中所显示的 Nb 晶体中的位错网络就是一般型小角度晶界的典型例子，可以看出随着晶界面取向的改变而改变其六边形网络的形状。

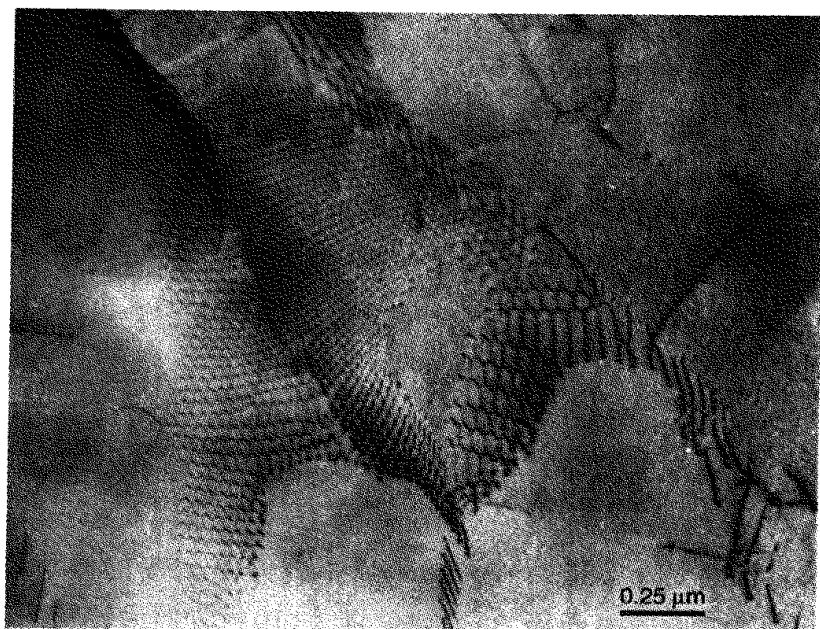


图 6-28 Nb 晶体中小角度晶界的六边形网络位错组态
网络形状的变化源于亚晶界面取向的变化

(引自 Berghezan A, Fourdeux A. 4ème Colloque Métallurgie-
C. E. N. Saclay. 1961, 127)

小角度晶界将单晶体分割成取向稍有不同的亚晶粒，在各向同性晶体中，它使进入晶体的光线的波平面发生畸变，对于各向异性晶体则使各部分的光轴不一致，因而对光学性质的影响较大。例如在固体激光中，当使用 Q 开关、声光调制、产生倍频时，输出线偏振光常常是必要的，小角度晶界存在引起的应力双折射及光轴变化都会使光波不能在同一偏振面内振动，从而使器件的输出功率大为降低。晶体中的位错虽然有可能通过长期退火而减少，但留下的位错可以排成低能量的位错界面而趋于稳定，因而要消除小角度晶界必须防患于未然，避免因各种原因将位错引入晶体内部。

6.4.2 大角度晶界

小角度晶界位错列阵的位错间距随着旋转角 θ 的增大而减小，当间距小到一定程度可与位错核心区的大小相当时，界面原子基本处于错排状态，单个位错已失去传统含义，以位错列阵或网络来描述晶界结构就不合适了。这一旋转角的限度大约在 10° 左右。我们称超过这一界限以致不能用位错模型来描述的晶界称大角度晶界。

早年提出的大角度晶界的结构模型，如过冷液体模型、小岛模型、无序群模型等，其基本构想都是认为晶界整体说来或局部说来是足够大的无序区，这些模型虽一再改进总不能圆满地说明有关晶界性质的实验现象。近年来，由于直接观察技术成功地揭示了大量有关大角度晶界结构的信息，计算机技术也已用来对晶界结构进行理论计算，特别是高分辨电镜术直接获得了许多晶界处原子排列组态的证据，对于大角度晶界结构的认识有了质的飞跃。

已被广泛认可的大角度晶界结构模型是重合位置点阵模型^[9]。如果将被晶界分隔的两部分晶体的点阵看成是能相互穿透的，不难想象它们将有一部分阵点是重合的，而且这些重合阵点也将构成一个新的点阵，这就是重合位置点阵。重合点阵与晶界面无关，但晶界面相对于重合点阵的取向关系对于界面结构却有重大影响。当界面为重合点阵的最密排面时，如同图 6-29 中标示为 ABCD 的晶界面中的 AB 和 CD 部分那样，其界面处具有最好的原子匹配度，错排较小，因而是界面能量最低的晶界。显然，并不是任意的取向关系都有高密度的重合阵点，对不同晶体结构的晶体而言，将有各自的特征取向关系，而且是不多的几种。如果以重合阵点密度的倒数 Σ 来表示重合度，那么对于面心立方晶体 Al 来说，以 $[110]$ 方向为转轴的晶界中，重合度高者有 3、9、11，分别对应于相对旋转 70.5° 、 39° 和 129.5° ，而这与对对称倾侧型晶界界面能测量中获得的三个能量极低值位置完全相同。

实际观测到的晶界从宏观上讲并非都与重合点阵的最密排面相符，但是高倍率的观测发现，在许多情况下即使是曲面状的晶界也常由一些较小的低

能量界面组合而成。晶界面取向对特殊界面的偏离可以通过界面台阶（如图 6-29 中的 BC 段）来调节，取向差对特殊取向差的偏离可以通过界面上的“结构位错”网络来容纳，其晶界从总体来讲仍处于较低能量状态。易知，大角度晶界的重合位置点阵模型中只有原子间距量级的厚度有局部错排，而不存在较大的错排或无序区。

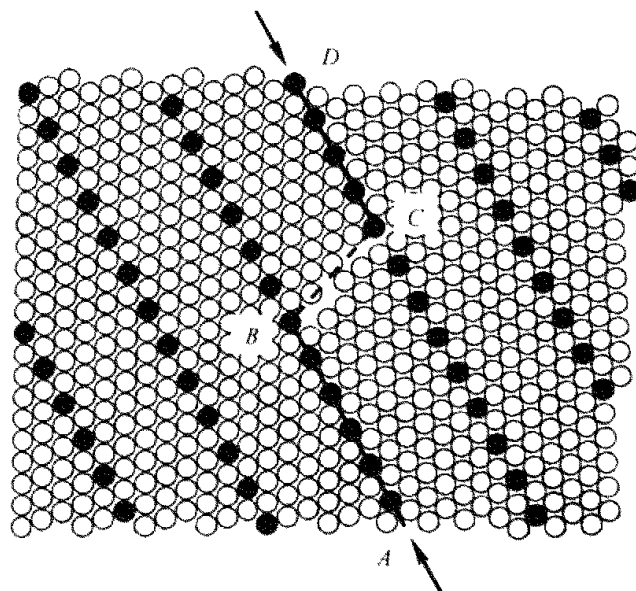


图 6-29 体心立方晶体中 $\Sigma=11$ 的重合位置点阵中的 $50.5^\circ/[110]$ 晶界模型

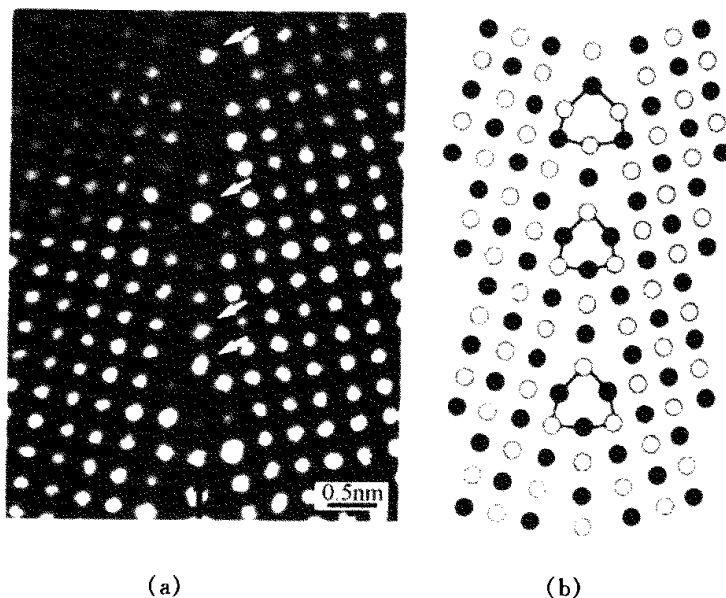


图 6-30 MgO 晶体中取向差为 24° 的 $\langle 001 \rangle$ 对称倾侧型晶界

(a) 高分辨电子显微像；(b) 结构单元示意图

(引自 Yan Y, Chisholm M F, Duscher G. et al. Phys. Rev. Lett. 1998, 81: 3675)

另一种已被大量观测结果所证实的晶界模型是结构单元模型^[10]。当取向差足够大以致不能用位错模型来描述时，可以想像，由于晶体结构的周期性，晶界面的错排状态也应具有某种周期性，通过原子位置调整而得到的具有最低交互作用能的界面组态应是一些特征的多边形原子组合，这就是结构单元。图 6-30 中显示的是 MgO 晶体中晶界结构单元组态的高分辨电子显微像及其结构单元示意图。

由于晶界区域包容了偏离正常点阵位置的大量原子，位错、杂质、第二相、甚或非晶材料等都能优先聚集到晶界上，因而晶界的存在将对材料的力学性质和输运性质等产生重大而复杂的影响。

6.5 相界^[11]

具有不同结构的两固相之间的界面称为相界。不同结构包括同质异构以及不同质材料，它们或结构对称性不同、或点阵参数不同、或键合类型不同，都使相界具有较复杂的结构组态。相界中界面完全有序，两相完全匹配者称共格相界；界面处失配通过弛豫使错配局限于错配位错处，其余大部分区域仅有甚小弹性畸变者称半共格相界；完全无序的界面则为非共格相界。

具有严格共格关系的相界是极为少见的。典型的例子是钴相变中出现的面心立方相与密集六方相的相界。由于两相具有相同的六方密堆积面，只在相界两侧具有不同的堆垛顺序，界面是两相共同的点阵面，因而是一种能量特别低的相界。在马氏体相变过程中或脱溶沉淀初期，某些相界也可以看成是共格相界，这时相界的一侧是原相作一适当畸变而来，但是平行于相界的晶面内的原子排列不存在任何变化。这种畸变称为不变平面应变，事实上，这是这类相中出现惯态面的晶体学基础。

半共格相界是极为常见的一类相界。随着微电子等现代技术的飞速发展，大量人工微结构材料被制备和应用，其中外延生长的薄膜材料和多层膜材料是极为重要的一类。由于相界结构对其性能有着重要影响，为材料科学工作者所密切关注，而其中最突出的相界就是半共格相界。

如果外延生长的基质晶体和外延层的晶格参数 a_s 和 a_0 略有差异，则其失配度为：

$$f = (a_s - a_0) / a_0 \quad (6-13)$$

在外延生长初期，这种失配可以用弹性形变来容纳以保持其共格，这时界面为共格相界。随着外延层厚度的增大，这种弹性畸变也将随之延续，保持共格将使外延层中积聚的弹性能愈来愈大。简单的计算可知，膜厚为 h 的薄膜内弹性应变为 ϵ 时的弹性能为：

$$E_{\epsilon} = \frac{2G_0(1+\nu)}{(1-\nu)} h\epsilon^2 \quad (6-14)$$

式中, G_0 为外延材料的切变模量。当膜厚过大, 以致弹性能量过高而不能为外延材料所容忍时, 外延层将倾向于恢复无畸变状态, 使弹性能尽量减小。但它与基质在界面处的失配重新出现, 错排能骤然升高, 幸好这可以通过在界面上产生错配位错列阵或网络, 将错配局限于位错处而降低相界能量, 于是形成如图 6-31 的半共格相界。

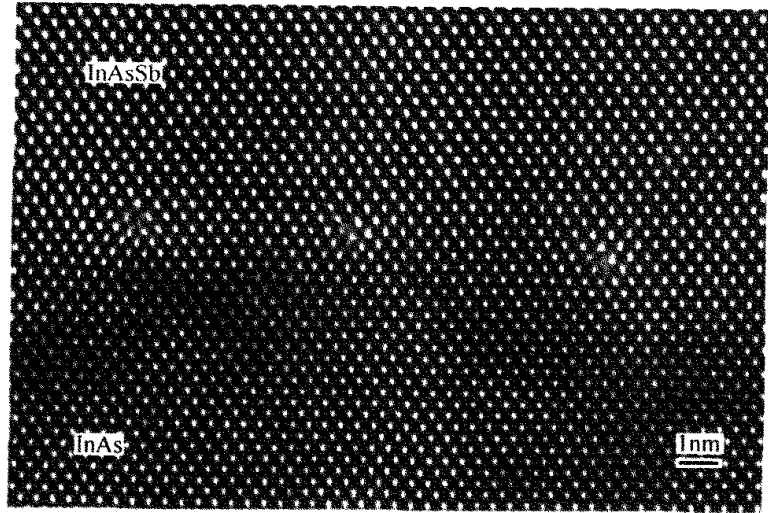


图 6-31 InAsSb/(001) InAs 界面上错配位错的 $\langle 110 \rangle$ 高分辨点阵像
(引自 Seong T Y, Norman A G, Booker G R, et al. Proc. MRS Meeting, 1989)

如果平行于膜面还残留有 ϵ' 大小的弹性应变, 那么错配位错的间距应为:

$$l = b / (f - \epsilon') \quad (6-15)$$

于是单位面积相界中错配位错网络的能量为:

$$E_d = \frac{G_0 G_s (f - \epsilon')}{\pi (G_0 + G_s) (1 - \nu)} \left(\ln \frac{R}{b} + 1 \right) \quad (6-16)$$

式中, G_s 为基质晶体的切变模量, R 为位错应力场范围, 如果 $h < 1/2$, 可视为 $R = h$ 。则系统总能量为:

$$\begin{aligned} E &= E_d + E_{\epsilon'} \\ &= \frac{G_0 G_s (f - \epsilon')}{\pi (G_0 + G_s) (1 - \nu)} \left(\ln \frac{h}{b} + 1 \right) + \frac{2G_0(1+\nu)}{(1-\nu)} h\epsilon'^2 \end{aligned} \quad (6-17)$$

对此求极小, 对应的 ϵ' 值为:

$$\epsilon^* = \frac{G_s b}{4\pi (G_0 + G_s) (1 - \nu)} \left(\ln \frac{h}{b} + 1 \right) \quad (6-18)$$

易知, 如果按此算出的 ϵ^* 等于或大于 f , 则薄膜将单纯通过共格弹性应变来容纳错配, 不产生位错; 如果算得的 ϵ^* 小于 f , 则将有一部分错配

$f - \epsilon^*$ 用错配位错来容纳。因而错配位错出现，即转变为半共格相界的条件为 $\epsilon^* = f$ 。对应的薄膜临界厚度为：

$$h_c = \frac{G_s b}{4\pi(G_0 + G_s)(1 - \nu)} \left(\ln \frac{h_c}{b} + 1 \right) \quad (6-19)$$

实践表明，金属膜的 h_c 值与理论估计大体相符，但半导体和氧化物薄膜中的 h_c 值通常比理论值大，这主要是由于后者位错成核激活能较高不易产生的缘故。

脱溶沉淀现象在许多合金材料中都能发生，由此产生的沉淀相可以取各种不同的形状，它们与母相间在界面处的失配会随着两相间取向的差异和界面取向的差异而变化，因而相界面的错配位错组态可以是相当复杂的，图6-32分别显示了两种合金材料中不同形态沉淀相与母相间界面的错配位错网络，可见它们具有不同的形态。这种界面位错在应力作用下可以成为晶体中的位错源，当然也可通过弹性交互作用吸引近旁的位错线，成为位错线的尾间。

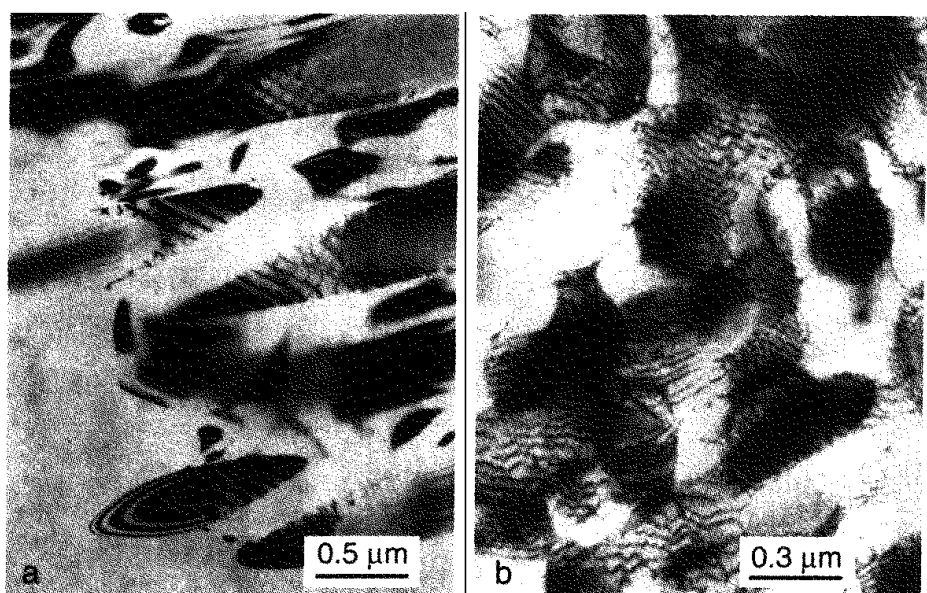


图 6-32 相界结构的电子显微衍射像

(a) 含 Cr 的 NiAl 合金中，富 Cr 柱状沉淀相表面的错配位错网络

(引自 Walter J L, Cline H E, Koch E F. Trans. AIME. 1969, 245: 2073);

(b) Cu-Ni-Fe 合金 Spinodal 失稳分解产物与母相间界面的错配位错网络

(引自黄孝瑛，侯耀永，李理，电子衍射分析原理与图谱，

济南：山东科学技术出版社，2000，211)

参 考 文 献

- 1 Porter D. A. Easterling K E, Phase Transformations in Metals and Alloys, 2nd ed. London:

- Chapman & Hall, 1992, 115
- 2 Blakly J. M. Introduction to the Properties of Crystal Surfaces. Oxford: Pergamon Press, 1973, 1
 - 3 Marks L D. Phys. Rev. Lett. 1983, 51: 1000
 - 4 Tiller W A. The Science of Crystallization: Microscopic Interfacial Phenomena. Cambridge: Cambridge University Press, 1991. 44
 - 5 黄孝瑛, 侯耀永, 李理. 电子衍射——分析原理与图谱. 济南: 山东科学技术出版社, 2000.215
 - 6 Sato M., Hiraga K., Sumino K. Jpn. J. Appl. Phys. 1980, 19: L155
 - 7 Cahn R. W. Adv. Phys. 1954, 3: 363
 - 8 李齐. 物理学进展. 1993, 1(2): 141
 - 9 Bollmann W. Crystal Defects and Crystalline Interfaces. Berlin: Springer-Verlag, 1970
 - 10 Sutton A. P. Vitek V. Philos. Trans. Roy. Soc. Lond. 1983, A309: 1, 37, 55
 - 11 冯端等. 金属物理学. 第一卷, 结构与缺陷. 北京: 科学出版社, 1987.463

第7章 多层次与非均质材料

撰稿人 濮 林 冯 端

凝聚态的物质具有众多的结构层次，它们之间的相互作用造就了丰富多彩、复杂异常的世界。本章主要介绍非均匀复相系统，以及逾渗和分形这两个在非均匀系统中起重要作用的基本概念^[1~4,6,7]。

7.1 非均质材料

7.1.1 概述^[1]

一般而言，材料科学总是从三个不同的尺度——即宏观的(macroscopic)、介观的(mesosopic, 广义上讲是指结构尺度处于宏观与微观的中间尺度，在 nm~ μ m 范围)和微观的(microscopic)层次上来研究材料结构特征与物性的关系。而材料的物性和这三个层次上的结构特征都有关联。

前面几章所讨论的材料从宏观尺度上看，材料显得均匀。但是光学显微镜乃至更现代的透射电子显微镜(TEM)、扫描电子显微镜(SEM)等技术已经揭示了在介观尺度内乃至微观尺度内，材料总是存在或多或少的不均匀性。常见的材料如晶体、陶瓷、玻璃、金属和高聚物等。本章所讨论的材料的不均匀性可分为两类。

① 宏观不均匀性。这类材料如不同相的混合物，也包括一部分人工复合材料。

② 介观不均匀性。在大于晶格常数的尺度内，晶粒生长、失稳分解(spinodal decomposition)和共析(coprecipitation)现象常造成这类静态的不均匀性。多晶陶瓷、玻璃-陶瓷等复合材料中常具有结构不均匀性，这类不均匀性也常被称为微结构(microstructure)、超微结构(super-microstructure)和纳米结构(nanostructure)。

在无机复相材料中，微结构不均匀性司空见惯。它们在很大程度上直接决定了这些材料的物性。我们知道，构成复相材料的各单一相按一定的统计规律分布在材料中，如果单一相的浓度、结构或分布改变了，则复相材料的宏观物性将相应改变。所以当前材料科学尤其是纳米材料科学的首要着眼点在于人工设计、制备和组装各类非均质材料，通过人为引入两类不均匀性而获得具有独特性质的材料。

在本章的第2、3节,我们将讨论逾渗(percolation)理论与分形(fractal)几何在描述非均质材料结构和相变中的应用。逾渗转变(transition)是一类几何相变。在非均质材料中,非均匀结构的某些参数,如某一相的体积百分数发生连续变化,就可以看作是一种逾渗转变。以金属-绝缘体复合材料为例,逐渐增加材料中金属相的体积分数,首先材料的微结构将发生很大变化,当金属相的体积分数增大到一定程度时,原先分散的金属相晶粒将形成一个连续的网络,即发生逾渗转变,此时金属相的体积分数称为逾渗阈值(percolation threshold)。可以预见,这一复合材料的电学性质将在逾渗阈值附近发生突变,即从不导电变为导电。因此,逾渗理论,如逾渗颗粒的几何标度理论,可用以描述非均质材料的微结构特性及结构相变。在表征无规非均质材料微结构特性及结构相变方面,分形几何也是一种强有力的工具。这些无规材料如多孔颗粒、多孔膜或多孔体材料、金属-绝缘体复合材料中逾渗型金属颗粒和材料的粗糙表面或界面。

7.1.2 非均质材料的微结构特征^[1]

多晶陶瓷和复相复合材料等非均质材料中往往具有许多均匀的单相微晶颗粒,它们有不同的尺寸和性质。了解它们的微结构特征对定量解决微结构与宏观物性的关联问题很重要。对微结构与宏观物性关联问题的研究可从零维、一维和二维的分析着手。对复相非均质材料而言,通常我们感兴趣的微结构特征有以下三个方面:

- ① 确定材料中所有存在的相,以及每相的含量和性质特征;
- ② 确定各相的结构形貌特征,如尺寸、形状、晶粒取向和分布;
- ③ 确定材料中晶界和其他结构缺陷如位错、微裂纹和包裹物的特性。

复相非均质材料中的某一材料相可以是晶体(各向同性或各向异性)、非晶或玻璃态,或为孔(把孔隙视为第二相)。每一相的相对含量可以表示为摩尔分数、质量分数或体积分数。

晶粒大小是表征复相非均质材料微结构特征的一个最重要的参量。尽管这一参量对材料物性的影响并不总是很清楚,但材料制备中还是有许多方法对其加以控制。晶粒形状和取向与原材料和制备方法有关。我们通常用有效半径 R 来表征球形、近球形或方形晶粒。有效半径 R 通常被表示为对数正态分布:

$$n(R') = \frac{1}{R' \sqrt{2\pi}\delta} \exp \left\{ - \left[\frac{\ln(R'/R)}{\sqrt{2}\delta} \right]^2 \right\} \quad (7-1)$$

其中 R 是晶粒的有效半径, δ 是偏差。对不同的材料,有效半径 R 可在几纳米至数十微米范围内变化。对非球形(椭球形)晶粒,可用其主轴的

半径 (a 、 b 、 c) 作为其大小的量度。其形状可用偏心率 $e_{p(0)}$, 和三个几何参量——退极化因子 (depolarization factors) L_x , L_y , L_z ($L_x + L_y + L_z = 1$) 来表征。

对长球形晶粒 (prolate spheroid, $a > b = c$):

$$\begin{aligned} L_x &= \frac{1 - e_p^2}{2e_p^3} \left[\ln \left(\frac{1 + e_p}{1 - e_p} \right) - 2e_p \right] \\ L_y &= L_z \\ e_p &= \sqrt{1 - \frac{b^2}{a^2}} \end{aligned} \quad (7-2)$$

对扁球形晶粒 (oblate spheroid, $a = b > c$):

$$\begin{aligned} L_x &= L_y \\ L_z &= \frac{1 - e_0^2}{e_0^3} (e_0 - \arctan e_0) \\ e_0 &= \sqrt{\frac{a^2}{b^2} - 1} \end{aligned} \quad (7-3)$$

上述的微结构特征参量, 如晶粒尺寸、尺寸分布、形状因子和体积分数, 可从对一实际为二维的样品定量测量来获得。

非均质材料的另一个微结构特征是单相晶粒的分布状态 (相分布)。复相材料的微结构将随着各相的体积分数的变化而变化, 所以相分布主要有三种几何形貌——分散晶粒结构、聚集晶粒结构和逾渗型晶粒团聚簇网络结构。考虑较为简单的两相系统, 它可以代表实际问题中的两相合金、复合陶瓷及多孔媒质 (将孔隙视为第二相)。按照第二相的百分比逐步增加的顺序, 图 7-1(a)~(d) 为其微结构的变化示意图。

① 分散晶粒结构 (dispersed grain structure)。当二相合金材料中含量小的相 (弱相, minor phase) 的体积分数较小时, 此相晶粒将随机分散在主相 (体积分数大的第二相, matrix phase) 中, 如图 7-1(a); 如果弱相的体积分数不太低, 则弱相的晶粒均匀地分散在主相中, 如图 7-1(b), 某些金属颗粒膜和玻璃-陶瓷材料就存在这种情况。

② 聚集晶粒结构 (aggregated grain structure)。当弱相的体积分数增大, 弱相晶粒将可能聚集而形成一定大小的晶粒团聚簇, 如图 7-1(c)。

③ 逾渗型晶粒团聚簇网络结构 (percolation-like cluster structure)。当弱相的体积分数增大到一临界值, 即逾渗阈值, 除了少数的晶粒团聚簇, 大多数的弱相晶粒将连接成一连续的随机走向的链状网络, 如图 7-1(d)。

图 7-1 所示的微结构基本上是各向同性的, 即材料中弱相晶粒的形状为

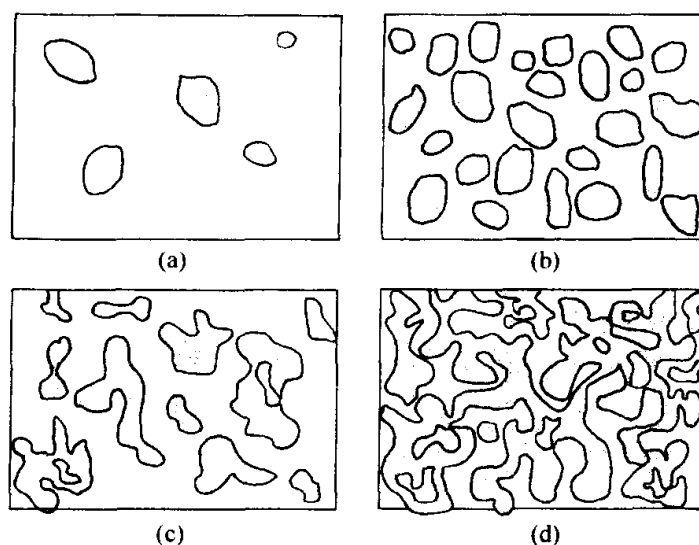


图 7-1 两相合金的微结构示意图

从(a)~(d)第二相的体积分数逐渐增大

球形或近球形，并且弱相晶粒取向随机、均匀地分布在主相中。进一步还可以将问题复杂化，考虑弱相晶粒的形状具有明显的各向异性，例如棒状、片状或扁球状的，这更接近于纤维增强复合材料、共晶及共析合金等的实际情况，并且这些弱相晶粒的取向分布有一定程度的一致性，那么这将造成材料微结构的各向异性(图 7-2)。

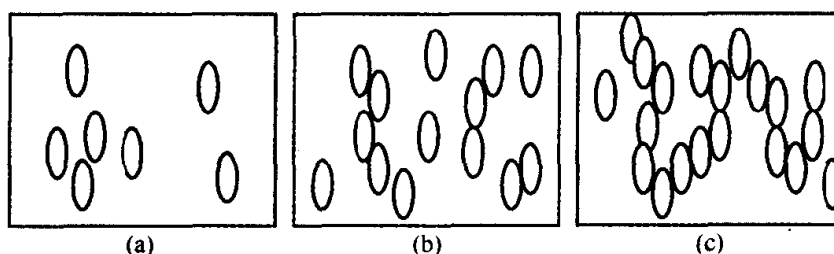


图 7-2 二相合金微结构各向异性示意图，

其中弱相椭球型晶粒取向一致分布

如果用方形晶粒(cubic grain)作为复相合金的缔合单元，而不像图 7-1 和图 7-2 那样采用球形或近球形的晶粒，那么我们将得到许多不同的晶粒连接方式。对于二相合金，有十种晶粒连接方式：0-0，0-1，0-2，0-3，1-1，1-2，1-3，2-2，2-3 和 3-3，如图 7-3 所示。

① 0-0，0-1，0-2 和 0-3 型对应于分散晶粒结构，其中 0-0 型连接与图 7-1(b)所示的分散晶粒结构等价，而 0-1，0-2 和 0-3 型对应于图 7-1(a) 和图 7-2(a)。

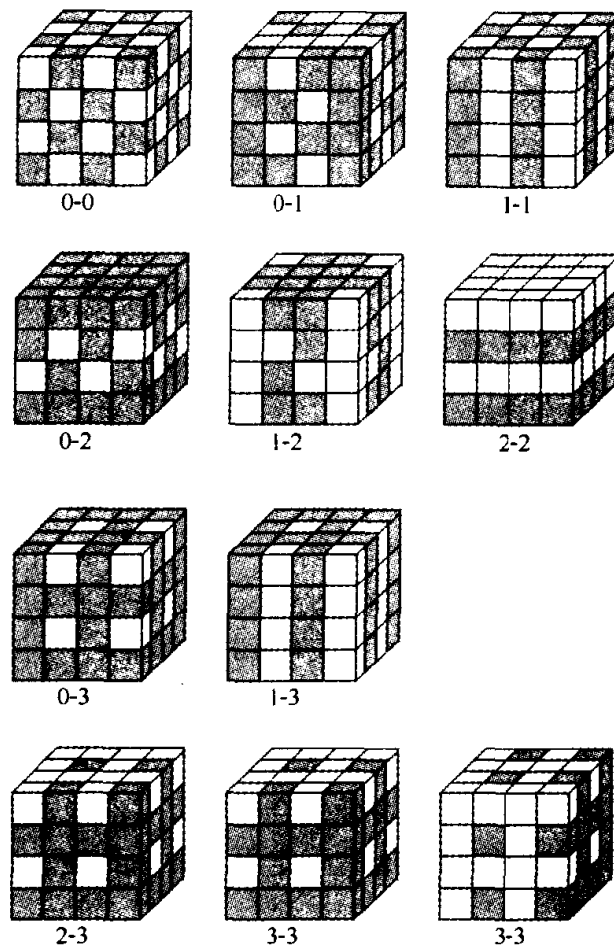


图 7-3 二相合金的十种晶粒连接方式。最后两个为 3-3 型连接的两个方向视图

(引自 Newnham R.E., Skinner D.P., and Cross L.E., Mater. Res. Bull. 1978, 13: 525)

② 1-2, 2-3 和 1-3 型对应于图 7-1(c)的聚集晶粒结构, 其中 1-3 型中弱相晶粒聚集成单链状; 而 1-2 和 2-3 型中弱相晶粒成密堆聚集。

③ 1-1 和 2-2 型是晶粒聚集结构的特殊情况, 其中弱相晶粒按一定的方向聚集而成片状结构。

④ 3-3 型对应与图 7-1(d)和图 7-2(c)的逾渗型晶粒团聚簇网络结构, 其中两相两相材料构成互穿透的三维网络, 即形成逾渗晶粒簇 (percolation cluster)。

从图 7-1~图 7-3 可以看到, 非均质材料的相分布是非常复杂的。搞清相分布的情况对建立微结构和物性的关联至关重要, 对于完全分散晶粒结构 [图 7-1(a) 和 (b)、图 7-2(a)、图 7-3 所示 0-0, 0-1, 0-2 和 0-3 型] 和片状聚集结构 (图 7-3 所示 1-1 和 2-2 型), 这一问题已经很好的获得解决。但是对于大多数晶粒聚集结构 [图 7-1(c) 和 (d)、图 7-2(b) 和 (c)、图 7-3 所示 1-2, 1-3, 2-3 和 3-3 型] 来说, 除非知道所有晶粒 (所有尺寸) 的 s -点晶

粒关联函数 (s -point grain correlation function), 否则解决这一问题几乎不可能。但对于高度有序的情况 (例如周期结构或准周期结构), s -点晶粒关联函数可以容易得到。实验测量的相分布对关联函数可以表示为 (pair-correlation function):

$$g_2(\mathbf{r}) = \frac{1}{N} \sum_{r_1} P_2(\mathbf{r}_1, \mathbf{r}_2) \delta(\mathbf{r}_1 - \mathbf{r}_2 - \mathbf{r}) \quad (7-4)$$

其中 $g_2(\mathbf{r})$ 是在位置 \mathbf{r} 处发现一个晶粒的几率。 $P_s(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_s)$ 是在同一晶粒团聚簇中发现 s 粒位于位置 $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_s$ 的几率。 $g_2(\mathbf{r})$ 可以通过光散射、小角中子散射和 X 衍射直接测量。一般而言, 我们对高阶关联函数知之甚少, 除非建立一些模型加以讨论。

7.1.3 有效媒质近似——二相合金微结构与物性关联问题处理实例

图 7-4 显示了一张复相合金的显微图片, 人们为材料在显微结构上如此复杂而感触甚深, 甚至于有些手足无措。过去, 物理学家习惯于处理一些简单的体系, 但凝聚态物质的深入研究需要和复杂体系打交道。传统的构筑模型来将问题简化在这里依然起了关键性的作用^[2]。

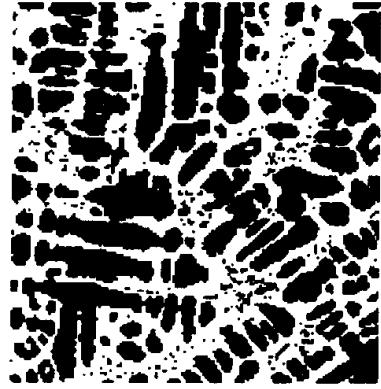


图 7-4 复相合金的显微照片

回到原来弱相基本是各向同性的情况 (见图 7-1)。如果知道了两相各自的基本物理参量, 那么这两相混合物的物理性质将如何呢? 19 世纪的物理大师麦克斯韦 (J.C. Maxwell) 给出了部分答案。我们来看他是怎样解决这一问题的。

假设两相的体积百分比分别为 c 和 $(1-c)$, 而它们的介电常数分别为 ϵ_1 和 ϵ_2 , 将弱相的一块区域理想化为一个球体嵌在具有有效介电常数为 ϵ_m 的均匀介质 (主相, 被称为有效媒质) 之中。如果再加一外场使有效媒质中的电场强度为 E_m , 根据静电学, 可求出嵌入球体的偶极矩与 $E_m(\epsilon_1 - \epsilon_m) / (\epsilon_1 + 2\epsilon_m)$ 成正比, 然后对原始结构的每一区域作类似的处理, 因而可以求出单位体积的总偶极矩 (即极化强度 P) 正比于:

$$P \propto c E_m \frac{\epsilon_1 - \epsilon_m}{\epsilon_1 + 2\epsilon_m} + (1-c) E_m \frac{\epsilon_2 - \epsilon_m}{\epsilon_2 + 2\epsilon_m} \quad (7-5)$$

这就等于设想将混合物球体嵌入在只有场强为 E_m 的有效媒质之中的结果。我们已经设想有效媒质应与嵌入的混合物球体在静电学上毫无区分, 即 $P=0$ 。根据这一条件, 即可求出有效媒质的介电常数 ϵ_m 。由于稳态电流场与静电场形式上完全相似 (图形上电流 i 线与电位移线 D 完全吻合), 所以电导率分别为

σ_1 及 σ_2 的两相媒质, 类似地可以求出有效媒质的电导率:

$$\sigma_m = \frac{1}{4} [(3c-2)\sigma_2 + (3c-1)\sigma_1] + \sqrt{[(3c-2)\sigma_2 + (3c-1)\sigma_1]^2 - 8\sigma_1\sigma_2} \quad (7-6)$$

如果涉及金属和绝缘体的混合物, 可令 $\sigma_2=0$, 则:

$$\sigma_m = \left(\frac{3}{2}c - 1 \right) \sigma_1, \quad c > \frac{1}{3} \quad (7-7)$$

结果表明 σ_m 与 c 成线性关系。若 $c \leq 1/3$, 则 $\sigma_m = 0$ 。这也是可以理解的, 导电相 (弱相) 如果成分太少, 将无法构成连通的区域来导电。这也就预示了下节要讨论的逾渗阈值问题。当然这里给出的只是有效媒质近似的粗糙的轮廓。在许多细节上, 后人已经作了不少改进, 这里不去细说。但这种近似虽然在一定范围给出了有用的结果, 也存在很多问题: 如果非均匀区域非常小, 两相的电导率差异相当大的情况下, 这一讨论就会失效。当然, 有关导通阈值附近的许多推论的有效性也是值得怀疑的。总之, 有效媒质近似是对复相材料的物性求平均的结果, 但在抹平之中, 难免将无序系统的某些特征也一同抹杀掉了, 这是方法本身的缺陷。

7.1.4 非均质介观结构实例

在物理学中, 将原子的尺度, 约为 $0.1\text{nm} (= 10^{-10}\text{m})$, 乃至更小的尺度称为微观尺度; 而将日常生活所接触的尺度, 约为 mm 到 m 乃至更大一些的, 称为宏观尺度; 而介于两者之间的还存在着所谓的介观尺度, 其尺寸约从微米到纳米量级 (图 7-5)。当前材料科学尤其是纳米材料科学的首要着眼点在于人工设计、制备和组装各类非均质材料, 通过人为引入 7.1.1 节所述的两类不均匀性而获得具有独特性质的材料。本节举出若干人工设计的介观周期、准周期和无序结构以期举一反三。

(1) 微米尺度结构 从不同的尺度来看, 同一材料具有不同的结构层次。以金属锻造材料为例 (如图 7-6 所示), 最低的层次起始于晶体结构单元, 其尺寸为 1nm 量级; 而在高一层次即介观层次, 位错为显著的结构特征, 它们以分立或纠结的形式存在, 其特征尺寸可以用位错间的平均距离 d ($\sim \rho^{-1/2}$, ρ 为位错密度) 来表征; 位错纠结可构成 $1\mu\text{m}$ 量级的位错“单元”; 在晶粒层次, 材料表现为晶粒的聚集体, 相邻的晶粒由晶界分开, 晶粒的典型尺寸为 $50\mu\text{m}$, 大的晶粒可达肉眼可见的 1mm 量级; 而锻造的原材料本身的尺寸可达 10cm 。可见对同一材料来说, 研究的层次不同, 材料将显示不同的结构特征。

如上所述, 通常的晶体由大量原子规则地排列成宏观晶体的, 在介观尺度上仅仅散布有一些位错或其他缺陷, 并不构成重要的结构层次。但也有

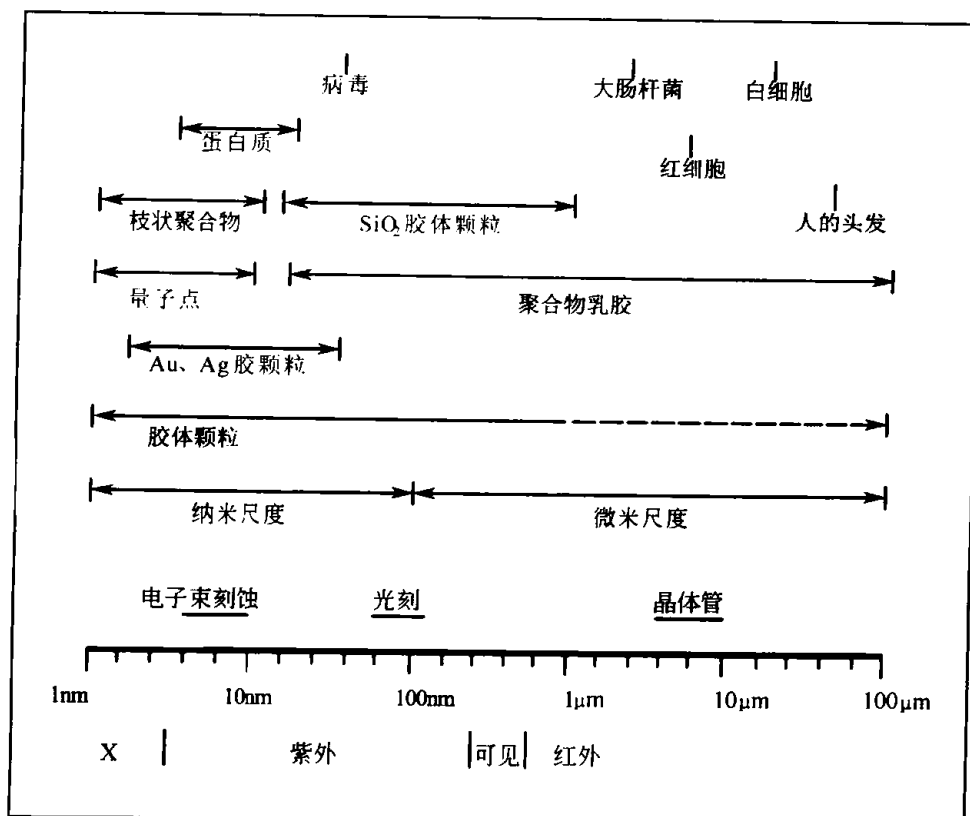


图 7-5 结构尺度示意图

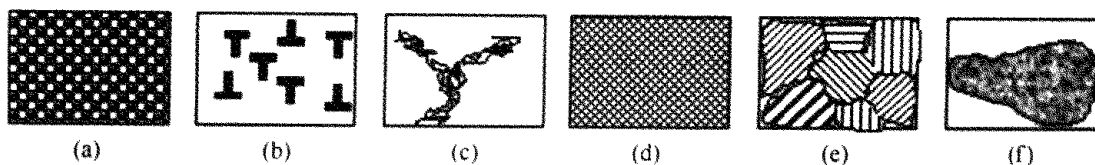


图 7-6 金属锻造材料的多层次结构特征

(a) 原子/晶体结构 (0.5nm); (b) 分立的位错 (10nm); (c) 位错纠结 (0.2 μm);

(d) 单个晶粒 (10 μm); (e) 多晶结构 (0.1mm); (f) 锻造原材料 (10cm)

(引自 Allen S.M. and Thomas E.L., The Structure of Materials, John Wiley & Sons, Inc., 1999)

例外，像自然界矿物中的蛋白石 (opal)。20 世纪 70 年代，澳大利亚学者对蛋白石进行了细致的研究，他们用扫描电子显微镜观察到蛋白石是 μm 量级的非晶态 SiO_2 小球堆积起来的周期结构 (图 7-7)，而这一结构的周期为 μm 量级，这一量级正好为可见光波波长的量级。因此可见光波可以对这一周期结构发生衍射，就像晶体结构对 X 射线衍射相类似。这种对可见光衍射效应的结果是其变彩效应的原因。白光是由各种波长的可见光混合而成的，当一束白光投射到相邻两个 SiO_2 球层的界面时，相邻界面对可见光的反射波之间会产生光程差，当白光中某一颜色的光的波长满足 Bragg 方程，

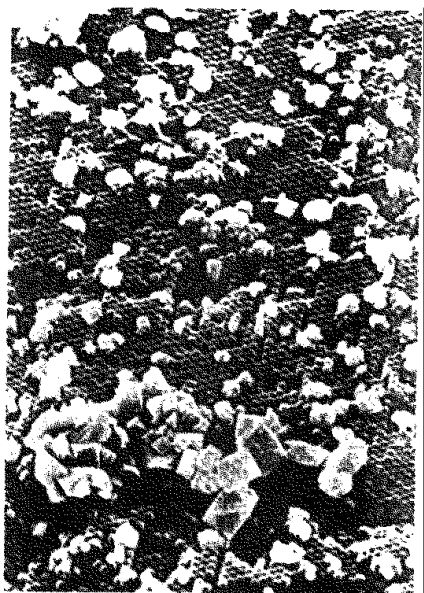


图 7-7 天然宝石蛋白石
结构 SEM 照片

(引自 Wenk H.R. et al., Electron
Microscopy in Mineralogy [M],
Berlin: Springer-Verlag, 1976)

$$2d\sin\theta = n\lambda \quad (7-8)$$

这里的 d 为相邻 SiO_2 球层的间距； θ 为入射光与界面的夹角； λ 为某一颜色可见光的波长； n 为整数，此种颜色的光波会衍射加强，而使宝石蛋白石呈现它的颜色。这是第一次从矿物研究中发现存在介观的周期结构。

由于宝石蛋白石具有极高的经济价值，矿物学家在发现了它的奥秘之后，便投入了人工制备蛋白石的研究，并取得了成就。图 7-8 为人工制备的蛋白石结构。这一结构材料可用于光子晶体。可以说人工蛋白石的合成工作为整个材料科学提供了一个新的方向。这一方向就是制备各种人工介观周期结构的材料。

从结构的角度来看，人工蛋白石是由单分散 SiO_2 胶体球自组装构成 3D 密堆结构 (fcc 结构) 的复相介电材料， SiO_2 相理论致密度为 74.0%，其中第二相为球间空隙(致密度为 26.0%)；因而可以通过改变 SiO_2 胶体球的直径而改变人工蛋白石的结构周期，从而实现介电常数的周期性调制。理论计算表明，要获得具有明显能隙的光子晶体，二相材料

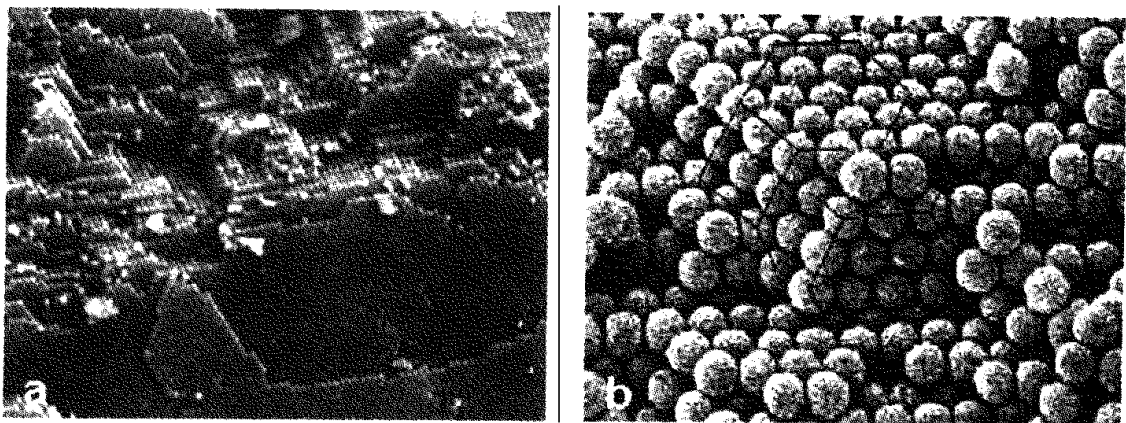


图 7-8 人工合成蛋白石晶体 SEM 照片：单分散 SiO_2
胶体球自组装构成 3D 密堆结构 (fcc 结构)

(a) 球体平均直径 440nm。可用做可见光波段的光子晶体 (引自 Lopez C., Blanco A., Miguez H., and Meseguer F., Opt. Mater. 1999, 13: 187); (b) 球体平均直径 300nm, 尺寸离散度为 5% (引自 Cheng B.Y. et al., Opt. Commu. 1999, 170: 41)

(如 SiO_2 和空气) 的折射率 ($n = \sqrt{\epsilon}$) 对比度即介电常数对比度要大。 SiO_2 材料的折射率还不够大, 因而可将人工蛋白石作为模板 (template) 用 C、Si、 TiO_2 来替换 SiO_2 , 而制得具有高介电常数比的反蛋白石型 (inverted opal structure) 介观周期结构材料 (图 7-9), 方法是将替代材料填充于蛋白石模板的 SiO_2 球间空隙, 然后溶解掉 SiO_2 球 (图 7-10)。显然这种反蛋白石型结构的材料理论致密度只有 26.0%, 因而进一步可将反蛋白石型材料作为结构模板, 把材料填充于它的球形空隙处而获得蛋白石型的材料。

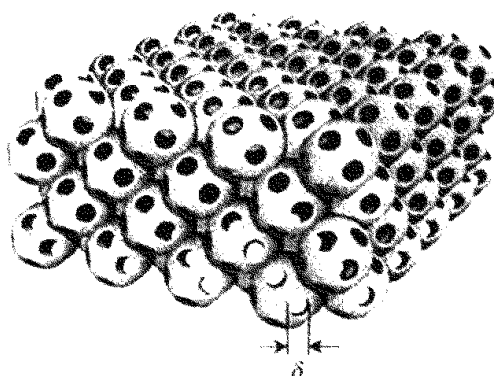


图 7-9 反蛋白石型材料

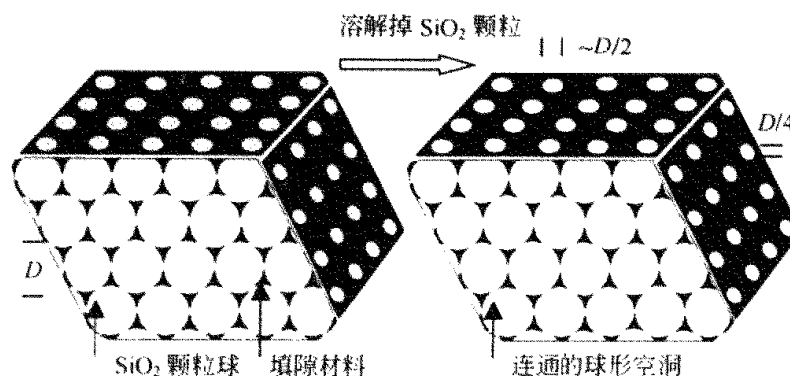


图 7-10 反蛋白石型结构模板制备示意图

(引自 Gates B., Yin Y.D., and Xia Y.N., Chem. Mater. 1999, 11: 2827)

理想的蛋白石或反型蛋白石是面心立方的介电超晶格, 类比于纳米尺度的晶体缺陷, 实际制备的这类介观周期结构也存在结构缺陷——所谓超晶格点缺陷、线缺陷和面缺陷。例如, 在蛋白石制备中引入少量更大或更小直

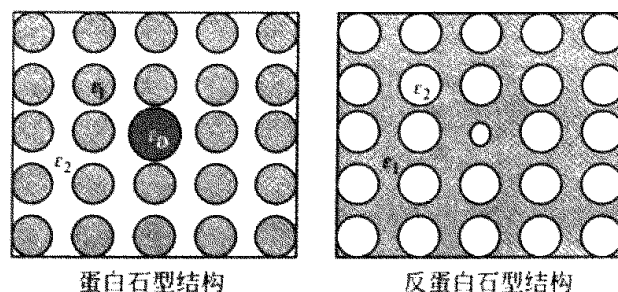


图 7-11 光子晶体介电超晶格中的点缺陷示意图

(引自 Joannopoulos J.D., Villeneuve P.R., and Fan S.H., Nature 1997, 386: 143)

径的颗粒球、或者引入异类材料颗粒，就可获得这类缺陷，如图 7-11 所示。不同于晶格缺陷的是，介观超晶格缺陷在光子晶体中有很重要的用途。

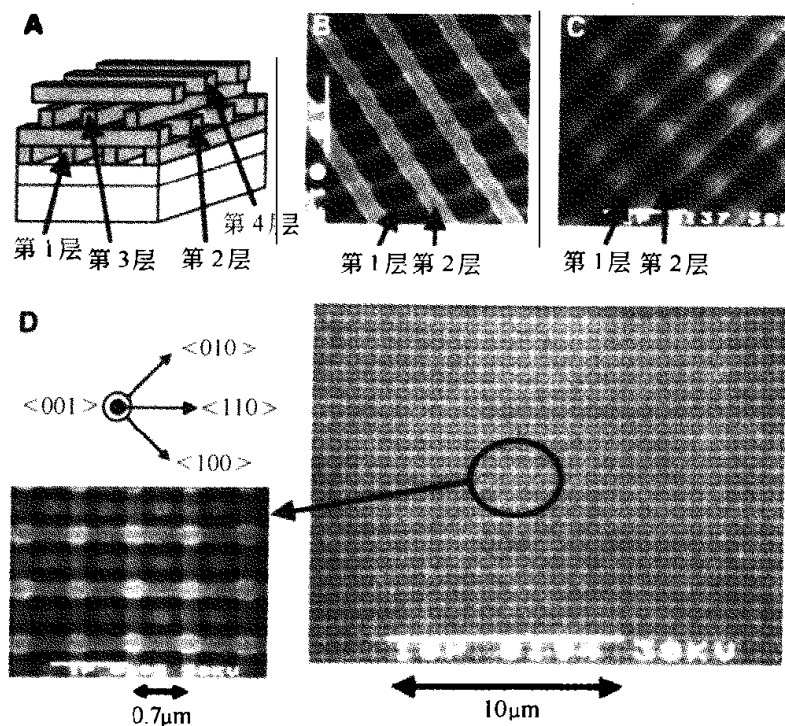


图 7-12 柴堆结构 (woodpile structure) 光子晶体，其中棒材为 GaAs 或 InP
(引自 Noda S., Tomoda K., Yamamoto N., and Chutinan A., Science 2000, 289: 604)

由于方法的局限性，目前通过分子自组装技术制得的 3D 介观周期结构类型有限，如只能获得密堆结构，且结构缔合单元多为球形。近期，用微米半导体线（图 7-12）和微米椭球颗粒作为结构基元制备介观周期结构的工

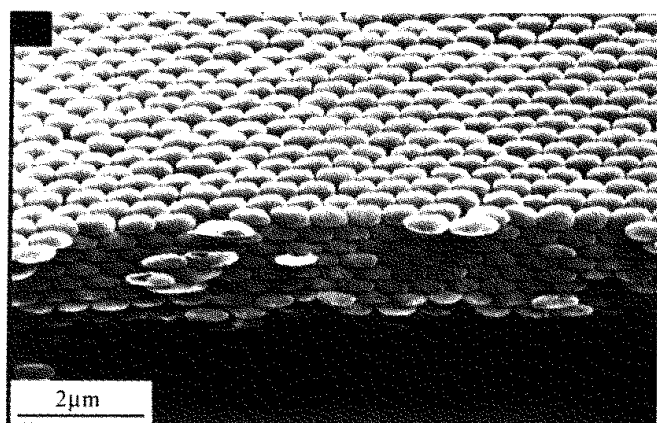


图 7-13 椭球形 TiO_2 颗粒（轴比 1.3、尺寸 300nm 量级）构成的光子晶体
(引自 Jiang P., Bertone J.F., and Colvin V.L., Science 2001, 291: 453)

作很有新意 (图 7-13)。

除上述分子自组装外, 还可以用刻蚀 (lithography)、薄膜外延生长和电化学阳极氧化 (electrochemical anodization) 等技术来制造介观周期、准周期结构材料。这类方法主要用以制造 2D 或 1D 介观结构, 结构的周期 (超晶格的“晶格常数”) 取决于材料的功能设计, 即与相应的声波或光波波长相匹配。

刻蚀方法在制备 2D 孔 (坑) 列阵和基于衬底的点列阵方面有其独到之处, 它可以控制结构基元 (柱型孔、坑和点) 的截面形状, 可实现所有类型的周期 (图 7-14) 或准周期排列 (图 7-15), 缺点是代价过高。

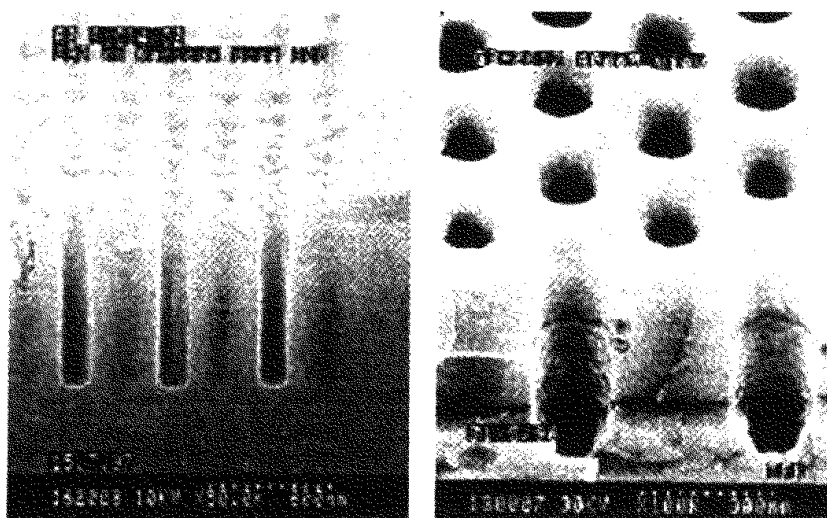


图 7-14 电子束刻蚀法制备的 2D GaAs/AlGaAs 孔列阵^[5]

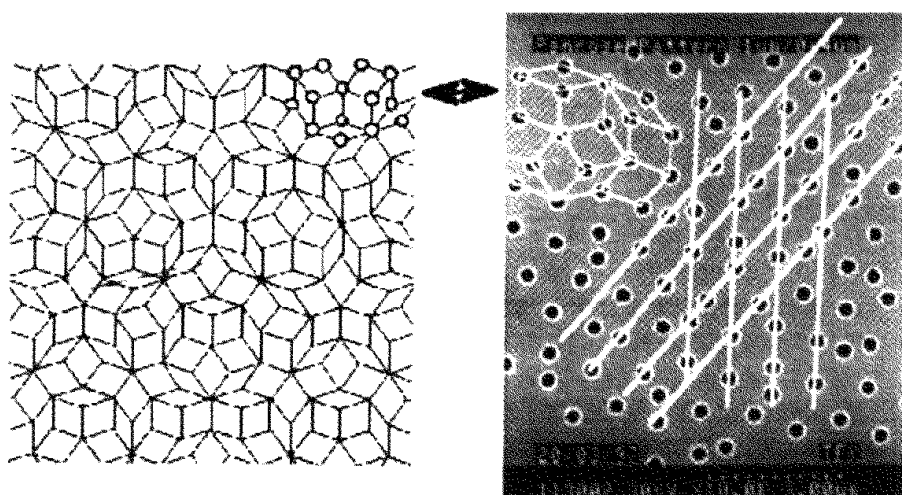


图 7-15 “光子准晶” (photonic quasicrystal): 2D 准周期超晶格结构
——圆柱形孔洞的 Penrose 拼砌^[5]

利用薄膜外延生长技术可以制备 1D 光学介电超晶格 (图 7-16), 通过控制每一层薄膜的厚度, 可以很容易地调制结构的周期性, 甚至可以制备具有分形特征的超晶格。例如结构周期满足 Fibonacci 或 Cantor 数列的超晶格^[5]。

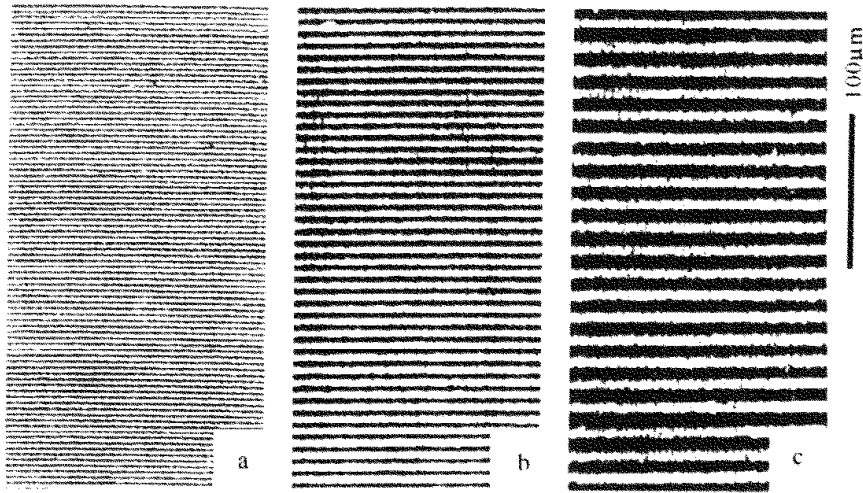


图 7-16 铌酸锂 (LiNbO_3 , LN) 1D 介电超晶格光学显微照片

其中超晶格周期分别为 (a) $2.7\mu\text{m}$; (b) $5.2\mu\text{m}$; (c) $15\mu\text{m}$ 。

这类结构材料具有非线性光学效应, 电-光效应及压电效应

[引自 Min N.B. (闵乃本), *Adv. Mater.* 1999, 11: 1079]

值得一提的是, 结合刻蚀和薄膜外延生长技术, 原则上可以制备各种形态的介观周期、准周期结构。例如棋盘结构 (图 7-17) 和前面所述的柴堆结构 (见图 7-12), 这类结构可能是实现真正意义上的 3D 光子晶体的一种有效途径 (Yablonovitch E., *Science* 2000, 289: 557)。

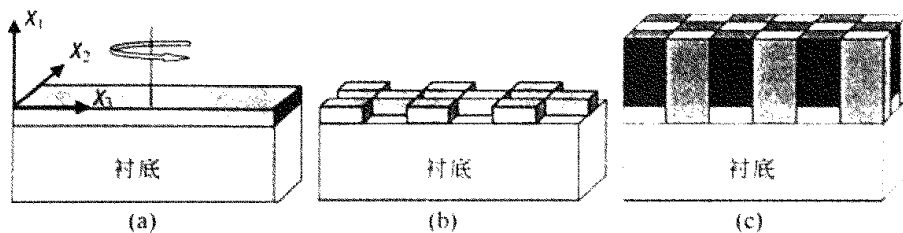


图 7-17 棋盘结构 (checkerboard structure) 光子晶体制备示意图

(a) 在衬底上外延生长一层 A 薄膜; (b) 把 A 薄膜刻蚀成所需结构花样;

(c) 在刻蚀的 A 材料模板上生长 B 材料, 方柱型 B 材料长在 A 材料柱间空隙处。

(参阅 Notomi M., Tamamura T., Kawashima T., and Kawakami S.,

Appl. Phys. Lett. 2000, 77: 4256)

从上面的讨论可以看出, 对于大多数微米尺度介观结构材料来说, 制备的关键在于首先要获得相应的结构模板, 其中以孔周期排列的多孔型模板 (porous template) 为主流。除了 3D 蛋白石型、反蛋白石型多孔模板, 电化学阳极氧化法制备的多孔氧化铝 (anodic porous alumina) 薄膜是一类重要的

2D 介观结构模板 (图 7-18)。多孔氧化铝模板中圆柱形孔的排列和蜂巢结构完全一样, 孔径 ($4\text{nm} \sim 0.5\mu\text{m}$) 由阳极电压控制, 孔的纵横比可达 1400。

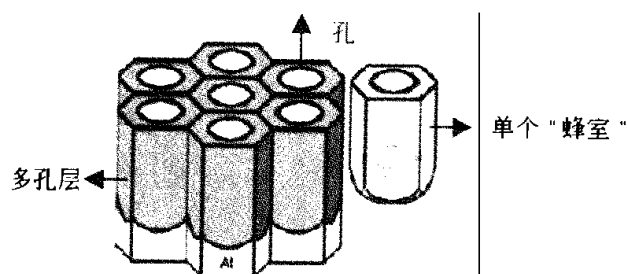
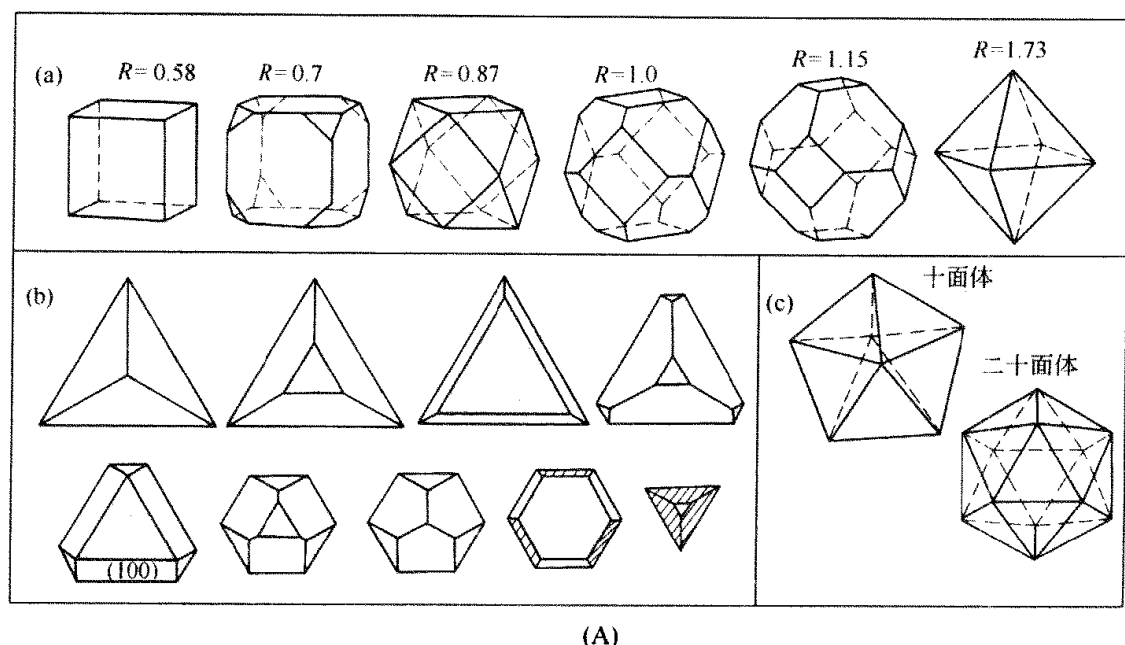


图 7-18 多孔氧化铝模板“蜂巢”结构示意图

(参阅 Diggle J. W., Downie T. C., and Goulding C. W., Chem. Rev., 1969, 69: 365)

(2) 纳米尺度结构 人工介观结构研究的另一重要方面是 nm (几 nm 或几十 nm) 尺度的结构, 这是微电子学技术向更小尺度的延伸。电子 de Broglie 波的波长在纳米以上量级, 因此在纳米尺度 ($1 \sim 100\text{nm}$) 结构中, 会呈现量子尺寸效应, 引起许多新的物理效应。特别是当代微电子技术已逐步逼近其物理极限, 取而代之的新技术必须是纳米技术。

纳米尺度显然要比蛋白石结构的尺度小得多, 但是同制备蛋白石结构一样, 化学自组装方法仍然是制备纳米尺度周期结构的有效途径之一。例如金属、半导体和氧化物型的自组装纳米晶超晶格结构 (self-assembling nanocrystal superlattices, 图 7-19~图 7-21), 它具有长程平移序甚至取向序。



(A)

图 7-19 多面体纳米晶粒

(a) 立方体及其衍生的多面体；(b) 四面体及其衍生的多面体；(c) 十面体与二十面体

[引自 Z. L. Wang (王中林), Adv. Mater. 1998, 10: 13]

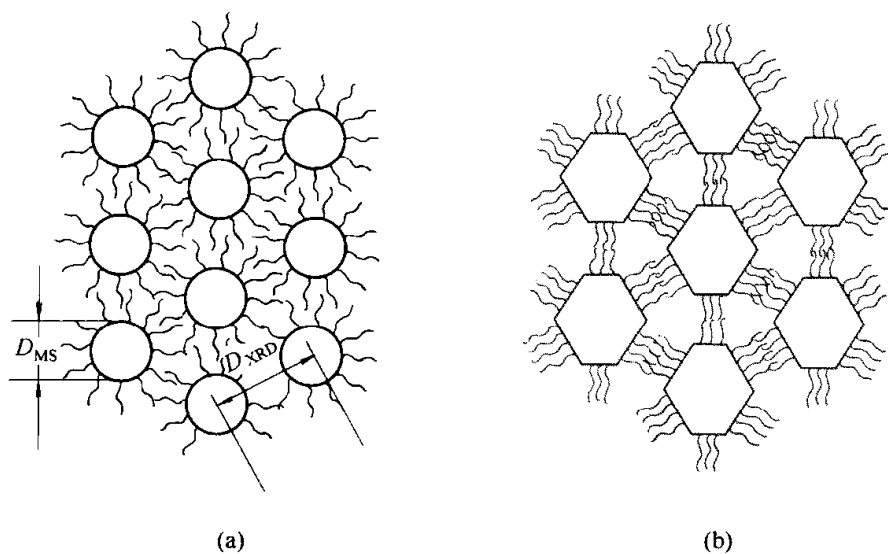
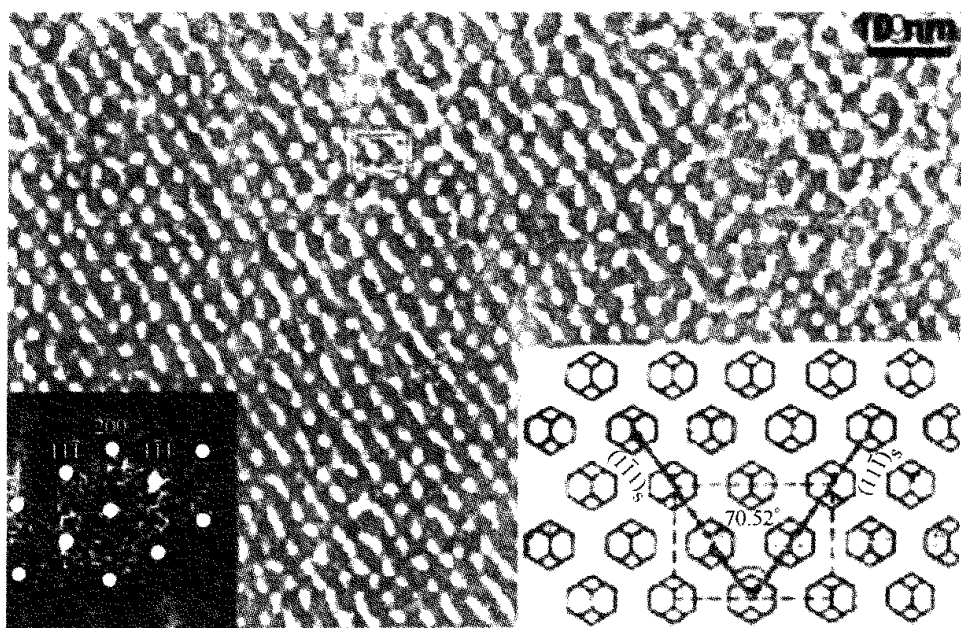


图 7-20 纳米晶粒自组装示意图

(a) 球形晶粒; (b) 多面体晶粒

[引自 Z.L.Wang (王中林), Adv.Mater.1998, 10: 13]

图 7-21 Au 纳米晶 3D 自组装 fcc 结构超晶格 $[011]_{NCS}$ TEM 照片

[引自 Z.L.Wang (王中林), Adv.Mater.1998, 10: 13]

其中自组装的结构单元为纳米尺寸的球形氧化物颗粒或具有规则对称外形的纳米正多面体晶粒，纳米颗粒间通过硫醇盐表面活性剂 (passivating thiolate surfacant) 耦合 (区别于原子间的化学键耦合)，颗粒间距 (即结构周期) 可调。纳米晶粒的外形对纳米超晶格的影响很大。

制备纳米尺度周期结构的另一重要方法是模板法。这类模板主要是多孔氧化铝 (anodic porous alumina, APA, 图 7-22)、有机玻璃 (PMMA, 聚甲基丙烯酸甲酯) 和聚苯乙烯 (PS) 模板。可以用电化学沉积 (electrochemical deposition) 等方法在孔中淀积不同的材料, 从而获得不同的 2D 超晶格材料。图 7-23 所示的是用多孔氧化铝为模板定向生长的 2D 纳米碳管阵列。

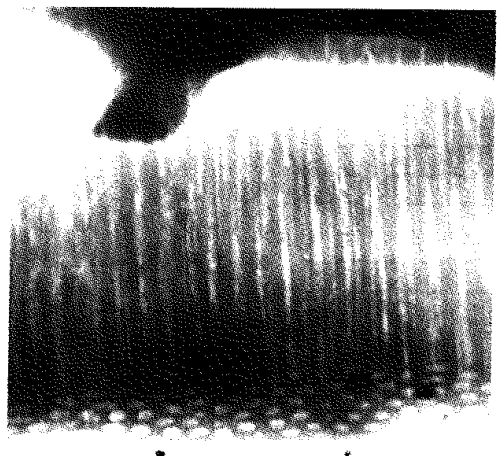


图 7-22 多孔氧化铝模板 TEM 照片, 其中孔径约为 15nm

[参阅 L. Pu, X. M. Bao, J. P. Zou, and D. Feng (濮林, 鲍希茂, 邹建平, 冯端),
Angew. Chem. Int. Ed. 2001, 40: 1490]

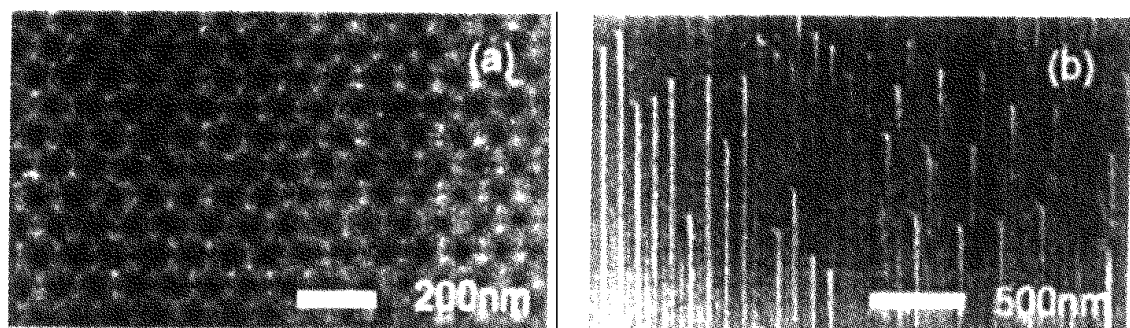


图 7-23 基于 APA 模板定向生长的纳米碳管阵列的 SEM 照片

(a) APA 模板; (b) 纳米碳管阵列

(引自 J. S. Suh, and J. S. Lee, Appl. Phys. Lett. 1999, 75: 2047)

7.2 结构的几何相变：逾渗

7.2.1 逾渗的概念

逾渗 (percolation) 的概念是由研究煤矿工人用的防毒面具而引入的。在这类面具中存在有多孔性的炭, 炭粒与孔隙分别相当于两相媒质中的两个相, 如果孔隙太小, 连通性不够, 气体就不可能透过, 这样, 就存在一个临界阈值, 超过它方能产生气体的渗透, 这现象就被称为逾渗。它和扩散是有

差别的，在扩散中无规性集中在流体之中，而在逾渗中，无规性却集中在固体骨架之中^[2,6]。

考虑图 7-24(a) 所示键逾渗问题，在通道网络中，黑线表示通道，虚线表示通道中流体的渗透。设每一通过被封锁的几率为 p ，我们希望了解流体在什么条件下可以穿过网络？基本条件显然就是存在无限长的能对穿的通道，出现这一通道的临界几率为 p_c ，即逾渗阈值。

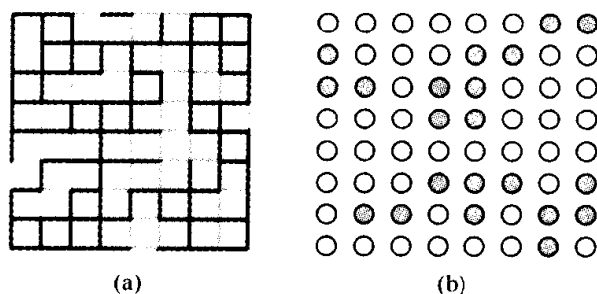


图 7-24 逾渗点阵的示意图^[6]

(a) 键逾渗；(b) 坐逾渗

也可以将无规性从点阵的键上转移到坐位上，图 7-23(b) 显示了白子和黑子构成的混合物，其占有几率分别为 q 与 p ($q + p = 1$)。同色的任意一对最近邻被认为是连通了，而黑子出现无限大集团的几率也被定义为逾渗阈值 p_c 。

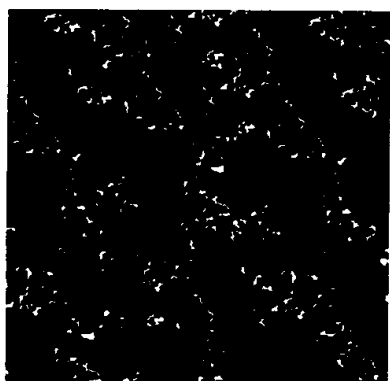


图 7-25 无限大集团的
计算机模拟图
(三角点阵坐逾渗计算
结果， $p_c = 0.5$)

渗流理论的基本任务在于计算每一坐位（或通道）属于无限大集团（或通道）的几率， P_∞ 作为 p 的函数，即 $P_\infty(p)$ ，算出 p_c 的具体数值。 $P_\infty(p)$ 的精确计算十分困难，通常要求近似解，而 p_c 却可以有一些精确的结果。图 7-25 给出了一些二维点阵逾渗阈值的计算结果。

逾渗模型建立之后，有关理论也大量展开。因而迫切需要有实验结果来对理论结果进行验证，寻求支撑。

B.J.Last 与 D.J.Thouless 于 1971 年首先利用导电纸上打了一系列小孔来模拟逾渗点阵，更加巧妙的是 B.P.Watson 与 P.L.Leath 的工作，是采用剪刀无规剪裁一块金属窗纱（大小为 1/4 英寸，包含 137×137 个网孔。1 英寸 = 2.54 厘米）来实现的。通过测量电导 σ 和切去的网点数（或切断的键数） p ，就可以测定逾渗阈值 p_c (见图 7-26)。

对于二维正方网格的键逾渗阈值为 $p_c = 0.5$ ，坐位逾渗阈值为 $p_c =$

0.59。电导 σ 与 p 成线性关系，但在 p_c 附近，实验就明确显示了和线性关系的偏离的临界区域，电导 σ 与 p 满足关系式：

$$\sigma = (p - p_c)^t \quad (7-9)$$

在 Watson 与 Leath 实验中测出 $t = 1.38$ 。在这一临界区域内的一系列物理性能（如电导、弹性模量、扩散系数等），又构成了需要理论来解释的问题。为了处理逾渗点阵上的扩散问题，法国物理学家 P.G. de Gennes，提出了迷宫中蚂蚁的有趣概念。设想将一个蚂蚁放在逾渗点阵上的某处，蚂蚁只能沿通道作无规行走，在 $p < p_c$ 时，无限长的通道已不存在，蚂蚁偏离原点位移 R 的均方值 $\langle R^2(T) \rangle$ （这里 T 为时间），将随 $T \rightarrow \infty$ ，两趋近一有限的饱和值；而在 $p > p_c$ 的情况下，存在有无限长的通道，则当 $T \rightarrow \infty$ ， $\langle R^2(T) \rangle \rightarrow \infty$ ，因而

$$\langle R^2(T) \rangle = ST = 2DT \quad (7-10)$$

这里的 D 为扩散系数， S 为步长。利用爱因斯坦关系，也可求出电导 σ 与 $\langle R^2(T) \rangle$ 之间的关系。从理论角度来看，逾渗实质上是一种几何的相变，逾渗阈值起了相变中临界温度所起的作用。逾渗阈值附近的行为类似于临界现象，可以借用临界现象的理论来计算一系列临界指数的数值。但是，从结构的角度看，逾渗转变并不仅仅是一种几何相变，这是因为材料的物性在逾渗阈值处有突变。这方面的具体问题读者可参阅相关专著。

7.2.2 逾渗的若干实例

有些分子单体溶在溶剂之中，聚合反应中有巨分子出现。若有 s 个单体聚合，即形成一个 S 聚体的大分子，当 $s \rightarrow \infty$ 时，即形成无限大的分子，被称为凝胶；当 s 为有限值时， S 聚体被称为溶胶。亚硅酸钠 (Na_2SiO_3) 溶于水时，形成“水玻璃”，就是溶胶-凝胶转变的一个实例。早在 1941 年（比逾渗模型提出要早 10 多年），有名的化学家 P.T. Flory 就提出了凝胶化的理论，他认为 S 聚体的键联接确定的分岔数 z 来进行，而且假设分叉总是树枝状的，不会形成闭合回路，数学上这被称为 Cayley 树。我们可以从图 7-27 中某一交叉点出发，到达相邻的分叉点，然后就有 $z - 1$ 个分支可供作延续的选择。如果通

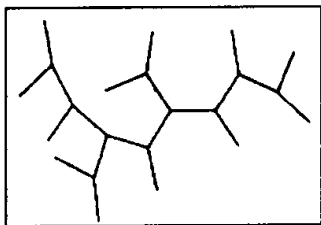


图 7-27 凝胶化的示意图
(Flory 论文中的原图)

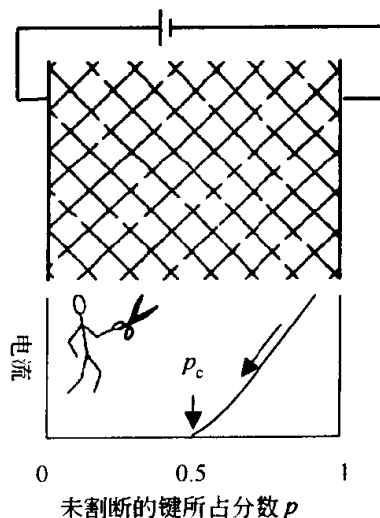


图 7-26 用剪刀无规切割金属网络验证逾渗现象^[6]

道的几率为 p ，那么只有满足 $p(z-1) > 1$ ，这些树枝方始有延续到无限远的可能性，而凝胶化的阈值为 $p_c = 1/(z-1)$ 。这实际上就是 Bethe 点阵（即无限大的 Cayley 树）的键逾渗阈值。这个事例说明采用简单直观的数学方法来解决，所得的结果往往可以和更普遍化的理论方法吻合。

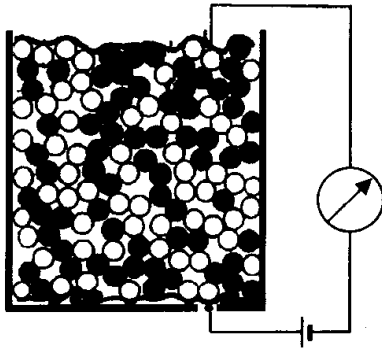


图 7-28 无规密堆两相球体的逾渗^[6]（黑球为导电的铝球，容器为玻璃）

图 7-28 所示的是一个位置无序结构中逾渗的实验，烧杯中放有塑料球和金属球的混合堆集。用铝箔在顶底两处作为电池和线路接通，通过改变金属球的百分数，观测电流的出现来确定逾渗的阈值。由于拓扑无规密堆结构不具有点阵的性质，其控制作用的临界阈值实质上是空间百分比（等于 p 乘上填充因子 ν ）的临界值 ϕ_c 。

进一步还可以将逾渗问题推广到连续统（continuum）无序的情况。图 7-29 画出一些轮廓线相当于等高线，白的区域对应于水面以上，阴影区域对应于水面之下。(a)、(b)、(c)分别对应于水库建成前后的情况：(a)相当于建水库之前，大地是连通的，有许多山，存在有分散的池塘；(b)相当于部分放水之后，水道业已贯通；(c)为水库建成之后，变成了千岛湖的图像，原来的陆地大部分被淹，只剩一些高的山峰，成为一个个孤立的岛屿。

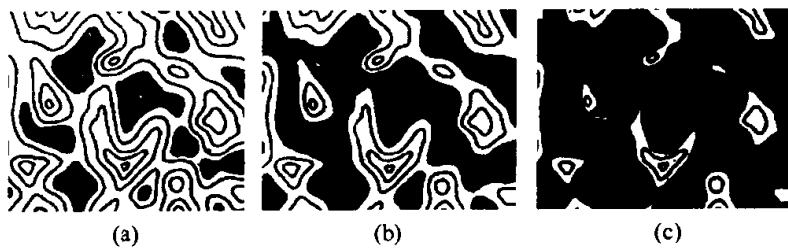


图 7-29 连续统的逾渗^[6]

逾渗阈值在(b)的附近，由于等高线所表示的高度分布是就高度的平均值而言，是统计性的分布，图(c)就像是图(a)所逆转，因而可以推断阈值处所对应面积 $\phi_c = 1/2$ 。这一模型可用以模拟粒子系统在无规势场中的行为，逾渗则对应于定域化到非定域化的转变^[6]。

7.3 无规结构：分形几何

7.3.1 支离破碎世界的描述

1883 年德国数学家 G.Cantor 设想将一根杠分为三段，将中段舍弃，再

对剩下的两段进行类似的处理，这一过程无穷尽地持续下去，其想像中的极限图形，被称为 Cantor 杆（图 7-30）。Cantor 之所以设想这种古怪的集合，无非是要举实例来说明存在一些传统数学方法难以处理的函数，一方面可无限分割，而另一方面又是不连续的^[2,3,7]。



图 7-30 Cantor 杆的构筑^[7]

Cantor 杆以外还有 Sierpinski 毯（图 7-31）。这些图形都具有自相似性；它们都具有错综复杂的精细结构，其几何性质很难用传统的几何术语来描述，因而获得了“病态图形”的称呼。

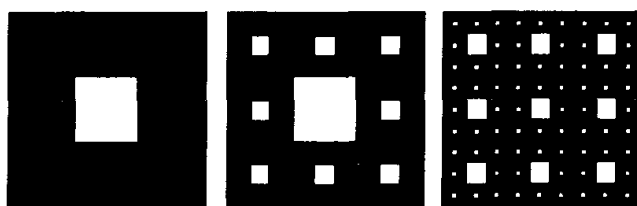


图 7-31 Sierpinski 毯的构筑^[7]

理解这些“病态图形”特殊含意的线索来自维数的考虑。空间或图像具有维数，其含义也来自数学家，按照 Euclid 理论，有一维的直线，二维的平面，三维的空间；推广出去，还可以有 $n (n > 3)$ 维的超空间；这是人所共知的常识。这里我们将 Euclid 的维数用 d 来表示。按照拓扑学，任何一个图形作橡皮式的形变而不撕裂的话，其拓扑性质保持不变。故而就引入了拓扑维数 d_T 。例如一条线，不管人们如何将它形变成千奇百怪的曲线或折线，其拓扑维数 $d_T = 1$ 保持不变；类似地，任何的曲面具有 $d_T = 2$ ；这也不难理解。尚有一种维数，是从自相似性出发来定义图形的维数：如果将任一图形放大 L 倍，对于 D 维的几何对象将能获得 K 个原来的几何对象，这样就有

$$K = L^D \quad (7-11)$$

对于正方形，当 $L = 3$ ， $K = 9$ ，即 $D = 2$ ；对于立方体，当 $L = 3$ ， $K = 27$ ，则 $D = 3$ 和直观的 Euclid 维数正好吻合。

1919 年数学家 F. Hausdorff 即以：

$$D = \frac{\ln K}{\ln L} \quad (7-12)$$

来定义几何对象的维数，被称为 Hausdorff 维数 D 。从另一观点来看，如 D 维的单元对象，分割为体积为 r^D 的 N 个小单元，就有 $r^D N = 1$ ，如此又可以得到和上式等效的维数表示式，即

$$D = \frac{\ln N}{\ln \frac{1}{r}} \quad (7-13)$$

D 并不限于整数。见图 7-29 所示 Cantor 杆，尺度增加 3 倍，单元数只为原来的 2 倍，这样

$$D = \frac{\ln 2}{\ln 3} \approx 0.6309 \quad (7-14)$$

而看图 7-30 所示 Sierpinski 毯，尺度增加 3 倍，单元数增加了 8 倍，即

$$D = \frac{\ln 8}{\ln 3} \approx 1.8928 \quad (7-15)$$

B. Mandelbrot 提出了分形 (fractal) 概念^[7]，据此，Hausdorff 维数 D 也常被称为分形维数。值得提醒一下的是，虽然具有分数 D 值的几何对象是分形，但也有一些整数 D 值的也具有分形特征，只需 $D > d_T$ ，而 D 一般不大于 d ，即 $D \leq d$ 。

分形结构既具有无序性，支离破碎，转折多端，具有明显的不连续性；但另一方面，它又是有序的，具有特殊的对称性，即自相似性，或标度不变性，对于缩小和放大，图像仍然保持不变。这种无穷嵌套的自相似结构，对我们来说并不陌生，如相变临界现象中出现的畴结构就具有这种特征；后来在聚合物的构型问题，逾渗阈值附近的无限大集团的结构问题，都看到满足自相似条件的几何图形。

7.3.2 物质结构中的分形

上节所述的分形几何学可用来分析若干自然界的现象。应该指出，由于自然界现象在空间上具有无规性；在时序上具有随机性，因此它们的自相似性（或标度不变性）只可能是统计性质的，而不可能是严格的。另外，自相似性也只可能在一定尺度范围内存在，必然存在有上、下截止半径：下截止半径肯定不会小于原子的尺寸，上截止半径肯定不会大于对象本身的尺寸。

我们可以从无规行走的 Brown 运动的粒子轨迹开始。Brown 运动粒子的轨迹虽然非常复杂，但仍然是一条线，所以其拓扑维数仍然为 $d_T = 1$ ，而其分形维数 $D = 2$ ，表明其轨迹最终将充满一个平面（见第 3 章图 3-19）。这一结果同样适用于聚合物熔体中的分子构型。至于聚合物稀溶液自回避行

走提供了一个更真实的模型。
当 $d = 3$, $R^{5/3} \propto N$, $D = 5/3$,
而 $d_T = 1$ (参阅第 3 章)。

同时, 分形几何也提供了描述多孔物质结构的有效方法, 凝胶结构的测定就是一个实例。我们来看硅气凝胶 (silica aerogel) 的情况 (图 7-32)。在这种凝胶中, 溶液为空气所取代, 形成孔隙百分比甚大的固体材料, 其玻璃态的 SiO_2 占 1%, 空气占 99%, 因而质量极轻, 密度低达 0.02g/cm^3 。就纳米尺度来看, 系由玻璃态 SiO_2 粒子所构成。从纳米到微米之间这一尺度范围内, 这些 SiO_2 粒子

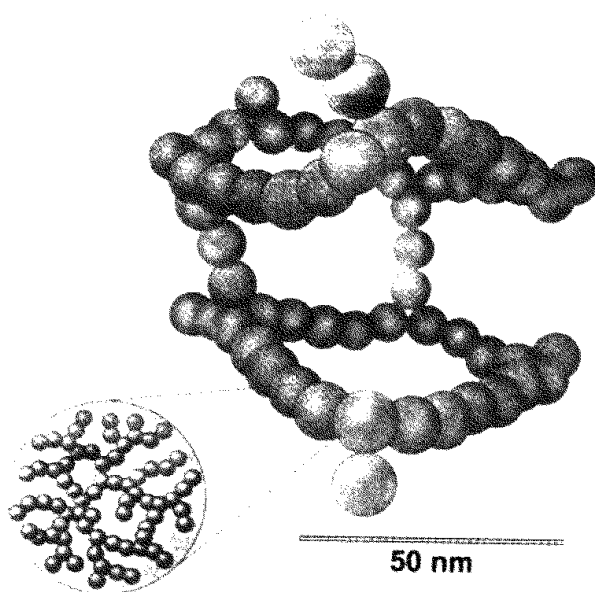


图 7-32 硅气凝胶分形结构示意图
(引自 Philip Ball. Made to Measure. Princeton: Princeton University Press, 1997)

汇集成多孔性的聚集体, 孔隙具有分形结构的特征。在微米尺度以上观察, 就显示为连续的固体。用 X 射线小角散射方法可以测定散射强度与波长的关系, 从而定其分形维数为 $D = 1.8$ 。

涉及物质的聚集生长过程的问题, 也时常发现分形的踪迹。例如铜从 CuSO_4 溶液中的电解沉积, 按照图 7-33 所示的装置, 可以观察到铜从点状阴极向外生长的分形特征。对这一过程进行计算机模拟, 可以看到粒子作无规的布朗运动, 一旦接触到中心或已经长出的部位, 随即粘住。这一模型被称为扩散限制聚集 (diffusion-limited aggregation, 简称 DLA) 模型。图 7-34 显示了二维 DLA 模型的结果, 分形维数为 $D = 1.71$ 。

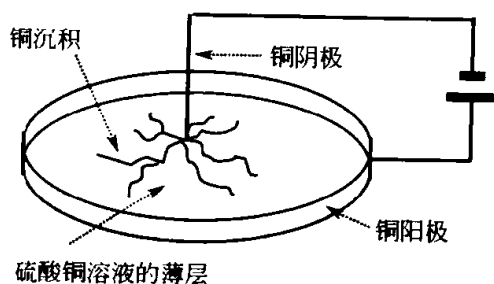


图 7-33 铜的电解沉积



图 7-34 DLA 模型的计算机模拟

参 考 文 献

- 1 Nan Ce-Wen (南策文). Physics of Inhomogeneous Inorganic Materials. Progress of Materials Science. 1993, 37: 1
- 2 冯端, 冯步云. 放眼晶态之外——漫谈凝聚态物质之二. 长沙: 湖南教育出版社, 1994
- 3 冯端, 金国钧. 凝聚态物理新论. 上海: 上海科学技术出版社, 1992
- 4 冯端, 金国钧. 凝聚态物理学中的基本概念. 物理学进展. 2000, 20: 1
- 5 Krauss T.F. and De La Rue R.M.. Photonic crystals in the optical regime——past, present and future. Progress in Quantum Electronics. 1999, 23: 51
- 6 Zallen, R., The Physics of Amorphous Solids. New York: John Wiley, & Sons, Inc., 1983. 中译本. 非晶态固体物理学. 黄昀等译, 北京: 北京大学出版社, 1988
- 7 Mandelbrot B. The Fractal Geometry of Nature., New York: Freeman, 1983

第 2 篇 材料的物性

第 8 章 理解物性的基本概念

撰稿人 韩 民

8.1 量子特征

8.1.1 波粒二象性

直觉上容易理解，电子是一种粒子，而光则是波。事实上，1897 年汤姆孙 (J.J. Thomson) 在实验中观察到阴极射线在电场和磁场中偏转，从而确定电子具有质量 m_e ($m_e = 9.1091 \times 10^{-31} \text{ kg}$) 和电荷 e 以后，电子就被确定为一种粒子了，因为它具有作为粒子的全部条件。而在经典光学以及麦克斯韦经典电磁理论中，光则总是被处理为电磁波，因为它具有波动的全部性质，如衍射、干涉和色散，光的颜色对应于波长 λ ，并且没有在牛顿力学中处理粒子运动所必须的质量。然而，根据 Hertz 在 1887 年所发现的光电效应以及 1905 年爱因斯坦对于光电效应的解释，人们发现光在某些方面也体现出粒子性。1927 年，科学家 C. Davission、L. Germer 和 G. P. Thomson 在实验中分别用单晶体反射与薄膜透射观测到电子衍射的图像，与 X 射线衍射花样完全类似，证实了布喇格条件在电子的实验中被满足。这些实验证明了电子具有波动的性质。

事实上，在此之前，1924 年，法国物理学家德布罗意 (Louis de Broglie) 基于宇宙的统一起源的信念，根据物理学定律的对应性的考虑，提出了电子的波粒二象性的假说：与一个动量为 p 、能量为 E 的粒子相联系的单色场的波长 λ 或频率 ν 通过普朗克常数相联系： $\lambda = \frac{h}{p}$ ， $\nu = \frac{E}{h}$ 。或引用波数 $k = \frac{2\pi}{\lambda}$ 和角频率 $\omega = 2\pi\nu$ ，并定义一个新的常数 $\hbar = \frac{h}{2\pi} = 1.0544 \times 10^{-34} \text{ J}\cdot\text{s}$ ，而将上述关系式写成更对称的形式： $p = \hbar k$ ， $E = \hbar \omega$ 。这样，电磁场和物质场，光与粒子的动力学性质（即动量和能量），获得了统一的定量表述。 $\lambda = \frac{h}{p}$ 因此被称为德布罗意波长。

根据德布罗意的波粒二象性关系, 可以估算与一个被电势 V 加速到动能 $E = eV$ 的电子的德布罗意波长为 $\lambda = \frac{h}{\sqrt{2m_e eV}} = \frac{1.23 \times 10^{-9}}{\sqrt{V} \text{ (V)}} \text{ (m)} = \sqrt{\frac{150.4 \text{ (eV)}}{E \text{ (eV)}}} \text{ (\AA)}$ 。加速到 $10^2 \sim 10^5 \text{ eV}$ 能量的电子的德布罗意波长为 $0.001 \sim 0.1 \text{ nm}$ ($1 \text{ nm} = 10 \text{ \AA}$) 与 X 射线的波长同数量级, 所以由电子衍射可以从晶体中得到与 X 射线衍射相同的花样。同样, 热中子 [在给定温度 (约 300 K) 与物质处于热平衡状态的中子] 的德布罗意波长约为 0.185 nm , 因此可以用中子对晶体进行衍射实验。与 X 衍射一样, 电子衍射和中子衍射都是研究晶体材料结构的最有力的工具。

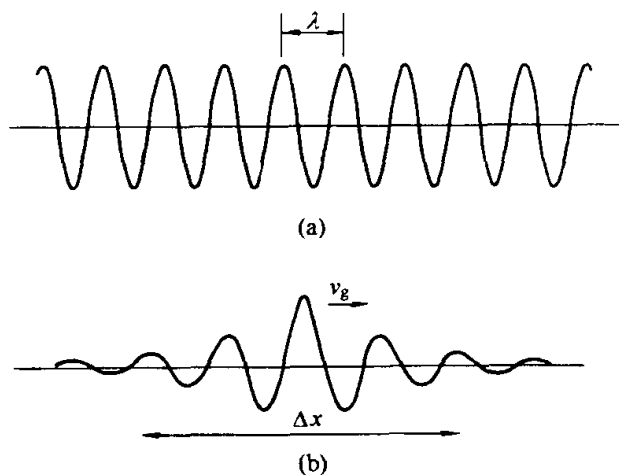


图 8-1 (a) 与一个未定域的粒子对应的连续波列; (b) 与定域在距离 Δx 之内的一个粒子对应的波包

利用德布罗意波粒二象性, 可用一个物质波 (德布罗意波) 来表示一个粒子。对于一个不受任何力作用的自由粒子, 具有确定的动量和能量, 对称性要求其德布罗意波振幅在整个空间都相同, 因此是具有恒定振幅的谐波 [图 8-1(a)]。而一个定域 (束缚) 于空间区域 Δx 内的粒子, 对应于一个如图 8-1(b) 所示的波包, 为了使波包在空间定域, 它必须由不同频率 (或波数) 的波叠加而成, 这样可使得波

包振幅或强度在区域 Δx 内很大, 而在此区域以外, 因不同频率的波的干涉过程而使得波包的振幅很快衰减。粒子的速度等于波包的传播速度, 即波包的群速度。因此, 如果波包扩展在 Δx 区域, 则组成此波包的干涉波的波数的值必须处于 Δk 范围内, 并且 $\Delta x \Delta k \sim 2\pi$, 或根据 $\Delta p = \hbar \Delta k$ 而改写成更一般的形式: $\Delta x \Delta p \sim h$ 。该关系式称为海森伯不确定原理 (Heisenberg's uncertainty principle)。它表明了自然界的一个基本事实: 一个粒子的位置和动量两者是不可能同时准确测知的。 Δx 称为粒子位置的不确定度, Δp 称为粒子动量的不确定度。 Δx 越大, 则 Δp 越小, 反之亦然。这就是说, 关于粒子的空间位置的信息越准确, 关于其动量的信息就越不准确。所以必须用一个延伸到整个空间的恒定振幅的谐波 ($\Delta x \sim \infty$) 来表示一个具有精确动量 ($\Delta p = 0$) 的自由粒子 [图 8-1(a)]。

不确定原理还可以表示成另一种形式。设粒子的能量的不确定度为 ΔE ，时间测量的不确定度为 Δt ，则 $\Delta E \Delta t \sim h$ 。例如，如果原子中的一个电子处于受激态，则它将以一定的几率发生辐射跃迁而进入能量较低的另一定态。电子处于定态的平均时间即为此定态的寿命，与其跃迁几率成反比，具有不确定度 Δt ，由不确定原理，定态的能量具有不确定度 ΔE ，称为该定态的能量宽度。可以认为 Δt 与受激态的寿命同数量级，因此受激态的寿命越短，其能量宽度就越大。由于能量宽度的存在，使得在原子跃迁中即使存在动能守恒所引起的反冲能量也可以产生共振吸收。

常规情况下，有两类决定材料物性的主导因素：一类是原子系统，通常作为经典粒子处理，反映了位置序或粒子序(性)的效应；另一类是电子系统，通常表现出明显的量子力学特征，反映了动量序或德布罗意波序(性)的效应。究竟哪一种序占主导地位，需要看体系的特征尺度或者粒子的平均间距 a 是否达到德布罗意波长 λ 的量级，如果 $a > \lambda$ ，则粒子序为主。德布罗意波长是温度的函数，如果令粒子的动能 $mv^2/2$ 等于平均热能 $3k_B T/2$ ，则等效地可以通过温度的界限来确定是粒子序亦或是波序占主导作用，该温度界限可用简并温度 T_0 来表征，并粗略地估算为： $T_0 \equiv \frac{h^2}{3mk_B a^2}$ ，如果 $T > T_0$ ，则粒子序占主导地位，反之，如果 $T \ll T_0$ ，则波序占主导地位。例如，取特征尺度为原子间的平均间距，在凝聚相其典型值为 \AA 的量级，由于电子的质量很小， m_e 约 10^{-27}g ，故简并温度 T_0 约高达 10^4K 。一般情况下， $T \ll T_0$ ，故电子系统的行为突出地反映了波序的效应。而原子系统，由于原子的质量为电子质量的 10^3 倍以上， T_0 约 $20 \text{K}/\text{\AA}$ ，一般的物质，在远高于 T_0 的温度就已固化，因而就凝聚物质的原子排列而言，粒子序应占主导地位，总可做经典的处理。轻元素氦则是一个例外。氦的简并温度 $T_0 \approx 5 \text{K}$ ，由于氦原子间很弱的范德瓦尔斯力，它们具有很低的液化温度 ($^4\text{He} \approx 4.2 \text{K}$, $^3\text{He} \approx 3.2 \text{K}$)，并能保持液态到绝对零度。所以氦液体表现出超流动性，这是原子系统的量子行为。而若干氦(^3He)原子构成的集团(团簇, cluster)，其稳定结构和能态也不是其他重惰性元素团簇由位置序决定的刚性球堆积壳层结构所能解释的，而体现出 ^3He 对单粒子本征态的满壳层占据，与超流动性一样，是典型的量子力学现象。

8.1.2 经典统计与量子统计

宏观上，材料是由大量粒子所组成的。在考察材料的宏观物理性质时，通常涉及的是大量粒子所组成系统的统计平均。考察由 N 个粒子所组成的孤立体系，每一个粒子可以一定几率处于能量为 E_1, E_2, E_3, \dots 的态。在一个特定的时刻，粒子分布在不同的态上，有 n_1 个粒子在能量为 E_1 的

态, n_2 个粒子在能量为 E_2 的态, 等等。由于粒子的相互作用, 粒子在资用能级上的分布是变化着的, 或者说表示 N 个粒子在资用能态之间的配分(分布)的数目 n_1, n_2, n_3, \dots 是变化着的。但对于系统的每一个宏观态, 总有一个比其他任何配分都更为有利的配分, 或者说, 给定系统的物理条件(粒子数、总能量), 就有一最可几的配分, 达到这个配分时, 就说这个系统处于统计平衡。通常获得的材料宏观的物理性质, 就是在统计平衡下的统计平均值。

有三种统计法。对于粒子性占主导的经典系统, 通常采用麦克斯韦-玻耳兹曼分布律, 构成经典统计。对于波动性占主导的系统, 需要采用量子统计。经典系统是由全同的但可区别的粒子所组成的系统。所谓全同粒子, 是指这些粒子具有相同的结构和组成; 所谓可区别, 是指在经典系统中每一个粒子在原则上确定的轨迹可以跟踪。例如处于晶格上的原子, 是可以确切地加以区分的, 是适合于用经典统计来处理问题的。在量子统计中, 粒子是全同的并且是不可区分的。所谓不可区分是指只能区分每一个能级上有多少粒子, 但不能区分是哪几个粒子。如果这些粒子遵从泡利不相容原理, 因而不能有两个粒子处于同一量子态 E_i (单粒子态, 其配分数或占有数 $n_i = 0$ 或 1), 系统的波函数必然是反对称的(交换两个粒子, 波函数反号), 满足所有这些要求的粒子称为费米子。费米子所遵从的统计法为费米-狄拉克统计法; 如果这些粒子不受泡利不相容原理的约束, 因此系统对于能够处于相同量子态 E_i 的粒子数目没有限制(占有数 $n_i = 0, 1, 2, \dots$), 描写粒子系统的波函数必然是对称的(交换两个粒子, 波函数不变), 满足这些要求的粒子, 为玻色子。玻色子遵从玻色-爱因斯坦统计。

在费米子系统, 单粒子态 E_i 的平均占有数为费米-狄拉克分布:

$$\bar{n}_i = \frac{1}{e^{(E_i - E_F)/k_B T} + 1} \quad (8-1)$$

或对于单粒子能级分布非常稠密的情形采用单粒子态密度的概念, 态密度分布遵从费米分布函数:

$$f(E) = \frac{1}{e^{(E - E_F)/k_B T} + 1} \quad (8-2)$$

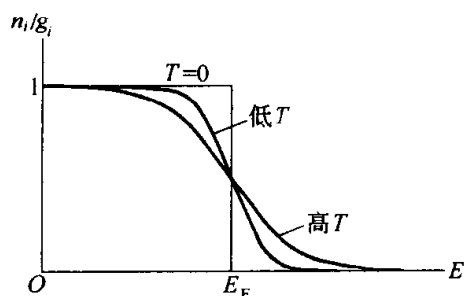


图 8-2 费米-狄拉克统计中三个典型温度的分布函数

式中的 E_F 为化学势, 常称为费米能(费米势, 费米能级), 在材料物理学中起着非常重要的作用。图 8-2 示出了三种不同温度下的平均占有数。当 $T = 0$ 时, 直到 $E = E_F$ 的所有能态全部被占有($n_i = 1$), 而对于所有 $E > E_F$ 的态全部空着($n_i =$

0)。因为在费米-狄拉克统计中，泡利不相容原理阻止粒子积聚在基能级上，因此零温度时，粒子占有最下面的一些资源能级，直到费米能 E_F ， E_F 表示该系统中费米子的最大能量。在较高温度时，较低能量的态通过吸收能量从而转移到（占有）能量大于 E_F 的态，但对于 $k_B T \ll E_F$ 的温度，因为所有低能量的态全部被占有，不相容原理不允许再把一些费米子加到这些态中去，因此只有自由能量接近 E_F 的那些费米子能够通过吸收较小的能量 $k_B T$ 而移至费米能以上的未被占有的态中，而处于距费米面较远的低能态的费米子则不能通过与周围环境交换能量而在费米面下能态间移动，除非它获得足够的能量（如吸收光子）而跃迁到费米面以上。图 8-3 形象地给出了这种过程。满足条件 $k_B \Theta_F = E_F$ 的温度 Θ_F 为费米温度。与费米能对应的波数 $k_F = \frac{\sqrt{2mE_F}}{\hbar}$ 为费米波数。

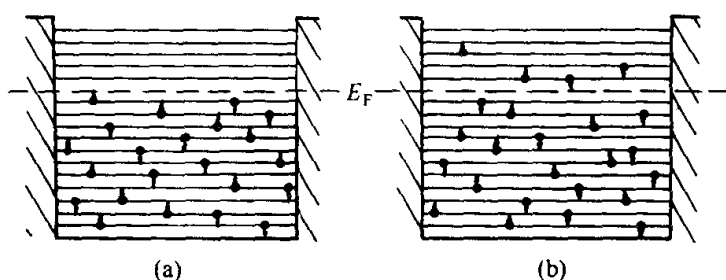


图 8-3 费米-狄拉克统计
(a) $T=0$; (b) $T>0$ 的有限温度

对玻色子系统，单粒子态 E_i 的平均占有数为玻色-爱因斯坦分布：

$$\bar{n}_i = \frac{1}{e^{(E_i - \mu)/k_B T} - 1} \quad (8-3)$$

或对于单粒子能级分布非常稠密的情形采用单粒子态密度的概念，态密度分布遵从玻色分布函数：

$$f(E) = \frac{1}{e^{(E - \mu)/k_B T} - 1} \quad (8-4)$$

式中的 μ 为化学势。玻色分布函数如图 8-4 所示。当 $E \rightarrow \mu$ 时， $f(E) \rightarrow \infty$ 。对于任何 E ，总应有 $\mu < E$ ，否则有些占据数成为负数而没有意义。所以如果单粒子能量的最低值选为 0，必有 $\mu \leq 0$ 。

引入一个参数 a ，可以将费米-狄拉克分

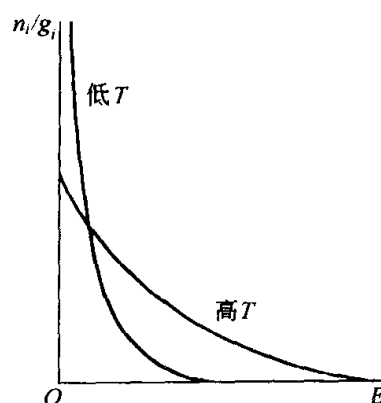


图 8-4 两种不同温度下玻色-爱因斯坦统计的分布函数

布、玻色-爱因斯坦分布和经典统计的玻耳兹曼分布统一写为：

$$\bar{n}_i = \frac{1}{e^{(E_i - \mu)/k_B T} + a} \quad (8-5)$$

其中， $a = \begin{cases} -1, & \text{玻色-爱因斯坦统计} \\ +1, & \text{费米-狄拉克统计} \\ 0, & \text{玻耳兹曼统计} \end{cases}$

当满足条件 $|\mu| \gg kT$ 时， $\exp(E - \mu)/k_B T \gg 1$ ，系统是非简并的，费米统计、玻色-爱因斯坦统计与玻耳兹曼统计在形式上是没有区别的，即量子统计可由经典的玻耳兹曼统计代替。

实验和理论表明，所有自旋为 $1/2$ 的粒子（电子、质子、中子、中微子、 μ 粒子等）都是费米子，而所有整数自旋的粒子（光子、介子等）都属于玻色子。容易理解，任意数量玻色子或偶数费米子组成的复合粒子（其自旋为整数）的行为如同玻色子，而奇数费米子组成的复合粒子（其自旋为半整数）则是费米子。例如，氢（ ^1H ）、 ^4He 、 ^7Li 原子等都是玻色子，而氘（ ^2H ）、 ^3He 、 ^6Li 原子都是费米子。

光子是一类重要的玻色子，在研究能级间的自发辐射跃迁和受激辐射跃迁中，玻色-爱因斯坦统计构成了光子在各能级统计分布的基础。对俘获在空腔中，与腔壁原子处于热平衡的电磁辐射——黑体辐射的分析，就是玻色-爱因斯坦统计应用于光子系统的著名例子。玻色把黑体辐射当作一种光子气体处理，研究系统中光子在各种能级上的分布。爱因斯坦断定，将各个光子分布于各个能级的过程中隐含的基本事实就是，光子是不可分辨的，即把系统看成是全同的且不可分辨的量子——玻色子，由玻色-爱因斯坦统计可以推导出黑体辐射的普朗克公式。由于光子在各能级的分布，对于每一频率，都有一个对应的黑体辐射强度，只取决于腔壁的温度。由于光子的不可分辨性，在同一能态上的粒子数只受统计分布所制约，在某能态上光子数可以无限的增加，并发生光子在动量空间的凝聚。例如通过受激发射，可以形成激光，其特征为大量光子占有同一状态，即具有相同的能量和动量，使激光束具有良好的相干性、单色性、方向性和高亮度，在现代科学技术中获得了广泛的应用。

粒子在低能级的凝聚是玻色子系统的重要性质。当化学势增大并趋向于最低能级（设为零）时，零能级的分布将趋于无穷，意味着粒子在最低能级的凝聚，即所谓的玻色-爱因斯坦凝聚（Bose-Einstein condensation, BEC）现象。定义一临界温度 T_c ，使得当 $T = T_c$ 时占据零能级以上各能级的粒子数 $N(E > 0) = \sum_{i \neq 0} \bar{n}_i$ ，等于总的粒子数 N_0 。当 $T < T_c$ 时，分布于激发态上的粒子数将减少，而将多余的玻色子全部推到基态上去，处于基态的粒子数为：

$$N(0) \approx N_0 \left\{ 1 - \left(\frac{T}{T_c} \right)^{3/2} \right\} \quad (8-6)$$

这些占据基态的粒子数将随着温度的降低而急剧增加，并最终使几乎所有的粒子都处于基态，这种现象即为玻色-爱因斯坦凝聚（BEC）。

BEC 是一纯的量子统计现象，发生于无相互作用的理想玻色气体或弱相互作用的均匀玻色气体中。大部分的实际的玻色子系统，总是或多或少地存在相互作用的。长期以来，科学家认为即使非常弱的相互作用，也总是导致粒子在发生 BEC 之前产生有序化凝聚成为固体，而非稀薄原子气体。到 1995 年，随着磁禁闭、激光冷却等技术的发展，E.Cornell 和 C.Wieman^[8] 等研究组分别在 ^{87}Rb 、 ^7Li 、 ^{23}Na 等碱金属及自旋极化 ^1H 的稀薄原子气体中观察到了 BEC。此后不到两年的时间里，凝聚于量子态的碱金属原子数增加了 3~4 个量级，从而给出了 BEC 的实在证据。为在稀薄原子气体中产生 BEC，需要将原子冷却并压缩于势阱中以使其德布罗意波长达到原子间距的量级，并使原子与容器壁没有任何热接触。所以 BEC 总是在超高真空中进行。首先将原子通过激光冷却的方法预冷到 1K 以下，然后用磁场或激光将原子在势阱中约束到数秒以上的时间，并在此时间内通过蒸发冷却的方法将原子冷却到 $\mu\text{K} \sim \text{nK}$ 的温度。由于理论和应用上的重要性，目前 BEC 的实验研究正方兴未艾。值得注意，W.Ketterle 等人从 BEC 的 ^{23}Na 气体中引出了具有同一动量态的原子束，即激射的原子束，亦被称为原子激光，具有一定的应用潜力，特别是精确的材料制备。

但相互作用的加强并不一定导致 BEC 的消失，液态氦就是一个显著的例子。由于氦具有轻的原子量和大的零点能，使之直到绝对零度都不会固化。因此历史上 BEC 最早在液态 ^4He 的超流现象上获得研究。氦在 4.22K 时液化，形成液态 He I 相，液态氦在 2.19K 时发生二级相变，并呈现反常的比热容，低于该相变温度的相叫做液态 He II 相，它是一种异于常规液体的特殊的液体，呈现超流状态（如，在狭缝间流过的液态氦的流速与压差无关，液体内部和液体与器壁之间无摩擦力）。实际上，液态氦是一种量子液体。1938 年，伦敦提出，液态 ^4He 在 2.19K 温度上发生的由正常液氦（He I）到超流氦（He II）的相变，可能就是在液氦中发生的 BEC 的一种表现形式。由此理论上可得到发生 BEC 的 T_c 约为 3.13K。按照伦敦的看法，处于零熵态（ $E=0$ ）的粒子，可能就是液氦的超流部分。

另外，由费米子构成的系统，不能直接实现 BEC 凝聚。但是两个费米子可以配对方式形成准玻色子，这些准玻色子也可以产生 BEC 现象，这就是金属合金乃至氧化物出现超导现象。因而 BEC 也是理解超导和超流现象的一把钥匙。

8.2 固态电子理论

8.2.1 金属自由电子理论

金属的电子结构属于金属物理的最基本问题之一。在简单金属的情形, 建立在近自由电子模型基础上的理论取得了和实验结果大致相符的结果。零级近似下, 简单金属的电子结构可采用特鲁德-索末菲的自由电子模型。根据这个极度简化的模型, 价电子完全公有化, 构成了金属中导电的自由电子, 离子芯与价电子的相互作用被完全忽略, 并且自由电子体系被视作电子间毫无相互作用的理想气体(电子气)。为保持金属的电中性, 可以设想将离子芯的正电荷散布于整个体积中, 恰好与自由电子的负电荷中和。这个模型被称为浆汁(jellium)模型。

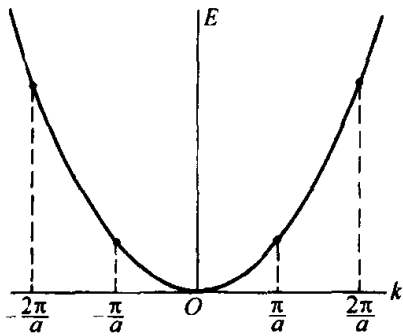


图 8-5 自由电子气的
抛物线型能带曲线

自由电子可视为波矢为 $k(k_x, k_y, k_z)$ 的平面波 $\Psi_k(r) \sim \exp(ik \cdot r)$, 如果金属样品的体积为 $V = L^3$, L 为样品边长, 则该金属样品可被看作一个势阱, 在势阱内部价电子可以自由运动, 该模型类似于量子力学中的方势阱中的定态的解, 根据波恩的周期性边界条件, 波矢 k 的诸分量只能为 $2\pi/L$ 的整数倍, 而由德布罗意波粒二象性, 电子的

动能与波矢之间有关系: $E = \frac{h^2}{8\pi^2 m} k^2$, 所以

电子可能占有的能态是量子化的。如果以

(k_x, k_y, k_z) 为坐标轴, 构成 k 空间, 则 k 在 k 空间作均匀分布, 而电子在 k 空间则呈球形分布, 等能面是以原点为球心的球面。图 8-5 绘出了电子能量与波矢的对应关系, 可见在 k 空间, 金属自由电子气模型具有抛物线型能带曲线。

金属中的电子是最典型的费米子, 考虑到电子的自旋, 允许每一个能级(允许的 k 态)能够容纳两个电子(一个自旋向上, 一个自旋向下)。因此由波矢到能量的转化关系, 可以得到态密度的表达式:

$$D(E) = \frac{V}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} E^{1/2} \quad (8-7)$$

此即抛物线型的能态密度曲线, 见图 8-6。

根据费米统计, 绝对零度下, 金属处于基态, 所有的电子占有不相容原理所允许的、最低的可能能级, 从 $k=0$ 的最低态开始, 从低到高, 依次填充。如果体积 V 中电子的总数为 N , 小于费米能级的总数, 则电子占有 $N/2$

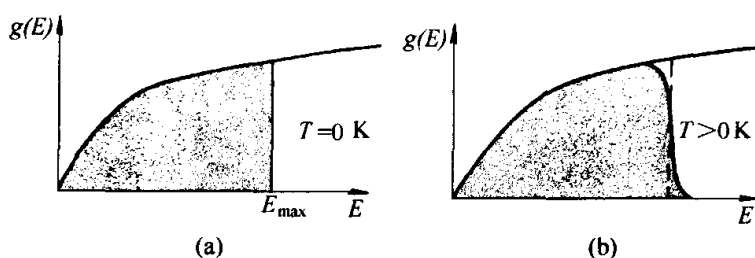


图 8-6 自由电子气模型的抛物线型能态密度曲线

(a) $T=0\text{K}$; (b) $T>0\text{K}$

个能量最低能态，而这些电子所占有的最高能级即为费米能 E_F ，根据以上的讨论，容易得到费米能的值为 $E_F = \frac{h^2}{8m_e} \left(\frac{3N}{\pi V} \right)^{2/3} = \frac{h^2}{8m_e} \left(\frac{3n}{\pi} \right)^{2/3}$ ， $n = N/V$ 为电子密度。由于 N 的数目很大，在 k 空间，占据区成为一个球，称为费米球，其半径称为费米波矢 k_F 。可以证明， $k_F^3 = 3\pi^2 n$ 。

自由电子模型对于金属传输性能（如导电和导热）等提供了定性解释。金属的高导电性是由于那些处于紧靠费米能的半占有状态上的电子漂移形成（外加电压对大多数电子不产生净效应，因为它们可能跃迁到的较高能态均已被填满）。金属的功函数是从最高的占有能级上取出一个电子所需的能量，在绝对零度时，即为费米能。在室温，只有很少的一些电子被激发到高于费米能，因此功函数在一个宽的温度范围内几乎是恒定的。

自由电子理论能满意地解释绝大多数金属的导电性，但不能正确解释绝缘体。

8.2.2 能带概念的引入

金属自由电子理论中忽略了离子的作用，认为电子是在金属内部的均匀势场中运动的。事实上，固体中电子是在由离子组成的非均匀势场中运动的。对此问题的量子力学处理，得到的能带理论可以用来对金属、半导体、绝缘体的电子结构进行统一描述。

考察由 N 个间距为 a 的正离子周期性排列所形成的一维晶体点阵，其势能如图 8-7 所示，显示出与晶体点阵相同的周期性。两个离子之间的空间势能为零，而在每个离子附近势能具有绝对值较大的负值。一个具有能量为 E_1 的电子，不能通过点阵自由运动，而只能局限于势垒之间的 AB 、 CD 等区域，因为相对于能量 E_1 ，势垒足够高和宽，使得电子穿透它的几率非常之小。这种情形对应于原子内部原子实内的电子，如构成铝的原子实的 $1s$ 、 $2s$ 、 $2p$ 轨道的电子，它们都是定域在特定铝原子的分立原子轨道上，从实际上看，它们的能量和波函数可以认为与孤立原子一样。一个具有较高能量

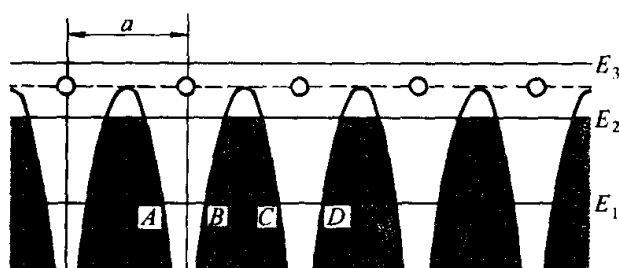


图 8-7 一维晶体点阵中的周期性势能曲线

E_2 的电子, 则不是很强地被束缚于某个特定的离子, 能够以较大的几率穿透势垒而在点阵中移动。而对于具有高于势垒的能量 E_3 的电子, 则不被束缚于任何特定的原子, 能以很大的自由度在整个点阵中运动, 例如铝的 $3s$ 、 $3p$ 价电子, 它们是晶体中离域于任何特定的离子的近自由电子。在晶体中, 它们要占据离域于任何特定离子的, 为整个晶体共有的能级。在含有 N 个铝原子的晶体中, 每个原子贡献一个 $3s$ 轨道, 最终形成一个包含 N 个紧密相间的能级的能带, 称为 $3s$ 价带, 以离域的形式存在。铝中的 $3p$ 能级同样也以离域的 $3p$ 能带的形式存在。

实际晶体是三维点阵, 其势能具有三维周期性。在上述一维点阵的分析可以推广到三维点阵。

当原子相距较大而孤立存在时 (原子间无相互作用), 每个原子的电子都处于原子的特定能级上。当原子间距缩小到晶体中的正常原子间距时, 属于两个原子的同一个能级的电子云就会发生重叠。在分子轨道理论中, 构成双原子分子的两个原子的某一原子轨道相重叠, 结果该原子轨道分裂成离域于两个原子的两个分子轨道, 如图 8-8 所示。对于 N 个原子的分子体系, 每一个原子能级将会分裂出 N 个分子轨道。随着分子轨道数目的增加, 各相邻分子轨道间的平均能隙必然减小, 如图 8-9 所示, 当 $N \rightarrow \infty$, 形成了连续的能带。因为价电子位于原子的最外层, 所以能级的分裂首先是从价电子开始的。

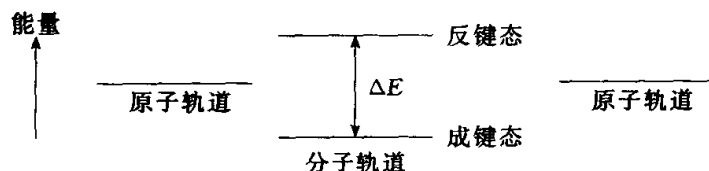


图 8-8 双原子分子的分子轨道

与分子轨道类似, 对于图 8-7 所示的点阵中, N 个近自由电子在周期性势场中运动, 每一个原子态的能级都分裂成 N 个相距很近的能级, 每一个能级都离域于晶体中的所有原子。能级的间隔和位置取决于离子的间距, 如

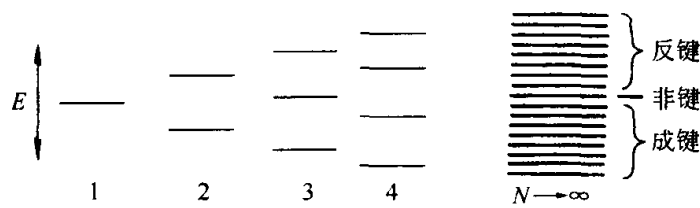


图 8-9 分子轨道理论中的能级分裂

图 8-10 所示。当 N 非常大时，各能级相距非常近，因此可以说它们形成一连续的能带。

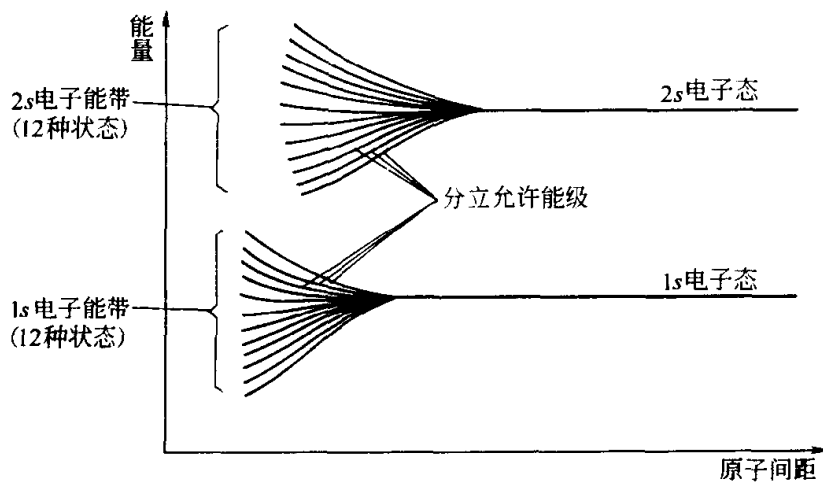


图 8-10 $N (= 12)$ 个原子的系统的电子能量随原子间距离的分布。

原子呈周期性分布，其 $1s$ 和 $2s$ 能级发生分裂形成能带

根据泡利不相容原理，每一个能级能够容纳两个电子，与一给定原子态对应的能带最多可容纳 $2N$ 个电子。这些能带，根据与它们有关的原子态的角动量值，记作 s 、 p 、 d 等。在给定的晶体点阵中，有多个能带，每个都对应于组成此点阵的原子中的一个能级。

“紧束缚近似”认为，电子是比较紧密地束缚于它们原来所属的原子的（例如过渡族金属的 d 电子）。一个通过“紧束缚近似”方法算得的能带结构示于图 8-11。一个特定能带的宽度与原子间距离有关，当离子间距减小时，能带开始

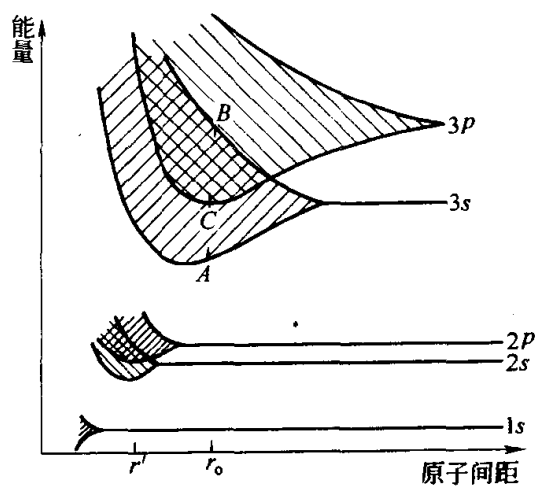


图 8-11 原子间距对原子能级和能带的影响

出现重叠。能带的重叠在解释一些固体性质中是重要的。

与孤立原子的原子实（充满内壳层）相联系的能带中的电子，在一定程度上是定域于各个特定原子的，在讨论固体的能带结构时可以不予考虑。在讨论固体的电子性质时，通常讨论价电子所占有的、与最高原子壳层对应的能带。这个最高能带如未被完全填满，则称为导带，如果完全填满，则称为价带，而将在价带上方的空带称为导带。

8.2.3 布里渊区与能态密度

固体的能带也可以通过考察电子在周期性势中的波函数而导出。点阵的周期性势的效应是改变了在自由电子模型中的自由粒子的平面波函数 $\Psi_k(r) \sim \exp(ik \cdot r)$ ，必须使之具有随点阵周期而改变的调制振幅 $u(r)$ ，而改写为 $\Psi_k(r) \sim u(r)\exp(ik \cdot r)$ ， $u(r)$ 具有点阵的周期。对于图 8-7 所示的线性点阵，波函数可写作： $\Psi_k(x) \sim u(x)\exp(ikx)$ ，并且 $u(x)$ 满足周期性条件 $u(x+a) = u(x)$ 。这样的处理构成布洛赫定理，而受调制的平面波 $\Psi_k(x) \sim u(x)\exp(ikx)$ 称为布洛赫波。布洛赫波函数表征的电子的能量 E 是波矢 k 的函数，一维点阵的 $E-k$ 曲线如图 8-12 所示。在某些 k 值处，能量是不连续的，对于周期为 a 的一维点阵，不连续点的 k 值出现在 $k = n \frac{\pi}{a}$ ， $n = \pm 1, \pm 2, \dots$ 。当 k 值不靠近间断点时，能量与图中虚线所示的自由电子模型的能量是一致的，因此在这些区域，电子通过点阵自由运动。只有在 k 值接近间断点时，点阵才影响电子的运动，其结果是产生能隙。连续能量区域构成能带，在图的右边标出。

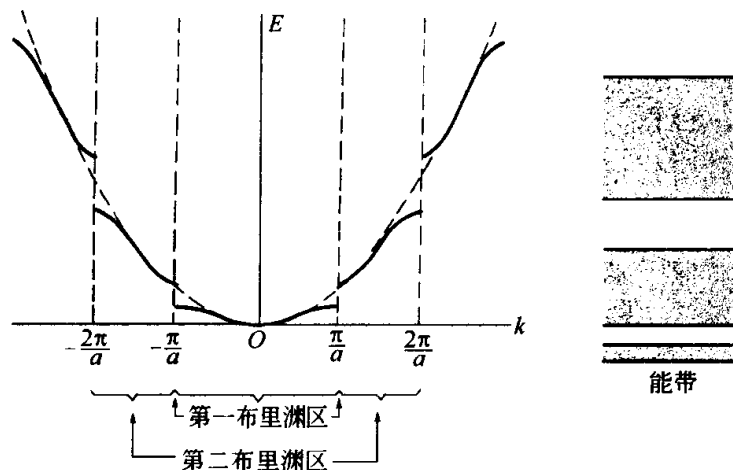


图 8-12 一维点阵的布里渊区和能带

在分析 $E-k$ 曲线时，常引入布里渊区的概念。所谓布里渊区是指在 k 空间中以倒格矢作倒格点，选取一个格点为原点，作由原点到各倒格点的垂直平分面，这些面相交所围成的多面体区域。包含原点的多面体称为第一布

里渊区。往外还有第二、第三，……布里渊区。在每一布里渊区所包含的波矢量数目等于晶体的原胞数，亦即包含了每一能带中的全部电子态。这样就可以用位于布里渊区内和周界上的 k 矢量来表征周期势场中的全部电子态。当 k 的变化限制在布里渊区内时，能量作连续变化，在布里渊区的周界上，才出现能量的不连续变化。因此，第一布里渊区与 $-\frac{\pi}{a} \sim \frac{\pi}{a}$ 之间的 k 值对应，在 $-\frac{2\pi}{a} \sim -\frac{\pi}{a}$ 和 $\frac{\pi}{a} \sim \frac{2\pi}{a}$ 之间的 k 值对应于第二布里渊区，等等。

分析表明，在布里渊区内，电子的速度是变化的，在能带顶和能带底，电子的速度都等于零，在能带的中间区域，电子的速度很接近于自由电子的速度。布里渊区间能量不连续的产生与电子在点阵中运动时受到类似于电磁波在点阵中的布拉格条件的散射有关。在晶体中由于内部的势是周期性的，电子可被势谷强烈地散射，所形成的次级电子不仅相互干涉而且还与初级电子相干，导致价电子能量不可避免地要做出修正，电子不能通过点阵。如果要过渡到更高的波矢，能量将有陡增，于是就产生了能隙。

在布里渊区的边界，能量是不连续的。对于三维晶体， k 空间是三维的。在 k 的不同方向上，布里渊区的边界具有不同的 k 值及相应的能量。两个相邻布里渊区之间的能量关系可以出现两种情况：如果第一布里渊区的所有方向的能态的最高能量均低于第二布里渊区的所有方向的能态的最低能量，则存在禁带，电子在 k 空间先填满第一布里渊区，再填充第二布里渊区。其能态密度曲线具有图 8-13(a) 所示的形状。由图 8-12 可见，在接近布里渊区的边界，能量 E 增加变慢，因此能态密度增高较快。当电子填充到最低能量的布里渊区边界时，能态密度达到最高值。此后，只有部分方向的 k 位于布里渊区内，因此状态密度降低，直到布里渊区完全填满，能态密度降为零。然后电子将填充第二布里渊区，并伴随着电子能量由第一布里渊区最高能量到第二布里渊区最低能量的跳跃。如果在不同 k 方向的布里渊区的边界值中，出现第一布里渊区的某些最高能量低于第二布里渊区的某些最低能量，则发生布里渊区的重叠。其能态密度如图 8-13(b) 所示。电子在填

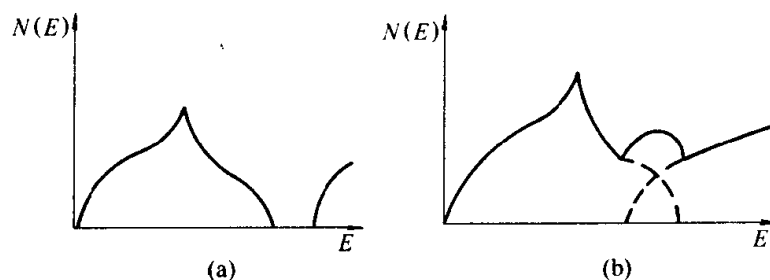


图 8-13 布里渊区的电子能态密度

(a) 相邻布里渊区之间存在禁带；(b) 相邻布里渊区产生重叠

充到第一布里渊区的最高能量之前，就进入第二布里渊区，继续填充第一布里渊区的剩余能态和第二布里渊区的低能状态。

固体能带论获得成功的重要理由之一，是它对于固体的显著不同的电学性质提供了简单的解释。根据能带理论，金属的能带结构如图 8-14(a) 所示。其特征为：最高占有带，即价带，仅部分充满。在图中，占有能级以阴影示出。在能带的费米能级附近的某些电子，在热激发下，或外电场的作用下，能够获得附加的能量，而过渡到该能带内的许多邻近的空能态中的任何一个，而不违背不相容原理。因此具有这种能带结构的物质，是电的良导体也是热的良导体。在独立地占有费米能级附近的状态上的电子可以运动，是这类最高的占有能带未被完全填充的固体成为电或热的良导体的原因。形成部分填充的能带有两种途径：电子数目不足以填满布里渊区，或布里渊区发生重叠，以至于在一个区还未填满时就开始填充下一个区。在金属晶体中，许多都存在布里渊区重叠的情形。

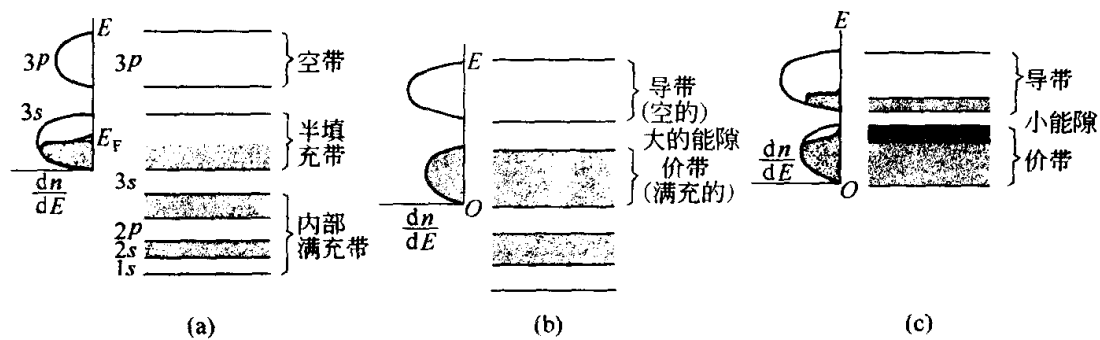


图 8-14

(a) 金属的能带结构；(b) 绝缘体的能带结构；(c) 半导体的能带结构

对于图 8-14(b) 所示的能带结构，一个布里渊区(价带)是填满的，并且不与下一个全空的布里渊区(导带)重叠，由于价带的所有能级全部被占有，电子的能量被“冻结”，即电子不能改变它们在能带中的态而不违背不相容原理，在价带中只有很少的电子具有足够的热能被激发到上面的空带上去。在一般外加电场下，不能加速价带中的电子，因而不能产生净电流。因此电导率很小，可以忽略。这种能带结构构成绝缘体的特征。

对于半导体，其能带结构与绝缘体类似，但在原子的平衡间距处，价带和导带间的能隙要小得多，使得把价带中最上面的电子激发到导带中去要容易得多 [见图 8-14(c)]。因此半导体可看作是这样的一些绝缘体，它们的价带和导带之间的能隙约为 1eV 或更小，因而比较容易实现把电子从价带中激发到导带中去。这是对应于纯态的半导体的情形，这种半导体称为本征半导体，其能态密度如图 8-15(a) 所示。如果半导体存在掺杂，则在填满的价

带和空的导带之间的能隙中有一杂质能级，如图 8-15(b) 所示。当温度升高时，电子由于热激发，可以从杂质能级进入导带，或者从价带顶部进入杂质能级，分别形成 n 型或 p 型半导体。值得注意，介于金属与半导体之间，还可能存在半金属：由于价带与导带在 k 空间的某些方向相互交叠，因而产生金属型的小口袋，导致了载流子浓度甚低的金属型导体，如 As, Sb, Bi 及石墨等。

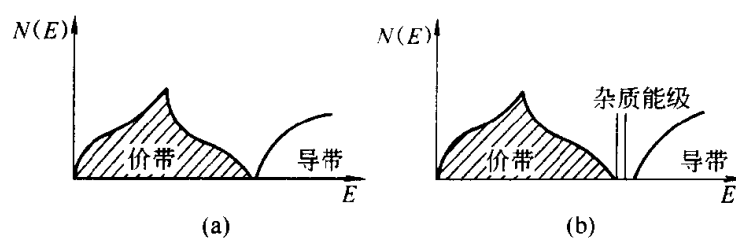


图 8-15 半导体的能态密度曲线
(a) 本征半导体；(b) 杂质半导体

8.3 晶格振动理论

8.3.1 简谐近似

固体中原子由于具有热能而围绕其平衡（点阵）位置而振动，称为晶格振动。晶格振动随着温度的升高而变得剧烈，如果认为晶体中原子的能量分布满足麦克斯韦-玻耳兹曼统计，则这种情况正等价于原子处于较高能量的几率随着温度的升高而增大，并对固体的热学、声学 and 光学性质产生显著的影响。

根据原子间相互作用势可知，当原子间距离大时，原子间是相互吸引的，当原子间距离较小时，原子间是相互排斥的，这些相反的作用力之间的平衡即意味着原子处于平衡位置。在平衡位置附近一对原子间的相互作用力与原子距离偏离平衡间距的大小成比例。因此，可以就晶格振动采用如图 8-16 所示的模型，将晶格设想为质量为 M 的原子通过弹性系数为 k 的弹簧相联系构成的网格。在不十分严格的情况下，可以将晶格振动系统看成由 N 个互相独立的一维简谐振子所组成（ N 为晶格的自由度数）。这种近似称为简谐近似。量子力学给出一维简谐振

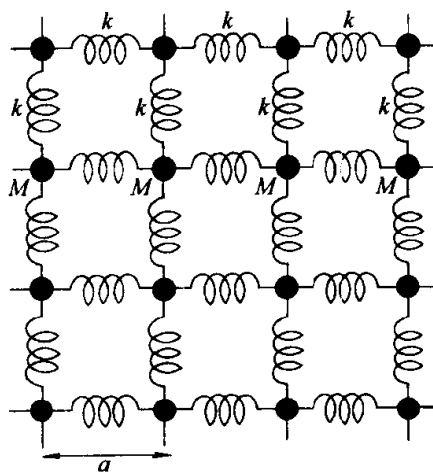


图 8-16 质量为 M 的原子通过力常数为 k 的弹簧相联系，构成晶格常数为 a 的简立方格子

动的允许的能量为: $\frac{1}{2}\hbar\omega$, $\left(1 + \frac{1}{2}\hbar\omega\right)$, $\left(2 + \frac{1}{2}\hbar\omega\right)$, $\dots\dots\left(n + \frac{1}{2}\hbar\omega\right)$, ω 为简谐振动的角频率, n 为整数。系统的最小能量为 $\frac{1}{2}\hbar\omega$, 称为零点能。所以系统的(热)能量可以为 $\hbar\omega$, $2\hbar\omega$, $\dots\dots n\hbar\omega$ 。为此, 引入一个能量为 $\hbar\omega$ 的晶格振动的简正模能量量子, 称为声子。

声子是作为晶格的激发量子引入的。一般地, 将特定晶格振动模式中携带弹性场能量单元的准粒子称为声子。由于简正模式是在整个晶格中传播的行波, 所以与此相应的声子不是局域粒子。声子具有自己确定的动量 $\hbar q$ (q 为波矢量), 因此根据不确定原理, 声子的位置是不确定的。可以将弹性声波看作携带波的能量和动量的声子流。声子的速度等于声在介质中传播的速度。一种谐振子的态 $\epsilon_n = n\hbar\omega$ 是由 n 个声子能量相加而成。显然由于 n 可以取任意的值, 所以声子是玻色子。

这些谐振子的能量分布可以由麦克斯韦-玻耳兹曼统计来描述, 即在温度 T 时处于热平衡状态的振子, 其平均能量为^①:

$$\bar{\epsilon} = \sum_n p_n \epsilon_n \quad (8-8)$$

p_n 为振子处于能级 $\epsilon_n = n\hbar\omega$ 的几率, 即玻耳兹曼分布 $f \sim \exp(-\epsilon/k_B T)$, 所以:

$$\bar{\epsilon} = \frac{1}{2}\hbar\omega + \frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1} \quad (8-9)$$

在简谐近似下, 一个原子可以独立地在三个空间维数上振动, 因而可以被看成为三个各自独立并等价的振子, 它们的能量是非简并的。

8.3.2 爱因斯坦模型与德拜模型

在爱因斯坦最早提出的关于固体比热的量子理论中, 即假定组成三维固体的 N 个原子的振动是互相独立的, 可以视为以相同频率作简谐振动的 $3N$ 个振子, 因此, 在热平衡状态固体的总能量为:

$$E = 3N\bar{\epsilon} = 3N\left\{\frac{1}{2}\hbar\omega_E + \frac{\hbar\omega_E}{e^{\hbar\omega_E/k_B T} - 1}\right\} \quad (8-10)$$

式中, ω_E 称为爱因斯坦频率, 而固体的比热容即为:

① 该平均能量也可以通过声子流的能量求和得到。根据玻色-爱因斯坦统计, 在温度 T 时处于热平衡态的平均声子(能量为 $\hbar\omega$ 的玻色子)数, 可以表示为: $n(\omega) = \frac{1}{e^{\hbar\omega/k_B T} - 1}$, 晶体的统计平

均能量就是这些声子能量之和, 即: $\bar{\epsilon} = \frac{1}{2}\hbar\omega + n(\omega)\hbar\omega = \frac{1}{2}\hbar\omega + \frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1}$ 。

$$c = \frac{dE}{dT} = 3Nk_B \left(\frac{\Theta}{T} \right) \frac{e^{\Theta/T}}{(e^{\Theta/T} - 1)^2} \quad (8-11)$$

其中 $\Theta = \frac{\hbar \omega}{k_B}$ 为爱因斯坦温度。

爱因斯坦模型能够给出正确的固体比热容的高温极限 $3Nk_B = 3R$ ，这个高温极限是杜隆-帕替首先在实验上发现的，并且可由经典理论解释。 N 个原子的固体包含 $3N$ 个振动自由度，每个振动的平均能量为 $k_B T$ ，总的振动能量为 $3Nk_B T$ ，因此比热容为 $c = 3Nk_B = 3R$ 。在高温时，振子被完全激发， $k_B T \gg \hbar \omega$ ，因为 $k_B T$ 与量子阶梯 $\hbar \omega$ 相比大得多，谱的量子性质变得无关紧要了，因此得到经典结果是可以预期的。从高温极限降低温度，在 $\hbar \omega \sim k_B T$ 之前，比热容可保持为 $3R$ 。此后量子效应变得显著，需要采用爱因斯坦模型。在绝对零度时，该模型给出的比热容趋于零。但爱因斯坦模型所给出的 $T \sim 0$ 时的比热容与温度的依赖关系与实验并不一致。模型预计的比热容比观察值偏低（在接近液氮温区，实验测得的比热容按 T^3 变化）。

显然，爱因斯坦模型中，原子相互独立地振动的假设是过于简略了。固体中的原子是存在相互作用的。一个原子的运动实际上要影响所有其他原子的运动（完全独立的原子是不可能达到热平衡的），因此必须将晶格的运动作为一个整体，即考虑晶格的集体振动模式。所以德拜提出的关于晶格原子振动的模型，将晶格看作是连续介质，晶格原子的振动在固体中形成弹性波（声波），这些原子的振动有 $3N$ 个独立的振动模式，每一种模式等价于一个谐振子，其平均能量由式(8-9)给出。这些振动具有相同的振幅和固定的位相关系，而振动频率由零连续分布到一最高频率，具有宽广的频谱（频率分布），并且满足声波的色散关系（频率与波数 $q = \frac{2\pi}{\lambda}$ 的关系式）：

$$\omega = v_s q \quad (8-12)$$

v_s 为波速。这样整个晶格的总振动能为：

$$E = \int \bar{\epsilon}(\omega) g(\omega) d\omega \quad (8-13)$$

$g(\omega)$ 是态密度函数，即 $g(\omega)d\omega$ 是在 $(\omega, d\omega + \omega)$ 内振动模的数目，可以写成：

$$g(\omega) = \frac{3V\omega^2}{2\pi^2 v_s^3} \quad (8-14)$$

V 为晶体体积，这样晶体的总能量表达式为：

$$E = \frac{3V}{2\pi^2 v_s^3} \int_0^{\omega_D} \omega^2 \left[\frac{1}{2} \hbar \omega + \frac{\hbar \omega_E}{e^{\hbar \omega_E / k_B T} - 1} \right] d\omega \quad (8-15)$$

ω_D 为频谱上端的截止频率，可根据模的总数等于整个晶体的自由度数的条

件，即：

$$\int_0^{\omega_D} g(\omega) d(\omega) = 3N$$

而定出为：

$$\omega_D = v_s (6\pi^2 n)^{1/3} \quad (8-16)$$

$n = N/V$ 为晶体的原子数密度。 ω_D 称为德拜频率。这样，根据德拜模型，晶体的比热容为：

$$c = \frac{dE}{dT} = 9R \left(\frac{T}{\Theta_D} \right)^3 \int_0^{\Theta_D/T} \frac{x^4 e^x dx}{(e^x - 1)^2} \quad (8-17)$$

其中引入了变量 $x = \frac{\hbar \omega}{k_B T}$ ，并定义德拜温度 $\Theta_D = \frac{\hbar \omega_D}{k_B}$ 。

由上述德拜模型能够给出正确的高温极限 $c = 3R$ ($T \gg \Theta_D$) 以及在低温 $T \ll \Theta_D$ 时的 T^3 关系。根据德拜温度，在低温时，仅少数的模被激发，这些模的量子能量 $\hbar \omega \geq k_B T$ ，即长波长的模。在一定的温度 T ， $\hbar \omega \gg k_B T$ 的振动模对热容量几乎没有贡献，热容主要来自长波长的振动模，而对于长波长，晶格可看作连续介质。爱因斯坦模型在低温下的偏差正是在于忽略了长波长的振动模。

德拜模型尽管取得了很大的成功，与实验发现仍有距离。德拜模型的近似在于假设了所有可能激发的模，连续介质的色散关系均成立，这导致了德拜温度 Θ_D 是与温度无关的量，但实验发现 Θ_D 随温度的改变可高达 10% 以上。

8.3.3 格波与晶格振动模

在德拜模型中，将固体作为一个连续体，采用了连续介质的线性色散关系 $\omega = v_s q$ ，当原子间距比波长小得多，这种近似是合理的。但当波长减小， q 增加时，晶格的不连续性就显得重要了，此时，原子会对波产生散射， q 越大，散射越强，结果使得色散关系偏离了线性。这时，必须将固体作为由分立的原子组成的体系来处理。晶体中的原子总是围绕其平衡位置不断振动，由于原子之间的相互作用，各个原子同时运动，所以必须考虑整个晶格的各个原子的振动，系统中所有原子以相同的频率（简正模）振动，这些振动相互间存在固定的位相关系，因此在晶格中存在着角频率为 ω 的平面波，这些波称为格波。

对于格波，数学上可以用如图 8-16 那样的弹簧连系着的质量为 M 的原子链模型来处理，通过直接解原子链的耦合的联立运动方程，可求得其振动模。由此得到所谓格波解。由于晶格的平移对称性，格波的解具有平面波的形式，系统的所有基元以相同的频率振动，这些振动模式即为简正模。简正

模的角频率与波矢之间具有确定的关系，即色散关系，或称振动频谱。

图 8-17 为一维布拉菲格子的 $\omega \sim q$ 色散关系 $\left[\omega = \omega_m |\sin(qa/2)|, \omega_m = \left(\frac{4a}{M} \right)^{1/2} \right]$ ，在 q 空间为周期等于 $\left(\frac{2\pi}{a} \right)$ 的正弦曲线（ a 为晶格常数），并且关于原点对称。在长波极限 $q \rightarrow 0$ ，色散关系近似为 $\omega = \left(\frac{\omega_m a}{2} \right) q$ ，为线性关系。在这个极限，晶格相当于一个弹性连续介质，格波与长波长声波等同。

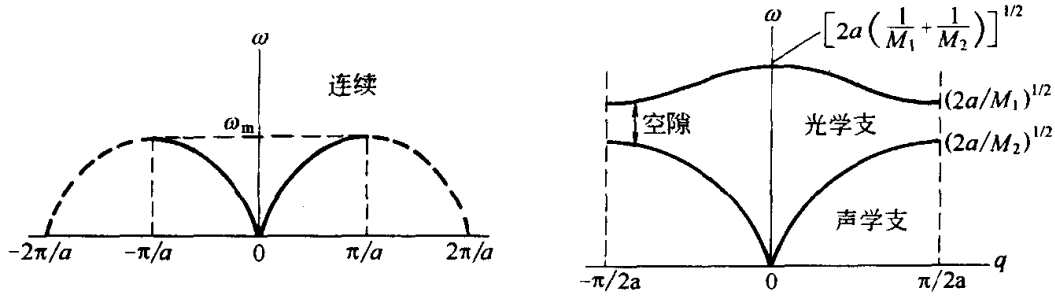


图 8-17 一维单原子链的 $\omega \sim q$ 色散曲线 图 8-18 一维双原子链的 $\omega \sim q$ 色散曲线

对于单胞由质量为 M_1 和 M_2 两种原子组成的双原子晶格，其 $\omega \sim q$ 色散关系为：

$$\omega^2 = a \left(\frac{1}{M_1} + \frac{1}{M_2} \right) \pm \sqrt{\left(\frac{1}{M_1} + \frac{1}{M_2} \right)^2 - \frac{4\sin^2(qa)}{M_1 M_2}} \quad (8-18)$$

a 为两相邻原子间距，与式中正、负号相对应，有两个色散关系，或者说两支色散曲线（见图 8-18），与式中负号相应的曲线通过原点（ $q=0, \omega=0$ ），与单原子晶格的色散关系曲线一样，在 $q \rightarrow 0$ 的长波极限表现为长波长的弹性波，故把这一支曲线叫做声学支，相应的振动模式称为声学模。它描述了原胞内原子质心的运动，即分子（实际上是整个晶格）如同刚体一样，质心来回振动。与式中正号相应的那支曲线，频率近似地为 $\left(\frac{2a}{M} \right)^{1/2}$ ，其值约位于红外光区。这支曲线称为光学支，相应的振动模式为光学模。它描写了原胞内原子的相对运动，在离子晶体中，由于离子的相对运动产生电偶极矩的振动，从而与电磁波产生强的耦合（共振），引起对远红外的强烈吸收。

对于三维单原子布拉菲晶格，每个单胞有一个原子，由原子的耦合运动方程可以得到如图 8-19 所示的三支色散曲线，所有这三支曲线都通过原点，表明它们都是声学支。这三支曲线具有不同的偏振。当波矢 q 沿高对称方向（如 $[100]$ 或 $[110]$ 方向），这些波可分成两支纯横波，一支纯纵波，分别记为 TA 支（横向声波）和 LA 支（纵向声波）。在一定的对称方向，会发生两个横向支重合，称为简并。当波矢 q 沿非对称方向，波可以不是纯

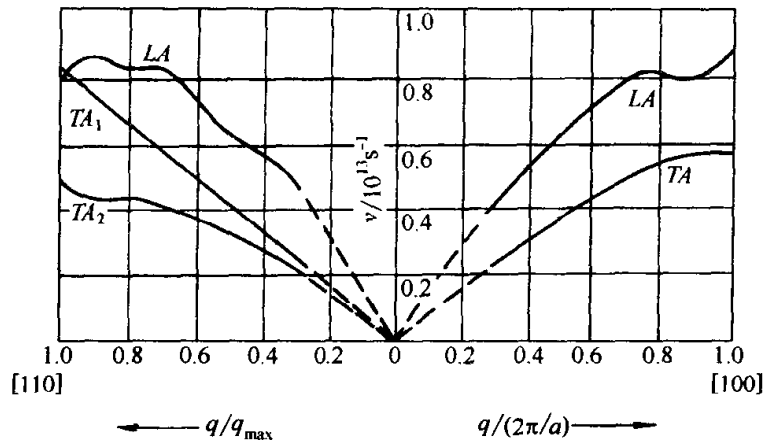


图 8-19 三维单原子布拉菲晶格的 $\omega \sim q$ 色散曲线
(此图给出的是 Al 晶体在 $[100]$ 和 $[110]$ 晶向的色散关系)

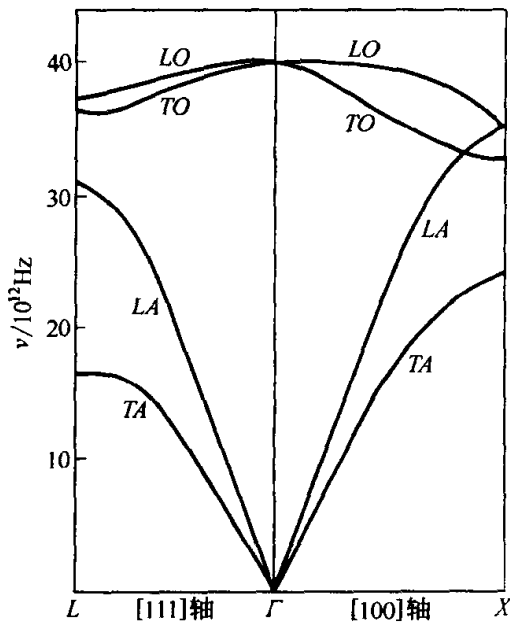


图 8-20 具有三维复式格子的金刚石晶体的晶格振动色散曲线

纵向或纯横向，具有混合的特征。所以，偏振和简并与相对于传播方向的晶体的对称性有关。

对于三维复式格子，如果每个单胞中有 n 个原子，则可得 $3n$ 支色散曲线，其中 3 支是声学支 ($3n - 3$) 支为光学支。声学支按其偏振可以分为横声学支 TA_1 、 TA_2 和纵声学支 LA 。光学支也可以分为纵向和横向，称为 LO (纵光学支) 和 TO (横光学支)。图 8-20 所示为面心立方晶格的金刚石晶体的色散曲线。其每个原胞中有 2 个原子，但在 $[100]$ 和 $[111]$ 轴上，其横声学支和横光学支都是双重简并的。

8.4 相互作用的电子体系

8.4.1 准粒子体系

金属的自由电子理论和固体能带理论，是基于单电子近似获得的结论。在固体物理中，将固体中的原子分为价电子和离子芯两部分。Born 和 Oppenheimer 在讨论分子中电子状态时认为，由于离子芯的质量远大于电子的质量，并且在常温下，离子芯总是围绕其平衡位置作小的振动（晶格振动），其速度远低于电子的运动速度，因此在讨论电子的运动时，可以将离子芯视

为静止不动的，从而将电子和离子芯作为相互独立的系统考虑，此即 Born-Oppenheimer 绝热近似。在金属自由电子模型中，离子芯只是提供均匀分布于金属所占空间的正电荷背景，以维持系统的电中性，而电子则被视为在正电荷背景上运动的自由电子。在能带理论中，则考虑单电子在离子芯提供的周期性势场和其他电子的平均场中的运动。上述处理中，完全忽略了电子与电子间，电子与离子芯之间和离子芯与离子芯之间的相互作用（关联）。这时，每个电子的能量状态除了必须服从泡利不相容原理，几乎与其他电子的占有状态无关。

实际上关联总是存在的。特别是电子与电子之间的关联，在不同的体系中，会有不同的物理表现，因而常常是不可忽略的。8.1.2 节中指出，电子是费米子，单电子态的平均占有数服从费米-狄拉克分布，因而在一个电子系统中不可能有两个电子具有相同的一组量子数，这一泡利不相容原理是固体电子结构理论的基础。电子-电子之间的相互作用主要有两种：受到屏蔽的电子之间的排斥力和通过晶格（声子）作为媒介而发生的吸引力。金属中的价电子在离子芯的正电背景中运动，这种正负电荷浓度相等的体系类似于等离子体。通常将正离子芯认为是固定不动的，由于电子间长程的库仑强相互作用，可发生电子的集体振荡，称为等离子体振荡（plasma oscillation），等离子体振荡的量子称为等离激元（plasmon）。在一般金属中，等离激元约为 $\hbar\omega \approx 12\text{eV}$ 。金属中的等离子振荡不会被电子的热运动所激发，此时，由于电子间的库仑排斥作用，使得一个电子周围的其他电子被排斥而远离，导致该电子周围的负电荷密度低于平均值，等效于在该电子周围出现等量的正电荷背景，包裹这个电子，从而屏蔽了该电子在远处的电场。这种屏蔽效应使得电子-电子间的相互作用成为：

$$V(r) = \frac{e^2}{r} e^{-r/r_D} \quad (8-19)$$

称为屏蔽库仑势。式中 r_D 为 Debye 屏蔽长度，与金属中电子间的平均距离相当。此外，同自旋的电子受泡利不相容原理所造成的相互排斥，即交换相互作用（见下节），也可在运动电子周围产生相同自旋密度降低的区域，从而对交换相互作用产生屏蔽。因此，电子-电子相互作用的结果，使一个电子周围由库仑和交换作用造成的电子贫乏区与电子一齐运动，构成准粒子。在通常情况下，准粒子具有独立粒子的许多特征。所以，基于单电子理论的能带模型在许多金属、半导体中取得成功。

8.4.2 强磁性与超导电性

电子间的相互作用还导致了一些重要的物理现象，如强磁性和超导电性。在多电子体系中还存在着自旋之间的相关性，或者说不同阵点上离子的

磁矩之间的库仑相关性。自旋间的关联是固体强磁性的物理起因，磁效应主要来自电子间的库仑相关，而非磁矩间的磁相互作用。后者要比前者小 u/c 的量级，因而是一种相对论相关。由电子产生的磁矩约为 $\mu_B (L + 2S)$ ， L 为电子的轨道角动量， S 为电子的自旋角动量。考虑磁矩-磁矩间的磁相互作用，其作用能 ΔE 约为 $\mu_0 \mu_B^2 / 4\pi a^2$ ， a 为偶极子间的距离。典型的值为 $\Delta E \approx 10^{-4} \text{eV} \sim k_B \cdot 1\text{K}$ ，因此在 1K 以上温度，磁效应就会被热涨落所掩盖。所以，仅仅考虑在外磁场中原子磁矩即原子的束缚电子的轨道运动和自旋运动，以及金属中传导电子的自旋引起的顺磁贡献，或磁场感生的轨道运动所引起的抗磁贡献，可以正确地解释顺磁性和抗磁性等弱磁材料的磁性，却不能解释铁磁性和反铁磁性等强磁材料的磁性。对这些材料，在常温下存在自发磁有序，磁矩自发地同向或反向排列。如果仅仅考虑磁矩间的磁相互作用，磁有序在常温下足以被热扰动所破坏而不可能存在。假如，材料的居里温度（由铁磁性转变为顺磁性的温度）为 $T_c \sim 1000\text{K}$ ，为保持相邻磁矩的自发有序排列所需克服的热涨落能量约为 $k_B T_c \sim 0.1\text{eV}$ ，这比磁偶极矩相互作用能要大几个数量级。也就是说，铁磁系统中磁矩间的定向耦合应来自于比磁偶极相互作用强得多的相互作用类型。1928 年，Heisenberg 指出，这种相互作用应是相邻电子自旋间的交换相互作用（exchange interaction）。根据 Heisenberg 模型，两个相邻电子间的静电作用能为：

$$E = K_e + V_{\text{ex}} \quad (8-20)$$

其中 K_e 为库仑积分能， V_{ex} 为交换积分能，它是趋向于使平行自旋分离的泡利不相容原理作用的结果，并被称为交换相互作用。根据不相容原理，在交换所有电子坐标（包括空间和自旋）的情况下，任意两个电子波函数是反对称的，也就是说如果空间波函数是对称的，则自旋波函数必须是反对称的，反之亦然。由于空间对称性和自旋对称性间的关系，自旋平行或反平行的两个态的能量是有差异的，为此引入交换积分能：

$$V_{\text{ex}} = -2JS_1 \cdot S_2 \quad (8-21)$$

S_1 、 S_2 为两个电子的自旋， J 称为交换常数。能量最小的稳定性条件要求 $V_{\text{ex}} < 0$ 。

例如，对于含有两个电子的氢分子系统， $J < 0$ ，因而自旋反平行的状态能量更低，其基态相应于反铁磁性状态。在固态系统中， J 也可能出现正值，此时，能量最低的稳定状态要求相邻两个原子的自旋平行，即处于铁磁性状态。如果 $J < 0$ 且两个原子的自旋不相等时，则出现亚铁磁性。

交换相互作用本质上是由波函数的对称性引起的静电效应，其强度基本上由库仑作用 $V \approx \frac{e^2}{4\pi r^2}$ 给出，因而交换相互作用是一种很强的相互作用。

自旋平行的电子倾向于远离，而在自旋反平行状态，电子可以靠得较近，且波函数显著交叠，产生强的库仑排斥，系统能量升高，因此自旋平行的状态系统能量更低，这正是 $J > 0$ 的铁磁系统。但固体中原子内的电子不仅受到其他电子的斥力，而且受到邻近原子核的引力（如氢最为典型），因此其竞争的结果，也可出现负的 J 值，即反铁磁和亚铁磁的情形。

上述发生于两个原子的自旋间的交换相互作用是最简单的情形，称为直接交换作用。实际的磁有序系统中存在更复杂的情形。典型的 Fe、Ni、Co 等金属的铁磁性是基于非定域导带电子的巡游电子交换作用；稀土金属的铁磁性与反铁磁性则是由于自由电子作为中介的间接交换作用；而在氧化物中铁氧体的亚铁磁性与反铁磁性则归源于超交换作用和双交换作用。请参阅第 10 章。

超导电性的根源在于电子由于吸引相互作用而配对。两个电子间的直接库仑相互作用是排斥的，所以要形成吸引力需要其他的媒介。成功地解释了常规超导体各种现象的 BCS[●] 理论就基于电子与晶格振动的耦合所形成的配对效应。

设想在费米面上的一个电子（电子 1），具有很高的速度 v_F ，当电子通过晶格时，将与晶格上的正离子发生相互吸引，但由于离子的质量大，正离子点阵的形变的发生相对于电子运动而言是极其缓慢的，达到最大形变的时间约为 $\tau \sim \frac{2\pi}{\omega_D} \sim 10^{-13} \text{s}$ （ ω_D 为晶格振动的德拜频率），而此时电子 1 已远离到 $v_F \tau$ 约 100nm 处。由于库仑屏蔽，在形变点阵处的另一个电子（电子 2）在电子 1 通过时，并不能显著地受到两个电子间的库仑排斥作用，但当电子远离后，形变的离子点阵将使其所在区域带有正电荷并对电子 2 产生吸引。按照晶格振动理论的语言，这种过程正是电子 1 发射声子，该声子很快为电子 2 所吸收，因而通过声子交换引起电子 1 和电子 2 之间的相互吸引。图 8-21 形象地示出了上述过程。

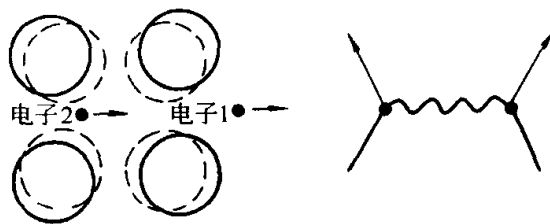


图 8-21 电子和电子之间通过声子交换引起的相互吸引

由于电子之间的这种相互吸引，在费米面附近的两个电子，特别是具有相反动量和相反自旋取向的电子 $k \uparrow$ 和 $-k \downarrow$ ，将形成一束缚态，获得最高

● BCS 理论是以该理论的创始人，Bardeen, Cooper 和 Schrieffer 而命名的。该理论由他们于 1957 年建立，并合获 1972 年诺贝尔物理学奖。

的结合能,从而使系统能量比不束缚时为低。这样两个配对电子称为库柏对 (Cooper pair)。库柏对为玻色子,因此在超导转变温度下,费米面附近的电子可两两配对存在,在动量空间凝聚成库柏对的体系,形成超导态。在绝对零度,在费米面附近能量小于 $\hbar \omega_D$ 范围的电子全部形成库柏对,就构成超导基态。库柏对结合的结果,在费米能级与库柏对的束缚态之间产生一个能隙,其值与库柏对的结合能相关。在绝对零度,超导能隙为^①:

$$\Delta_0 = 4\hbar \omega_D e^{-[2/g(E_F)V']} \quad (8-22)$$

$g(E_F)$ 为正常金属在费米能级处的态密度, V' 为电子-晶格相互作用强度。粗略地, $\Delta_0 \sim \omega_D$, 正是一个典型声子的能量。在超导能隙下方的电子为超导电子 (库柏对), 能隙上方的电子为正常电子。超导电子具有零熵 (完全有序), 不受散射, 并有很长的相干长度 (coherence length, 约 10^3 nm), 因而具有无限大的电导率。正是它们赋予超导体的超导电性。在绝对零度的超导基态, 在费米面附近能量小于 $\hbar \omega_D$ 范围的全部电子凝聚于超导能隙以下。而在低于临界温度 T_c 的有限温度, 部分库柏对会由于热激发被拆散成为正常电子, 从而使超导体内具有两种载流子: 库柏对与单粒子态的正常电子。但只有库柏对对超导电性产生贡献。

8.4.3 强关联电子体系

准粒子概念和能带理论在解释常规导体、半导体和绝缘体的物理性质时是非常成功的。然而能带理论并不总是左右逢源的。当晶格常数加大、能带变窄时, 电子间的库仑屏蔽减弱, 电子间的相互作用将变得重要。单电子模型的能带理论将因电子间强的关联而失效, 典型的如过渡族金属化合物 NiO 、 MnO 、 CoO 等。例如, MnO 中费米能位于半满的 $3d$ 带, 其下 O^{2-} 的 $2p$ 能带是满带, 并且与 $3d$ 带不重叠。因此根据能带理论, 该化合物应为金属, 具有没有能隙的金属性能带。但是实际上该化合物中由于电子被电子间的强的库仑排斥所限制, 不能像金属中那样自由地在格点上移动, 从而表现为绝缘体, 这种固体称为 Mott 绝缘体。有些 Mott 绝缘体会随温度升高而从绝缘体转化为金属, 称为 Mott 转变。

Mott 转变被认为是一种典型的窄能带现象。以 Na 为例, 其 $3s$ 能带是半满的, 根据能带理论, 不论其晶格常数 a 如何变化, $3s$ 带总是半满的, 因而仍能保持为导体。然而, 事实上并非如此。当 a 增加到临界值 a_c 时, 电导率突然降为零, 使 Na 变为绝缘体, 即发生金属-绝缘体转变。

在能带理论中, 假定金属中的传导电子是非定域的, 扩展于整个晶体

① 根据 BCS 理论, 能隙随温度的变化遵从以下关系: $\Delta(T) = \tanh \left[\frac{T_c \Delta(T)}{T \Delta_0} \right]$ 。

中，称为扩展态。这是造成金属导电性的原因。能带理论中允许两个或更多的电子存在同一个格点。考虑电子间的关联，在严格意义上这是不可能的。当能带较宽时，电子可以调整其动能，补偿库仑能的增加，使得电子间的关联可以忽略。当 a 增加时，能带宽度减小，电子动能也相当小，便不足以补偿库仑能了。这时，晶体中电子运动不再是布洛赫形式那样扩展到整个晶体，而是以各格点为中心的定域化，即处于定域态。由于轨道是定域化的，晶体中的原子的行为像自由原子的情况一样，电导率降为零，成为绝缘体。

可见，传导电子由扩展态进入定域态，导致材料由导体变为绝缘体。通过改变材料的有序度，可使电子运动从扩展态进入定域态，材料由金属转变为非金属，这类金属绝缘体转变为 Anderson 转变。而在 Mott 转变中，电子运动从扩展态到定域态的转变则源于电子-电子作用强度的效应。

Hubbard 首先讨论了窄能带中的电子关联问题。他假定，当两个自旋相反（泡利不相容原理要求）的电子占据同一原子轨道时，相互作用为 U ，否则相互作用为零。通常将电子间的库仑相互作用 U 大于能带宽度 B 的体系称为强关联电子体系。MnO 晶体等即属于这种强关联电子体系。如图 8-22 所示，由于电子间的库仑相互作用，使得原来不存在关联的（或关联弱时）半满的 d 带分裂成间隔为 U 的上下两个能带，分别成为上 Hubbard 带和下 Hubbard 带。当原子间距较大时，能带宽度很窄（极端的情形，当原子成为孤立时，能带转化为能级），上下 Hubbard 带相互分离，并且下 Hubbard 带全满而上 Hubbard 带全空， d 电子是定域态，因而体系为绝缘体。当温度升高，会使下 Hubbard 带的电子激发到上 Hubbard 带，并在下 Hubbard 带中留下空穴，电子-空穴由库仑相互作用而形成束缚态成为激子。在一定温度下，激发到上 Hubbard 带的电子足够多，将屏蔽电子-空穴间的束缚，导致 d 电子的退定域化，从而使 Mott 绝缘体过渡到金属态。或者，当原子间距减小，造成能级展宽，上下 Hubbard 带发生交叠，转化为部分充满的能

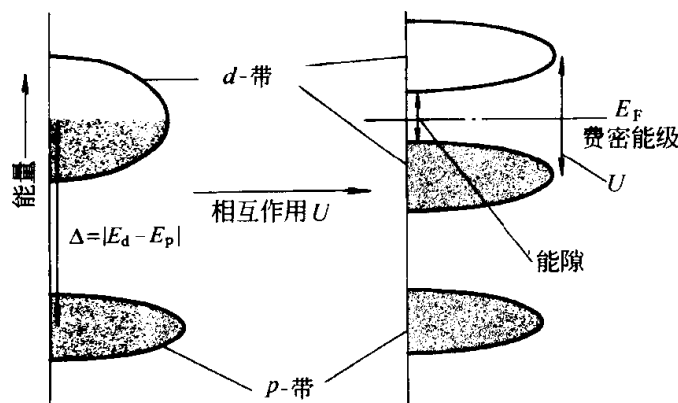


图 8-22 Mott 绝缘体的能级示意图

带, 使 Mott 绝缘体转变到金属态。

正如能带绝缘体(半导体)可以通过掺杂来控制其导电性能, Mott 绝缘体也可用掺杂来控制其导电性能。而掺杂的 Mott 绝缘体的性能相当特别: 例如具有钙钛矿结构的铜氧化物高温超导材料, 正是一类典型的强关联体系。这类高温超导材料结构上的共性是都具有由 CuO_2 层为主体的超导单元。 CuO_2 层的电子间存在很强的关联, 例如, 其 d 电子的库仑关联能 U 高达 $8 \sim 10\text{eV}$, 远大于相应的带隙。因而这些钙钛矿结构的铜氧化物在纯净状态(或称母化合物)为典型的 Mott 绝缘体, 并表现出反铁磁有序。超导电性通常通过掺入少量杂质而得到。在母化合物中的掺杂使得在价带出现空穴, 从而载流子浓度增加, 在一定的最佳掺杂量下, 材料具有高临界温度的超导电性。而在转变温度以上, 较高浓度的载流子使材料成为金属性。通常认为, 对于这类高温超导体材料的强关联电子的研究, 是最有可能接近其超导机制的途径。而锰氧化物的庞磁电阻现象又是掺杂 Mott 绝缘体作为强关联电子导体的另一例证。这表明对强关联电子系统的研究有可能发掘出一些新型具有优异电子性能的材料。

参 考 文 献

- 1 冯端等著. 金属物理学·第一卷. 北京: 科学出版社, 1987
- 2 黄昆, 韩汝琦编著. 固体物理学. 北京: 高等教育出版社, 1988
- 3 Kittel C. Introduction to Solid State Physics. 7th ed., New York: John Wiley & Sons Inc., 1996
- 4 Rosenberg H. M. The Solid State. London: Oxford University Press, 1975
- 5 Ziman J. M. Electrons in Metals. London: Taylor & Francis Ltd, 1962
- 6 Hummel R. E. Electronic Properties of Materials. 2nd ed., Berlin Heidelberg: Springer-Verlag, 1993
- 7 Edwards P. P. and Johnston R. L. et al., A Perspective on the Metal-Nonmetal Transition. Solid State Physics Adv. In Research and Application, New York: Academic Press, 1999, 52: 229
- 8 Ketterle W. Physics Today, 1999, 52(12): 30
- 9 Hall H. E. Solid State Physics, New York: John Wiley & Sons Ltd., 1974
- 10 Anthony A. R. Solid State Chemistry and its Applications, New York: John Wiley & Sons Inc., 1984
- 11 Tokura T. Correlated Electrons: Science to Technology, JSAP International, 2000, 2: 12

第9章 输运性质^[1]

撰稿人 袁仁宽

本章介绍的内容是固体中的载流子在外力作用下的运动（载流子的输运）及其由此产生的一些物理现象。导电性是物质最重要的输运性质之一，它源于载流子在电场作用下的迁移运动。电导率是描述物质导电性强弱的物理参数，物质的室温电导率最低可达 $10^{-20}(\Omega \cdot \text{cm})^{-1}$ ，最高可达 $10^5(\Omega \cdot \text{cm})^{-1}$ ，彼此相差 25 个量级，如此大的变化范围在各种物性中堪称之最。人们按电导率的大小将材料分成金属、半导体和绝缘体三大类，如图 9-1 所示（另有一类性质特殊的超导体将在第 12 章中专门介绍）。

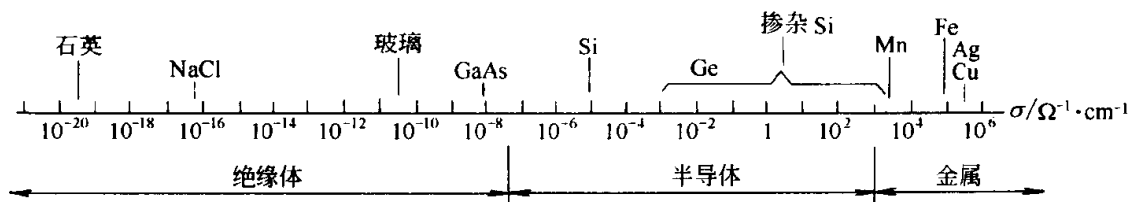


图 9-1 金属、半导体和绝缘体的室温电导率

载流子除了在电场作用下发生迁移运动外，还可以在其他场和力的作用下运动。本章将涉及到载流子在温度梯度场、浓度梯度场和洛伦兹力作用下的运动，并将涉及到因费米能级（化学势）之差产生的运动。

不同类型的物质具有不同的载流子，载流子在物质中的运动受到不同因素的影响而产生各种效应。人们可以通过施加一定的外界作用来控制载流子的运动，以实现所希望的功能。为此，人们必须对载流子在外力作用下运动的规律和机理有足够的认识。

固体中的输运，除了载流子的输运外，还应包括其他的输运，如原子的输运等，但这不在本章介绍的范围之内，在其他章节中将有有关叙述。

9.1 金属与合金的输运性质

金属是室温导电能力最强的材料，其室温电导率在 $10^3 \sim 10^5(\Omega \cdot \text{cm})^{-1}$ 范围内。为改善材料的某些性能，人们将金属制成合金。金属及其合金中的载流子是电子，本节介绍了电子的经典电导理论和量子电导理论，并将理论

与实验进行对照, 讨论了金属及其合金的电导率随温度的变化以及与材料成分和结构的关系。电子在运输的过程中有时伴有热能的交换, 这便产生了热电效应。由于在半导体材料中也会产生热电效应, 故在本节一并作了介绍。

9.1.1 经典电导理论

为方便起见, 我们以一价钠晶体为例讨论金属的电导。钠晶体中, 每个钠原子可以提供一个自由电子, N 个钠原子可以提供 N 个自由电子, 从而形成“电子气”。自由电子的浓度等于原子密度 n , 为 10^{22} 量级。在一定温度下, 这些电子作无规则的热运动, 没有定向的流动。当有电场 E 存在时, 电子产生定向运动, 形成电流。在电场力 $-qE$ 的作用下, 电子作加速运动, 速度应愈来愈快, 电流也应愈来愈大。然而, 实际情况并非如此, 电流通常维持一稳定值。这说明在电子的运动过程中有阻力存在, 电子流动的阻力来源于与晶格原子的相互作用(碰撞)。假设电子所受阻力的大小正比于电子的漂移速度 v , 则有如下关系式

$$m_0 \frac{dv}{dt} + \gamma v = -q|E| \quad (9-1)$$

其中, m_0 为电子质量, q 为电子电量, γ 为阻尼系数。在稳态情况下, $v = v_0$, $dv/dt = 0$, 则有

$$\gamma = \frac{-q|E|}{v_0} \quad (9-2)$$

将此值代入式(9-1), 可得瞬态方程

$$m_0 \frac{dv}{dt} + \left(\frac{-q|E|}{v_0} \right) v = -q|E| \quad (9-3)$$

假设, $t = 0$ 时, $v = 0$, 则瞬态方程的解为

$$v = v_0 \left[1 - \exp\left(-\frac{t}{\tau_0}\right) \right] \quad (9-4)$$

其中

$$\tau_0 = \frac{m_0 v_0}{-q|E|} \quad (9-5)$$

τ_0 是每个电子与晶格相继两次碰撞的平均时间间隔, 被称为弛豫时间。由此可得稳态的电子漂移速度 v_0 和电流密度 J

$$v_0 = \frac{-q\tau_0}{m_0} |E| \quad (9-6)$$

$$J = nv_0 q = \frac{nq^2\tau_0}{m_0} |E| \quad (9-7)$$

则材料的电导率为

$$\sigma = \frac{J}{|E|} = \frac{nq^2\tau_0}{m_0} \quad (9-8)$$

由此可知, 电子浓度 n 愈高, 弛豫时间 τ_0 愈长, 则材料的电导率 σ 愈高。弛豫时间 τ_0 正比于平均自由程 l , 平均自由程 l 是电子与晶格相继两次碰撞间所走过的平均路程。

$$l = v_0 \tau_0 \quad (9-9)$$

9.1.2 量子电导理论简介^[2]

量子电导理论与经典电导理论的一个重要区别在于, 经典电导理论认为, 在外电场的作用下所有 (自由) 电子都对电流有贡献; 而量子理论认为, 只有费米能级附近的电子才对电流有贡献。为讨论方便起见, 假设温度为绝对零度, 按照统计理论, 费米能级以下的状态占满了电子, 费米能级以上的状态全空。在没有外电场作用的情况下, 所有电子的速度都在速度空间中一球心位于原点的圆球内, 电子的最大速度为球面上的速度, 即能量为费米能的电子所具有的速度 v_F , 此速度称之为费米速度。此球称之为速度空间的费米球, 费米球的表面称之为费米面。费米球内被电子填满, 其电子数等于状态数。图 9-2(a) 是无外场时二维速度空间电子速度示意图。因无外场存在, 故在各个方向上, 电子速度大小的分布相同。由于每两个速度大小相同、方向相反的电子的运动互相抵消, 因此所有电子的净流动为零。当有外电场 E (沿 X 方向) 存在时, 整个费米球沿与电场相反的方向发生移动。此时, 绝大部分电子的运动仍然因两两大小相等方向相反而互相抵消, 剩下一部分在费米面附近 [如图 9-2(b) 中斜阴线所示] 的电子的运动未被抵消, 从而形成电流, 其电流密度

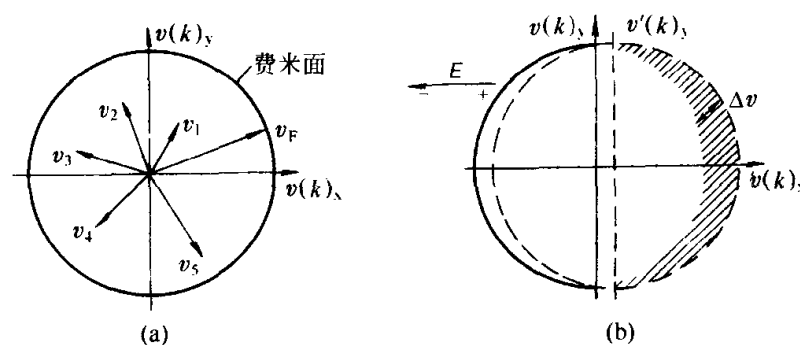


图 9-2 二维电子速度空间示意图

(a) 平衡时; (b) 外加电场时

$$J = -q v_F N' \quad (9-10)$$

其中 N' 为斜阴线部分的电子数 (单位体积, 下同)

$$N' = N(E_F) \Delta E \quad (9-11)$$

式中, $N(E_F)$ 为费米面处单位能量间隔中的电子数。由如图 9-3 所表示的 $N(E)$ 曲线可知, 费米面处 $N(E_F)$ 值很大, 因此较小的 ΔE 就可产生较大的

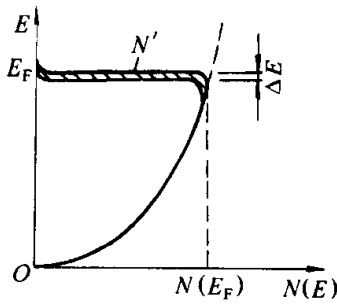


图 9-3 自由电子数 (单位 E , 则 t 时刻的 Δk 为

体积) 随能量的变化
 $N(E)$ 及 N'

$$\Delta k = \frac{-qt}{\hbar} |E| \quad (9-13)$$

因此

$$\frac{d(\Delta k)}{dt} = \frac{F}{\hbar} \quad (9-14)$$

考虑到因散射引起的阻力, 则有

$$\hbar \left(\frac{d}{dt} + \frac{1}{\tau'_0} \right) \Delta k = F \quad (9-15)$$

其中, 弛豫时间 τ'_0 为电子相继两次被散射的平均时间间隔。在静电场 E 的作用下, 系统处于稳态, 即 $d(\Delta k)/dt = 0$ 。则

$$\Delta k = \frac{F\tau'_0}{\hbar} = \frac{-q\tau'_0}{\hbar} |E| \quad (9-16)$$

对于自由电子

$$E = \frac{\hbar^2 k^2}{2m_0} \quad (9-17)$$

费米面附近, v 为 v_F , 则有

$$\frac{dE}{dk} = \frac{\hbar^2 k}{m_0} = \frac{\hbar p}{m_0} = \hbar v_F \quad (9-18)$$

由式 (9-18) 和式 (9-16) 可得因外电场 E 的作用而引起的费米面的移动量

$$\Delta E = \left(\frac{dE}{dk} \right) \Delta k = -q\tau'_0 v_F |E| \quad (9-19)$$

将此式代入式 (9-11), 再代入式 (9-10) 可得

$$J = q^2 v_F^2 N(E_F) \tau'_0 |E| \quad (9-20)$$

由于只有平行于电场方向的运动才对电流有贡献, 所以必须将 v_F 投影到 X

轴上。如此得

$$J = \frac{q^2 v_F^2 N(E_F) \tau'_0}{3} |E| \quad (9-21)$$

则有

$$\sigma = \frac{q^2 v_F^2 N(E_F) \tau'_0}{3} \quad (9-22)$$

由上式可知, 材料的电导率正比于费米面处单位能量间隔的电子数 $N(E_F)$ 。

由图 9-4 可以看出, 不同种类的材料应具有不同的电导率。一价金属材料的

格振动是晶体中两类彼此独立的散射源，因此有

$$\rho = \rho_{\text{不完整性}} + \rho_{\text{晶格振动}} \quad (9-24)$$

剩余电阻率是由晶体的不完整性引起的，它不随温度改变。晶体的不完整性有杂质和缺陷两种，则

$$\rho_{\text{剩余}} = \rho_{\text{不完整性}} = \rho_{\text{杂质}} + \rho_{\text{缺陷}} \quad (9-25)$$

晶体缺陷包括空位、间隙原子、晶粒间界和位错等。材料的加工处理过程对材料结构的完整性有十分重要的影响，经过淬火的材料会有大量的缺陷，因此电阻率高。如果将这样的材料进行退火再结晶处理，使大量缺陷消失，晶体的完整性大大提高，电阻率将显著下降。

晶格振动引起的散射与晶体不完整性引起散射的明显不同在于前者对温度的强烈依赖性，温度愈高，晶格振动愈激烈，电子的平均自由程愈小，弛豫时间 τ_0 愈短，散射作用愈强，电阻率愈高。图 9-5 所示铜电阻率随温度变化的实验结果与式(9-8)和式(9-22)所表示的经典和量子电导理论结果相一致。

合金是固态溶体，简称固溶体。如图 9-5 所示，当纯金属铜中掺入镍形成铜镍合金时，电阻率随溶质镍浓度的升高而升高。铜镍合金电阻率随温度的变化仍为线性关系，并其斜率保持为纯铜的值。可见，镍的掺入对晶格振动的散射作用几乎没有影响。但由于增加了材料的不完整性，对材料的剩余电阻率却产生了显著的影响。首先，因原子大小不同，从而改变了晶格参数，使得散射作用增强；又因为原子价不同，改变了晶体中的电荷分布，也增加了对电子的散射。此外，由于原子价的不同，亦会改变电子浓度，从而影响费米能，继而影响 $N(E_F)$ 的值。图 9-6(a) 给出了含有不同价溶质的银合金电阻率的实验结果，数据表明，在溶质浓度较低（例如，1% 原子）的情况下，合金电阻率的变化 $\Delta\rho$ （相对于纯银的电阻率）按照平方关系随

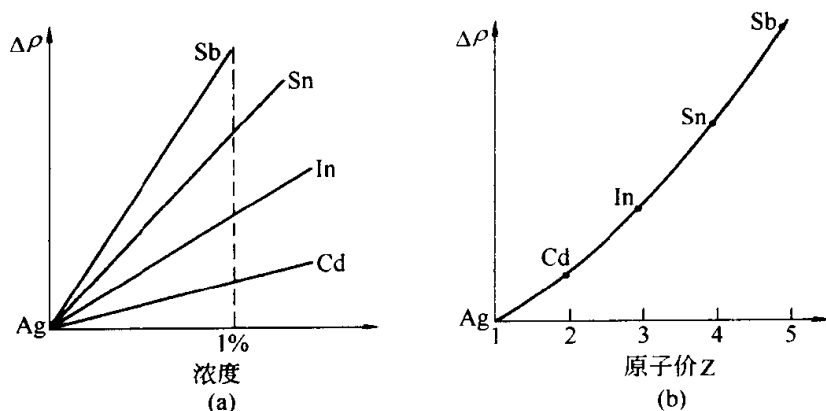


图 9-6 不同溶质的银合金的电阻率

(a) $\Delta\rho$ 与溶质浓度的关系；(b) $\Delta\rho$ 与溶质原子价的关系

溶质与溶剂间原子价差的增加而增加，如图 9-6(b) 所示。

在合金(固溶体)中，溶质原子通常是无序分布的，因此，散射将随溶质原子数目的增加而增强。倘若溶质原子呈周期性排列，例如对于 A、B 原子比例为 1:1 的合金，如果 A、B 原子交替地占据格点位置，材料的电阻率将会大大下降。这样的晶格周期性，即长程有序性，只有在 Cu_3Au 、 CuAu 和 Au_3Mn 等合金材料中观察到，图 9-7 给出了

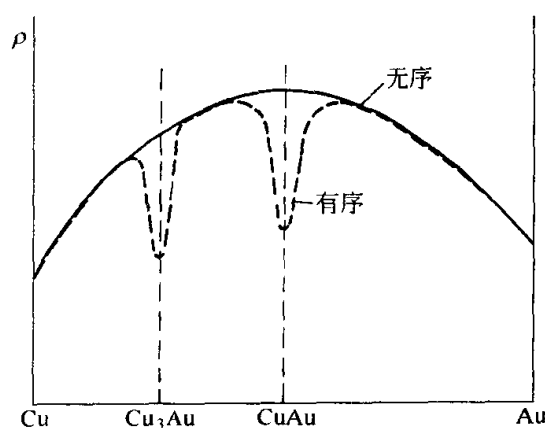


图 9-7 有序和无序铜-金合金的电阻率与成分比的关系

金、铜合金的情况。对于这样的合金材料，只有在一定的温度下才能发生无序状态与有序状态之间的转变，此温度称之为转变温度 $T_{\text{转变}}$ 。对于 Cu_3Au ， $T_{\text{转变}}$ 为 395°C 。在略低于 $T_{\text{转变}}$ 的温度下退火或从高于 $T_{\text{转变}}$ 的温度缓慢冷却，可出现上述长程有序状态。相反，如果在高于 $T_{\text{转变}}$ 的温度下退火，则会使该有序性消失。有些合金，如 CuAu ，显示出很强的长程有序性，甚至能一直保持到接近熔点。对于具有长程有序倾向的合金材料，如要使其在室温下呈现无序状态，只需将其从略高于 $T_{\text{转变}}$ 的温度突然放入冰盐水中淬火。

9.1.4 热电效应^[3,4]

热电效应（有时也称为温差电效应）首先是在如图 9-8 所示的装置中被发现的，其中，BC 为导体 I，AB 和 CD 为另一种导体 II，两种导体在 B 和 C 处联接，如此构成一热电偶。如果 B 和 C 这两个接点处的温度不同，则 A 和 D 之间会出现电位差，即温差电动势，该效应称之为塞贝克效应（Seebeck effect）。A、D 间的电位差 ΔV 取决于 B、C 间的温度差 $\Delta T = T - T_0$ ，与 A、D 处的温度无关。并有温差电动势

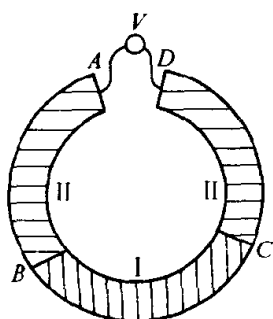


图 9-8 观察塞贝克效应和珀尔帖效应的装置

效应（Seebeck effect）。A、D 间的电位差 ΔV 取决于 B、C 间的温度差 $\Delta T = T - T_0$ ，与 A、D 处的温度无关。并有温差电动势

$$\Delta V = a(T - T_0) + b(T - T_0)^2 \quad (9-26)$$

$$\Theta = \frac{\Delta V}{\Delta T} = a + b(T - T_0) \quad (9-27)$$

其中， Θ 为温差电动势率（或微分温差电动势）。对于由某些材料构成的热电偶， b 值基本上为零。将热电偶的一端 B（或 C）放在某固定的参考温度（例如 0°C 的冰水）中，热电偶的另一端 C（或 B）放在被测

温度处，通过测量温差电动势 ΔV 可以准确迅速地测得温度，测量灵敏度高，且价格低廉。铜-康铜（55%铜和 45%镍的合金）是常用的热电偶，可在 $-180 \sim 400^\circ\text{C}$ 温度范围内使用。

珀尔帖效应（Peltier effect）是另一种热电效应，它正与塞贝克效应相反，当电流流过如图 9-8 所示由两种导电体组成的 $\text{II}/\text{I}/\text{II}$ 结构时，在 I 和 II 的两个接点 B、C 处，一处温度降低（放热），另一处温度升高（吸热）。对此现象可作如下分析，如果导电体 I 的功函数 Φ_{I} 大于导电体 II 的功函数 Φ_{II} ，由于功函数是费米能级与真空能级之间的能量差，则导电体 I 的费米能级 E_{FI} 低于导电体 II 的费米能级 E_{FII} 。假设电流由 A 流向 D，则电子由 D 流向 A。当电子由导电体 II 经过联接点 C 流向导电体 I 时，电子由高费米能级流向低费米能级，电子能量降低，向周围放出能量（放热）；接着，当电子由导电体 I 经过联接点 B 流向导电体 II 时，电子由低费米能级流向高费米能级，电子能量升高，从周围吸收能量（吸热）。如果电流方向相反，则变成 C 处吸热，B 处放热。利用珀尔帖效应可实现致冷和致热。

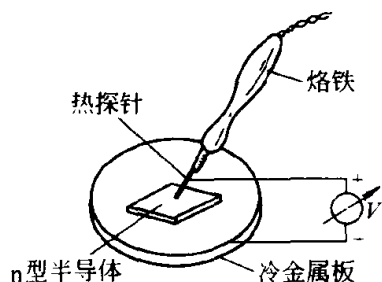


图 9-9 热探针法测定
导电类型

图 9-8 中的导电体 I 和导电体 II 可以是金属，也可以是半导体，即在金属和半导体中都能观察到热电效应，包括塞贝克效应和珀尔帖效应。利用半导体的热电效应可以方便地测定半导体材料的导电类型，其测量装置如图 9-9 所示。当半导体材料为 n 型时，多数载流子为电子（见 9.2.2 节），热探针（上电极）附近的电子热运动较强，向下扩散，使热探针附近缺少电子，而下电极附近积累电子，从而形成

极性如图 9-9 所示的电动势。当半导体材料为 p 型时，多数载流子为空穴，热探针附近的空穴热运动较强，向下扩散，使热探针附近缺少空穴，而下电极附近积累空穴，从而形成极性相反的电动势。如此便可以通过所产生电动势的极性来判断半导体材料的型号。从上述分析可以看出，由电子和空穴产生的温差电动势率的符号是相反的。

9.2 半导体的输运性质^[4]

半导体材料的室温电导率在 $10^{-8} \sim 10^3 (\Omega \cdot \text{cm})^{-1}$ 之间。半导体材料的另一重要特征是其对光照、温度、电场、磁场、压力等外界作用的敏感性，以及对自身成分和结构的灵敏性。半导体材料对外界作用的敏感性具体表现在诸如光电、热电、压电、光致发光、电致发光等现象和效应中。例如，光照可以在半导体中产生电子-空穴对，使电阻率下降。半导体材料性质对自

身成分和结构的灵敏性使得人们可以通过一定的手段来改变材料的成分和结构以控制材料的性质。例如，通过掺入施主杂质或受主杂质可以向半导体材料引入电子或空穴。电子和空穴这两种具有相反电荷的载流子在电场、磁场、温度梯度场、浓度梯度场等的作用下将发生运动。半导体中的电子和空穴在运输的过程中可相遇复合而放出能量，例如发出光子。半导体中载流子如此丰富多彩的运输现象及其相关效应使得半导体材料可以用来制作具有各种功能的器件。

9.2.1 半导体的能带

硅是目前应用得最多的半导体材料。在硅原子的外层电子态结构中有两个 $3s$ 态和六个 $3p$ 态，随着原子间距离的缩短，杂化了的 s 、 p 态 (sp^3 态) 能带发生交叠，且又分裂成两个能带，如图 9-10 所示。每个硅原子有四个价电子，这些价电子正好填满下面一个能带，该能带称为价带 (valence band) 或满带；上面的一个能带完全空着，称为导带 (conduction band)。价带与导带由禁带隔开，其能隙 (禁带宽度) 为 E_g 。晶体中原子的排列具有周期性，其中电子的状态用 k 来表征。晶体的结构是各向异性的，因此，在 k 空间中不同方向的能带有不同的形状。图 9-11 给出的是沿 k_x 方向的能带结构，其中 Γ 点和 X 点分别为 k 空间的原点和 k_x 方向的布里渊区边界。

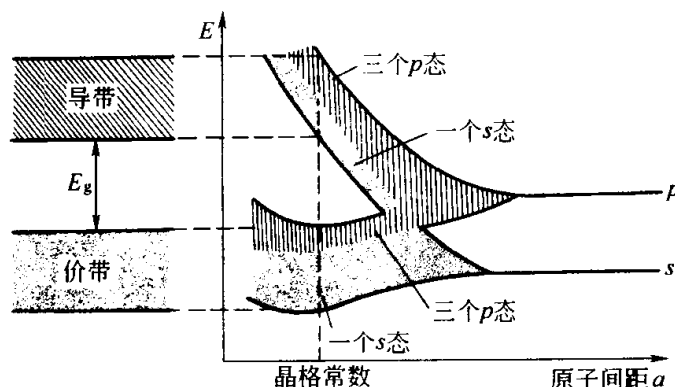


图 9-10 随着原子间距的缩小能级展宽成带，
继而发生交叠和分裂 (共价元素晶体)

锗晶体的情况与此基本相同，只需将 $3s$ 换成 $4s$ ，将 $3p$ 换成 $4p$ 。所有共价元素晶体都有类似的 sp^3 杂化能带。然而，禁带宽度却有所不同，表 9-1 给出了由 IV 族元素构成的晶体材料的禁带宽度 E_g 值。禁带宽度是材料的重要物理参数，对材料的性质有重大影响。金刚石的禁带宽度为 5.48eV ，是绝缘体；硅和锗的禁带宽度分别为 1.12eV 和 0.67eV ，都属半导体；灰锡的禁带宽度为 0.08eV ，呈金属性。

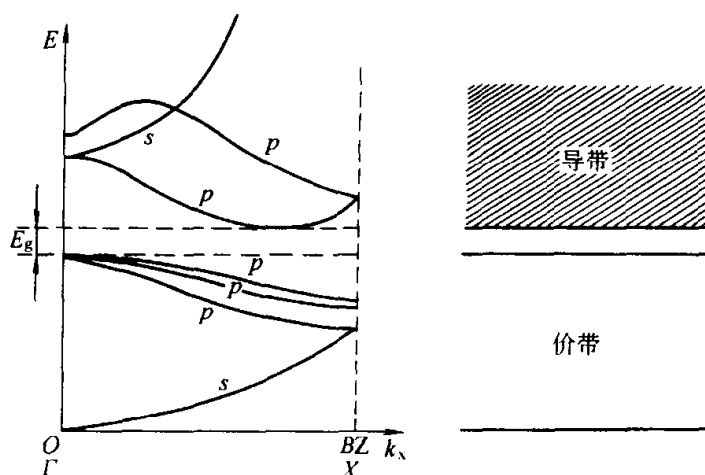
图 9-11 k 空间中 k_x 方向硅的能带结构

表 9-1 一些共价元素晶体的禁带宽度 (0K)

元 素	E_g/eV	元 素	E_g/eV
C (金刚石)	5.48	Ge	0.67
Si	1.12	Sn (灰锡)	0.08

9.2.2 载流子和费米能级

严格地说，只有在绝对零度时，硅和锗中的价带才是被填满的，导带才是全空的，此时材料的电导率才为零。随着温度的升高，价带中的部分电子跃迁到导带，并在价带中留下等量的空穴。电子从价带向导带的跃迁称之为带间跃迁。导带中的电子和价带中的空穴具有相反的电荷，在电场的作用下沿着相反的方向运动，它们都能导电，都是载流子。这种由本征热激发产生的

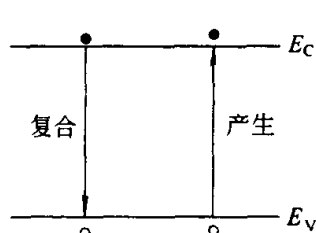


图 9-12 载流子的产生与复合

的载流子称为本征载流子。本征载流子一方面不断地因热激发而成对产生，另一方面也不断地因复合而成对消失，如图 9-12 所示，如此，在一定温度下保持一平衡的载流子浓度。室温下半导体中本征载流子浓度 n_i 很低，在硅中 n_i 只有 $10^{10}/\text{cm}^3$ ，其本征电导率很低。要提高半导体在室温下的电导率需要进行掺杂，掺入了杂质的半导体称为掺杂半导体

(doped semiconductor)。图 9-13 为硅晶格的二维示意图，每个硅原子有四个价电子与相邻的四个硅原子形成共价键。如在硅晶格中掺入少量的五价原子(如磷原子)以取代硅原子，见图 9-13，磷原子有五个价电子，除了用四个价电子与相邻的硅原子形成共价键，还剩余一个价电子，该价电子只要具有很小的能量(离化能)就能摆脱磷原子的轻微束缚，由施主能级 E_D 进入导带 [如图 9-14(a) 所示]，在硅晶格中自由运动。因此，硅中的杂质磷原子

可以提供导电电子，被称为施主杂质。掺有施主杂质的半导体为 n 型半导体。与此相反，如果将三价原子（如硼原子）掺入硅中，便可提供载流子空穴，此为受主杂质，相应的杂质能级称之为受主能级，如图 9-14(b) 中的 E_A 所示。掺有受主杂质的半导体为 p 型半导体。

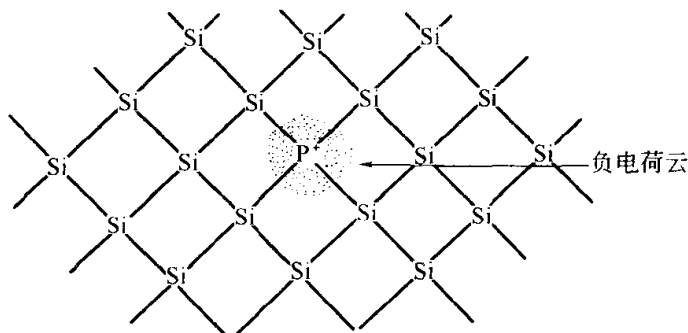


图 9-13 掺有磷原子的硅晶格的二维示意图

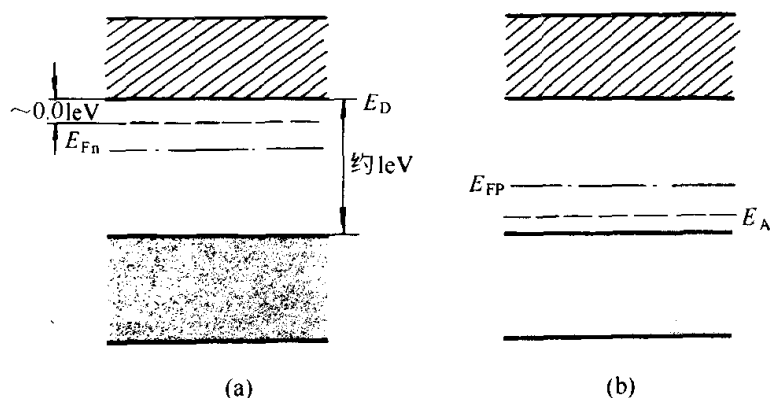


图 9-14 掺杂半导体的能带图

(a) n 型半导体；(b) p 型半导体

半导体中的平衡电子浓度 n_0 可由下式计算得

$$n_0 = \int_{E_C}^{\infty} f(E) g_C(E) dE \quad (9-28)$$

其中， $f(E)$ 是电子的统计分布函数，它给出能量为 E 的电子态上电子占据的几率； $g_C(E)$ 为导带底附近的状态密度，即导带底附近单位能量间隔的量子态数； E_C 为导带底的能量。在一般情况下， $f(E)$ 为费米分布

$$f(E) = \frac{1}{\left[1 + \exp\left(\frac{E - E_F}{k_B T}\right) \right]} \quad (9-29)$$

其中， k_B 为玻耳兹曼常数； E_F 为费米能级，该参数描述了电子的填充水平，能量为 E_F 的电子态的占据几率总是 1/2。对于非简并半导体，费米能

级离开导带较远, $E - E_F \gg k_B T$, 对此, 费米分布退化为玻耳兹曼分布

$$f(E) = \exp\left[\frac{-(E - E_F)}{k_B T}\right] \quad (9-30)$$

由于 $f(E)$ 随能量 E 的升高而指数下降, 所以式 (9-28) 的积分只需考虑导带底附近的贡献, 并且, 也正因为如此, 式 (9-28) 中的积分上限原应是导带顶, 现可改为 ∞ 。根据自由电子近似可以得到

$$g_C(E) = \left(\frac{4\pi}{h^3}\right)(2m_n^*)^{3/2}(E - E_C)^{1/2} \quad (9-31)$$

式中, $h = 2\pi\hbar$, m_n^* 为电子的有效质量, 其值见表 9-2。

表 9-2 室温下锗、硅和砷化镓的一些参数值

各项参数	E_g /eV	m_n^*	m_p^*	N_C /cm ⁻³	N_V /cm ⁻³	n_i /cm ⁻³ (计算值)	n_i /cm ⁻³ (测量值)
Ge	0.67	$0.56m_0$	$0.37m_0$	1.05×10^{19}	5.7×10^{18}	2.0×10^{13}	2.4×10^{13}
Si	1.12	$1.08m_0$	$0.59m_0$	2.8×10^{19}	1.1×10^{19}	7.8×10^9	1.5×10^{10}
GaAs	1.428	$0.068m_0$	$0.47m_0$	4.5×10^{17}	8.1×10^{18}	2.3×10^6	1.1×10^7

将式(9-30)和式(9-31)代入式(9-28)可得平衡时电子浓度

$$n_0 = N_C \exp\left[\frac{-(E_C - E_F)}{k_B T}\right] \quad (9-32)$$

这里, N_C 为导带底的有效状态密度

$$N_C = \left(\frac{2}{h^3}\right)(2\pi m_n^* k_B T)^{3/2} \quad (9-33)$$

同理, 平衡时的空穴浓度 p_0 应为

$$p_0 = \int_{-\infty}^{E_V} [1 - f(E)] g_V(E) dE \quad (9-34)$$

其中, E_V 为价带顶的能量, $g_V(E)$ 为价带顶附近的状态密度

$$g_V(E) = \frac{4\pi}{h^3}(2m_p^*)^{3/2}(E_V - E)^{1/2} \quad (9-35)$$

按照同样的方法可得

$$p_0 = N_V \exp\left[\frac{-(E_F - E_V)}{k_B T}\right] \quad (9-36)$$

这里, N_V 为价带顶的有效状态密度

$$N_V = \left(\frac{2}{h^3}\right)(2\pi m_p^* k_B T)^{3/2} \quad (9-37)$$

其中, m_p^* 为空穴的有效质量。对于本征半导体, 导带的电子数等于价带的空穴数, 即

$$n_i = p_i \quad (9-38)$$

$$n_i = N_C \exp \left[\frac{-(E_C - E_{Fi})}{k_B T} \right] \quad (9-39)$$

$$p_i = N_V \exp \left[\frac{-(E_{Fi} - E_V)}{k_B T} \right] \quad (9-40)$$

式中, E_{Fi} 为半导体的本征费米能级

$$E_{Fi} = \frac{E_C + E_V}{2} + \left(\frac{k_B T}{2} \right) \ln \left(\frac{N_V}{N_C} \right) \quad (9-41)$$

室温下 N_C 和 N_V 的值如表 9-2 所示, 因室温下 $k_B T$ 约为 0.026eV, 故式 (9-41) 简化为

$$E_{Fi} \approx \frac{E_C + E_V}{2} \quad (9-42)$$

即室温下本征费米能级近似位于禁带中央, 将式 (9-41) 代入式 (9-39) 和式 (9-40) 可得

$$n_i = p_i = (N_C N_V)^{1/2} \exp \left(\frac{-E_g}{2k_B T} \right) \quad (9-43)$$

由此可见, 本征半导体的载流子浓度 n_i 和 p_i 将随禁带宽度 E_g 的增加而指数下降。当材料的禁带宽度大到一定程度就成为绝缘体, 例如金刚石的禁带宽度为 5.48eV, 为绝缘体。由式 (9-43) 我们可以看到, 半导体中的本征载流子浓度随着温度升高而指数上升, 这便是半导体的热敏性。根据此制成的半导体热敏电阻可用于温度的自动测量和控制。对于 n 型半导体, 费米能级位于禁带上半部, $E_C - E_{Fn} < E_{Fn} - E_V$, 由式 (9-32) 和式 (9-36) 可知, 电子浓度 n_0 大于空穴浓度 p_0 , 因此, 电子为多数载流子, 简称多子; 空穴为少数载流子, 简称少子。对于 p 型半导体, 费米能级位于禁带下半部, $E_C - E_{Fp} > E_{Fp} - E_V$, 由式 (9-32) 和式 (9-36) 可知, 空穴浓度 p_0 大于电子浓度 n_0 , 则空穴为多数载流子, 简称多子; 电子为少数载流子, 简称少子。由式 (9-32) 和式 (9-36) 可得

$$n_0 p_0 = N_C N_V \exp \left(\frac{E_g}{k_B T} \right) \quad (9-44)$$

此式表明, 某一半导体材料的平衡电子和空穴浓度的乘积 $n_0 p_0$ 为一常数, 与掺杂情况无关, 它只是温度的函数。并且, 由式 (9-43) 可知, 此常数即为该温度下本征载流子浓度的平方

$$n_0 p_0 = n_i^2 \quad (9-45)$$

值得注意的是, 能带图中的能量坐标是按电子能量选取的, 坐标向上移动, 电子能量升高; 而空穴的能量正好与此相反, 坐标向下移动, 空穴能量

升高。

9.2.3 迁移率和散射

载流子在单位电场作用下的漂移速度称为载流子的迁移率 μ ，只取正值，它描述了载流子的导电能力，

$$\mu = \left| \frac{v}{E} \right| \quad (9-46)$$

其中， E 为电场强度， v 为载流子的漂移速度。半导体中有两种载流子，电子和空穴，因此其电导率 σ 为

$$\sigma = \frac{J}{|E|} = (n\mu_n + p\mu_p)q \quad (9-47)$$

式中， J 为电流密度， n 和 p 分别为电子和空穴的浓度， μ_n 和 μ_p 分别为电子和空穴的迁移率。这里， q 为电子电量，电子和空穴具有相反的电荷，分别为 $-q$ 和 q 。影响载流子迁移率的是各种散射作用，其中最重要的是晶格散射和杂质散射。温度愈高，晶格振动愈激烈，晶格散射愈强。因此，温度愈高，载流子迁移率愈低，如图 9-15(a) 所示。另一方面，如上节所述，本征半导体中的载流子浓度随着温度的升高按照指数规律迅速增加，如图 9-15(b) 所示。因此，本征半导体的电导率随温度的升高仍呈上升趋势，如图 9-15(c) 所示，这与金属中的情况截然相反。在低温（例如液氮温度）下，晶格振动较弱，晶格散射作用也较弱，因此，杂质散射作用显得突出。在 n 型或 p 型半导体中，离化的施主或受主起着重要的散射作用，是影响低温下迁移率的主要因素。

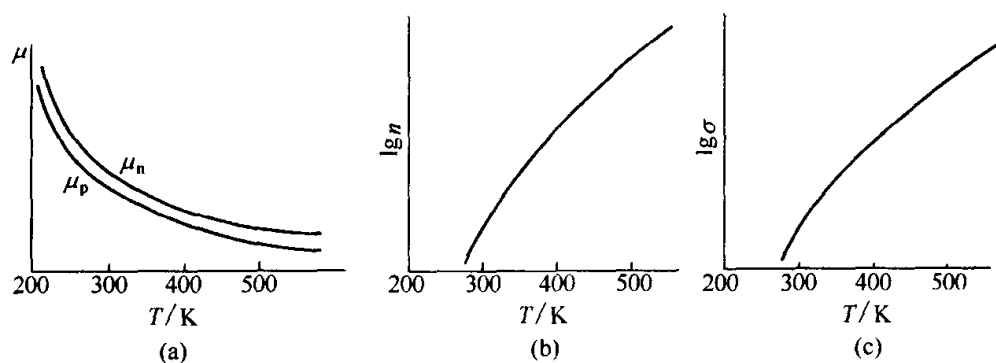


图 9-15 以下各量随温度的变化

(a) 电子和空穴迁移率；(b) 本征半导体载流子数；(c) 本征半导体电导率

掺杂半导体中，载流子由两部分组成，一部分由本征激发产生，另一部分由杂质离化提供，故不同掺杂浓度 N_d 的半导体材料显示出不同的 σ - T 曲线，如图 9-16 所示。施主和受主杂质的离化能较小，通常在较低温度下都已全部离化，因而在本征激发较弱的低温范围内，载流子主要由杂质离化提

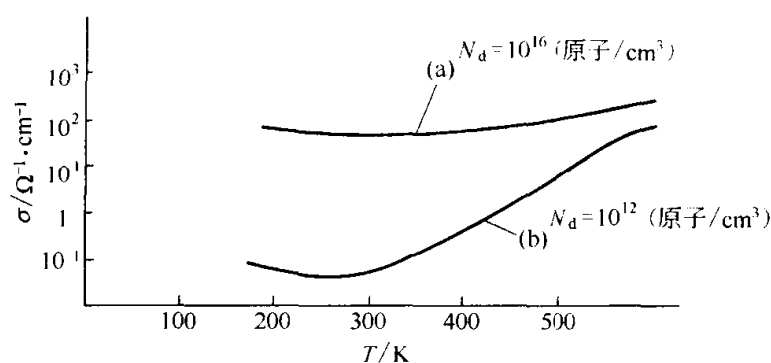


图 9-16 掺杂半导体的电导率随温的变化

(a) 高掺杂; (b) 低掺杂

供, 故掺杂浓度较高的样品具有较高的电导率。随着温度的升高, 晶格散射加强, 电导率一开始随温度升高有所下降; 可是当温度进一步升高时, 本征激发加强, 本征载流子浓度升高, 并使本征载流子浓度逐步超过掺杂引入的载流子浓度, 成为主要的载流子, 最终两样品的电导率随温度升高而上升趋于相同的值。

9.2.4 霍尔效应

将一半导体放在磁场中, 磁场沿 Z 方向, 磁感应强度为 B_z 。如在半导体中沿 x 方向通以电流, 电流密度为 J_x , 则沿 y 方向将产生一电场, 这称之为霍尔效应, 如图 9-17 所示。该 y 方向的电场称之为霍尔电场, 其电场强度为

$$E_y = R J_x B_z \quad (9-48)$$

其中比例系数 R 称为霍尔系数。因为 J_x 、 B_z 和 E_y 都是易测量的量, 所以

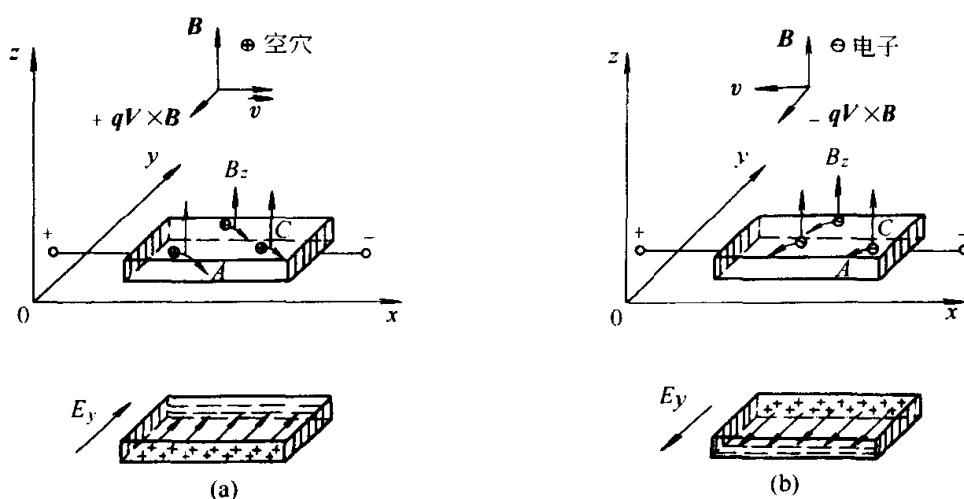


图 9-17 霍尔效应

(a) p 型半导体; (b) n 型半导体

霍尔系数 R 很容易利用下式由实验测得

$$R = \frac{E_y}{J_x B_z} \quad (9-49)$$

下面以 p 型半导体为例来讨论霍尔效应的物理实质。在 x 方向电场的作用下 p 型半导体中的空穴以漂移速度 v_x 沿 x 方向运动，由此产生的电流密度为

$$J_x = pqv_x \quad (9-50)$$

由于 z 方向有磁场存在，沿 x 方向运动的空穴受罗伦兹力的作用向 $-y$ 方向偏转，从而使样品的 $-y$ 端 (A 面) 积累正电荷， $+y$ 端 (C 面) 积累负电荷，如此电荷积累形成霍尔电场，见图 9-16(a)。在稳定的情况下，该霍尔电场对空穴的作用应与罗伦兹力相抵消

$$qE_y = qv_x B_z \quad (9-51)$$

因此

$$E_y = v_x B_z \quad (9-52)$$

将式(9-50)和式(9-52)代入式(9-49)可得霍尔系数

$$R_p = \frac{1}{pq} \quad (9-53)$$

由此可知，对于 p 型半导体，霍尔系数 R_p 为正。由于在 x 方向电场的作用下，电子漂移的方向与空穴相反，且电荷符号也相反，故在磁场中受罗伦兹力的作用，电子和空穴都偏向 $-y$ 侧。因此在相同条件下，n 型半导体与 p 型半导体中产生的霍尔电场的方向相反，见图 9-16(b)。由式 (9-49) 可知，霍尔系数 R 的符号亦相反。对于 n 型半导体， R_n 为负

$$R_n = \frac{-1}{nq} \quad (9-54)$$

由式 (9-48) 可知，霍尔电场强度 E_y 正比于外加磁场的磁感应强度 B_z ，而霍尔电场强度 E_y 又正比于降落在样品 y 方向两端的霍尔电压 V_y ， V_y 是很容易精确测量的量，因此利用霍尔效应可制成磁场计，可用来测量磁感应强度。另外，导电类型和载流子浓度是半导体材料的重要信息，由式 (9-53) 和式 (9-54) 可知，通过霍尔效应的测量，一方面由霍尔系数的符号可以判断半导体的导电类型；另一方面由霍尔系数的绝对值可以确定载流子浓度 n 或 p 。对于硅材料，通过霍尔效应可测量的最低杂质浓度为 10^{12} cm^{-3} 量级，因为硅晶体的原子浓度为 $10^{22} / \text{cm}^3$ 量级，其杂质浓度测量的相对精度可达 10^{-10} ，如此高的精度是任何一种化学分析方法无法与之相比的。另一方面，根据霍尔效应制成的霍尔器件还可用来制作非接触开关和传感器等，这在计算机和自动控制系统中被广泛应用。此外，在强磁场、极低温的条件下，在半导体界面二维体系中，人们观察到整数和分数量子霍尔效应，这对深入认识量子化现象起了很大的推动作用。

9.2.5 p-n 结^[5]

当 p 型半导体和 n 型半导体接触在一起时便形成 p-n 结。由于在接触前 n 型硅的费米能级 E_{Fn} 比 p 型硅的费米能级 E_{Fp} 高, 如图 9-18(a) 所示, 一旦接触在一起, 电子将由 n 型区流向 p 型区 (这种载流子的流动是由载流子的浓度差引起的, 是载流子的扩散运动), 并在 n 型区表面层留下带正电的离化施主; 同时, 空穴将由 p 型区流向 n 型区, 并在 p 型区的表面层留下带负电的离化受主。因此, 在 p-n 结界面的两侧形成空间电荷层, n 型硅一侧为带正电的空间电荷层, p 型硅一侧为带负电的空间电荷层, 如图 9-18(b) 所示。如此空间电荷层中有电场存在, 电场的方向是由 n 型区指向 p 型区。该电场不是外加的, 而是自身产生的, 故称之为自建电场。由此电场引起的载流子的漂移运动的方向正与上述扩散运动的方向相反, 当由载流子漂移运动引起的漂移电流与由载流子的扩散运动引起的扩散电流相等时, p-n 结两边达到平衡, 有相同的费米能级 E_F , p-n 结中既无净

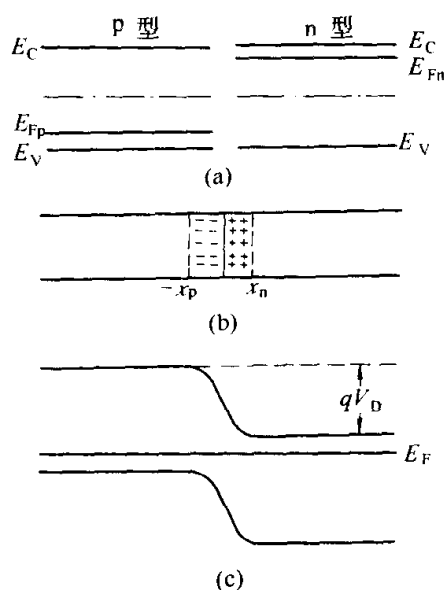


图 9-18 p-n 结的空间电荷区和势垒
(a) 接触前的能带图; (b) 和 (c) 接触形成的空间电荷区和势垒

的空穴流, 也无净的电子流。由于 p-n 结空间电荷区有静电场存在, 这就使能带发生弯曲, 在 p-n 结两边形成电势差 V_D , 此被称为 p-n 结接触电势差。如图 9-18(c) 所示, 该接触电势差在 p-n 结处形成势垒, 阻挡着 n 型区的电子 (多子) 向 p 型区的运动, 也阻挡着 p 型区的空穴 (多子) 向 n 型区的运动。接触电势差 V_D 对应于能带的弯曲量 qV_D , 此能带的弯曲量应等于 n 型硅与 p 型硅在接触前的费米能级之差。

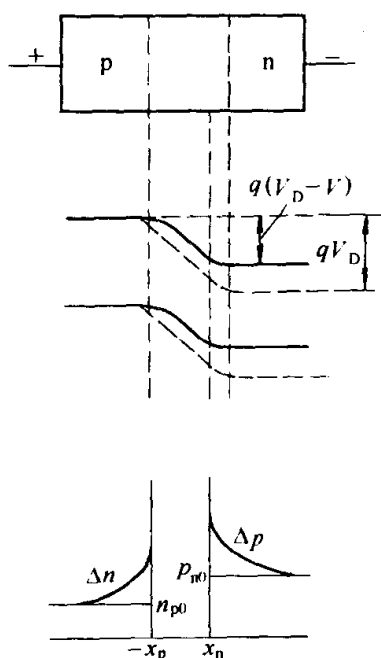


图 9-19 正向偏置 p-n 结的能带图和少子注入

的空穴流, 也无净的电子流。由于 p-n 结空间电荷区有静电场存在, 这就使能带发生弯曲, 在 p-n 结两边形成电势差 V_D , 此被称为 p-n 结接触电势差。如图 9-18(c) 所示, 该接触电势差在 p-n 结处形成势垒, 阻挡着 n 型区的电子 (多子) 向 p 型区的运动, 也阻挡着 p 型区的空穴 (多子) 向 n 型区的运动。接触电势差 V_D 对应于能带的弯曲量 qV_D , 此能带的弯曲量应等于 n 型硅与 p 型硅在接触前的费米能级之差。

$$qV_D = E_{Fn} - E_{Fp} \quad (9-55)$$

由于空间电荷层有强电场存在, 其中的电子和空穴都被电场扫向两边而耗尽, 因此空间电荷层被视为载流子的耗尽层, 其电阻率远大于 p 型和 n 型区的体内。所以, 当接上外加偏压时, 外加电压 V 基本上全部降落在空间电荷层上。如果外加偏压的正端接 p 型区一边, 负

端接 n 型区一边, 则相对应于 p 型区而言, n 型区的电势下降, 电子能量升高, 能带上移。如此外加偏压的作用使 p-n 的势垒降低, 由 qV_D 降为 $q(V_D - V)$, 能带弯曲减小; p-n 结界面处空间电荷区 (势垒区) 宽度变窄, 电场强度减小, 见图 9-19; 载流子的扩散电流大于漂移电流, 净电流由 p 型

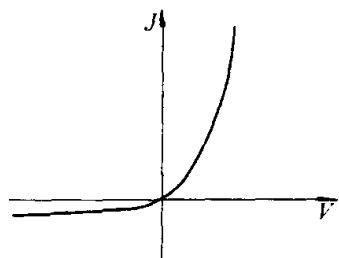


图 9-20 理想 p-n 结的伏-安特性

区流向 n 型区。并且, 该电流随着偏压的升高而迅速上升, 如图 9-20 所示。此时的偏压称为正向偏压, 此时的电流称为正向电流。在此情况下, p 型区的空穴注入 n 型区, 形成附加的空穴浓度 Δp ; n 型区的电子注入 p 型区, 形成附加的电子浓度 Δn , 如图 9-19 所示。 Δp 和 Δn 都是非平衡载流子, 它们是在正向偏压下电注入产生的。由于注入的都是少数, 故称之为正向少数注入。

如果外加偏压的方向相反, 正端接 n 型区, 负端接 p 型区, 则相对于 p 型区而言, n 型区的电势上升, 电子势能下降, 能带下移。如此外加偏压的作用使 p-n 结的势垒升高, 由 qV_D 变为 $q(V_D - V) = q(V_D + |V|)$, 能带弯曲加大, p-n 结空间电荷区宽度变宽见图 9-21, 电场强度增加, 载流子的漂移电流大于扩散电流, 净电流由 n 型区流向 p 型区。由于空间电荷区的电场方向是由 n 型区指向 p 型区, 所以该电流是在电场作用下, p 型区的电子 (少数) 被扫向 n 型区, 同时, n 型区的空穴 (少数) 被扫向 p 型区。因为少数浓度很低, 所以电流很小, 此时的偏压称为反向偏压, 此时的电流称为反向电流。随着反向电压的升高, 空间电荷区的电场强度加大, n 型区和 p 型区的少数一旦扩散到空间电荷区的边界就被电场扫向 p-n 结的对面, 致使 p-n 结的反向电流最终达到饱和, 如图 9-21 下面所示。不难得出, 理想 p-n 结的电流-电压关系为

$$J = J_S \left[\exp\left(\frac{qV}{k_B T}\right) - 1 \right] \quad (9-56)$$

其中 J_S 被称为反向饱和电流

$$J_S = q \left(\frac{D_n n_{p0}}{L_n} + \frac{D_p p_{n0}}{L_p} \right) \quad (9-57)$$

式中, D_n 和 D_p 为 p 型区电子和 n 型区空穴的扩散系数, n_{p0} 和 p_{n0} 为平衡时 p 型区电子和 n 型区空穴的浓度, L_n 和 L_p 为 p 型区电子和 n 型区空穴的扩散长度。由图 9-20 可知, 在一定的偏压范围内, p-n 结在绝对值相同的偏压下的正反向电流比很大, 这称之为 p-n 结的整流效应。其正反

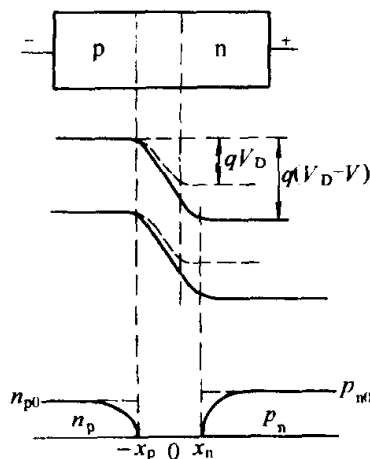


图 9-21 反向偏置 p-n 结的能带图和少数分布

向电流之比称为整流比。如果反向偏压进一步加大，到一定程度，p-n 结会发生击穿，反向电流将突然猛增，如图 9-22 所示，其中 V_{BR} 为击穿电压。

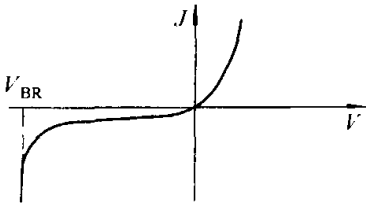


图 9-22 p-n 结击穿

p-n 结是电子器件和电路的重要基本单元，例如，利用 p-n 结的整流效应可以制成整流管；利用 p-n 结的反向击穿特性可以制成稳压管；利用 p-n 结的光电效应可以制成光电管和太阳能电池；此外，将两个 p-n 结背靠背紧紧地耦合在一起，从而组成 p-n-p 或 n-p-n 晶体管，可以实现放大、开关等功能，如此等等。

9.2.6 金属与半导体接触

不但两种不同型号的半导体接触在一起具有整流作用，金属与半导体接触在一起也可具有整流作用。事实上，早在 p-n 结被发明之前，人们就已发现氧化铜矿石与金属接触具有整流作用，可用来作为检波器，并用其制成了早期的半导体收音机——矿石收音机。

并非任何金属与半导体接触都具有整流效应，这依赖于材料的功函数，如前所述材料的功函数 Φ 为材料的费米能级与真空能级之差。假设某金属的功函数 Φ_M 大于某 n 型半导体的功函数 Φ_S ，如图 9-23(a) 所示，若将两者放在一起彼此接触，由于半导体的功函数较小，费米能级较高，则电子将由半导体流向金属，并在半导体表面留下带正电的离化施主，从而形成一层带正电的空间电荷层。金属表面的负电荷和半导体表面的正电荷在空间电荷层中形成电场，电场方向由半导体指向金属。空间电荷层的电势降使得半导体表面层的能带由外向内向下弯曲。当能带弯曲到使半导体的费米能级与金属的费米能级相平时，由半导体向金属的电子流动停止，系统达到平衡，如图

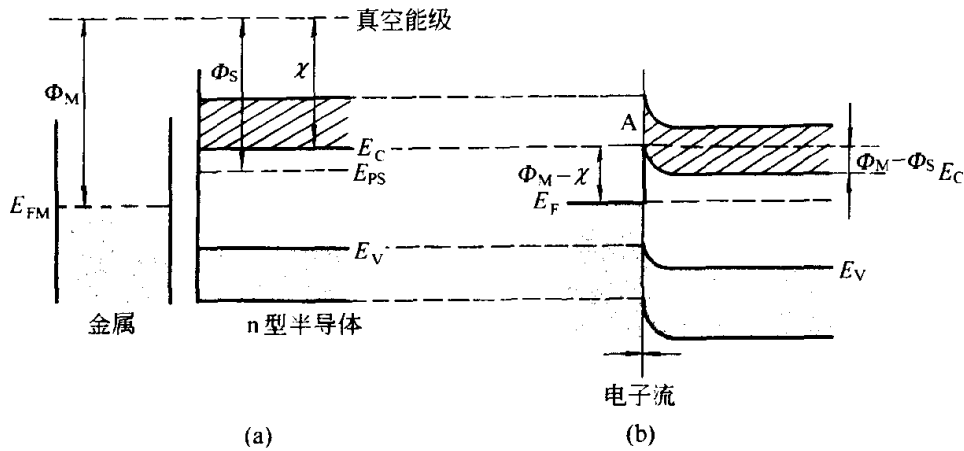


图 9-23 金属与 n 型半导体的整流接触
(a) 接触前；(b) 接触后

9-23 (b) 所示。在半导体一侧，表面能带高于体内，形成一表面势垒。表面势垒高度 qV_S 等于接触前金属与半导体的费米能级之差。

$$qV_S = E_{FS} - E_{FM} = \Phi_M - \Phi_S \quad (9-58)$$

该电势差 V_S 被称为金属与半导体接触的接触电势差。在金属一侧，表面势垒的高度为

$$qV_M = \Phi_M - \chi \quad (9-59)$$

其中 χ 为半导体的亲和能，它是半导体导带底与真空电子能级之间的能量差。虽然金属的电子浓度大于半导体的电子浓度，但由于金属表面的势垒高度大于半导体表面的势垒高度，结果使两侧间的电子流动达到平衡。

如果加上偏压，金属接负，n 型半导体接正，从而使金属一侧电子势能升高，费米能级 E_{FM} （在非平衡情况下的费米能级也称为准费米能级）上升；半导体体内电子势能降低，费米能级 E_{FS} 下降，半导体表面能带弯曲加强，表面势垒的高度和宽度增加，如图 9-24(a) 所示。在如此反向偏压下，只有很小的电流流过。如果改变偏压的极性，金属接正，n 型半导体接负，金属的电子势能降低，费米能级 E_{FM} 下降；半导体体内电子势能升高，费米能级 E_{FS} 上升，其能带图如图 9-24(b) 所示。在如此正向偏压下，半导体表面势垒高度和宽度将随偏压的增加而减小，电流将随偏压的增加而迅速地加大。图 9-25 为其正反向偏压时的伏-安特性示意图。

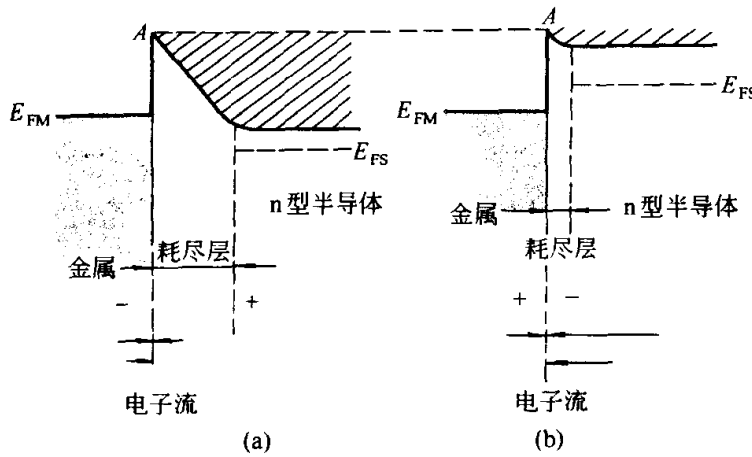


图 9-24 偏压下的金属-半导体接触

(a) 反向偏压；(b) 正向偏压

下面根据热电子发射理论讨论该金属-半导体 (M-S) 整流接触的伏-安特性。由于金属一侧界面势垒的高度是固定的，为 $\Phi_M - \chi$ ，因此，电子由金属流向半导体而产生的电流密度为

$$J_{MS} = -qC_{MS}T^2 \exp\left[\frac{-(\Phi_M - \chi)}{k_B T}\right] \quad (9-60)$$

其中, C_{MS} 为一系数。而半导体一侧界面势垒的高度 $\Phi_M - \Phi_S - qV$ 将随偏压 V 改变, 电子由半导体流向金属而产生的电流密度为

$$J_{SM} = qC_{SM}T^2 \exp\left[\frac{-(\Phi_M - \Phi_S - qV)}{k_B T}\right] \quad (9-61)$$

其中, C_{SM} 为另一系数。流过界面的总电流密度

$$J_{\text{总}} = J_{SM} + J_{MS} \quad (9-62)$$

由于 $V=0$ 时, $J_{\text{总}}=0$, 则 $J_{SM} = -J_{MS}$, 即

$$C_{MS} \exp\left[\frac{-(\Phi_M - \chi)}{k_B T}\right] = C_{SM} \exp\left[\frac{-(\Phi_M - \Phi_S)}{k_B T}\right] \quad (9-63)$$

由式(9-60)、式(9-61)、式(9-62) 和式(9-63) 可得

$$J_{\text{总}} = J'_S \left[\exp\left(\frac{qV}{k_B T}\right) - 1 \right] \quad (9-64)$$

其中

$$J'_S = CT^2 \exp\left[\frac{-(\Phi_M - \Phi_S)}{k_B T}\right] \quad (9-65)$$

这里, $C = qC_{SM}$ 。以上讨论的是金属与 n 型半导体接触, 且 $\Phi_M > \Phi_S$ 的情况。对于金属与 p 型半导体接触, 且 $\Phi_M < \Phi_S$ 的情况, 亦具有整流特性, 其分析基本同上, 只需将 n 型半导体的电子换成 p 型半导体的空穴即可。比较式(9-56)与式(9-64)以及图 9-20 与图 9-25 可知, M-S 结与 p-n 结具有相似的整流特性。但由于 M-S 结的输运过程只涉及一种类型的载流子, 不存在少子的积累和因复合而消失的过程, 故响应速度快, 可用作微波器件。并且由于结的一侧为金属, 具有好的散热性, 故适合在大功率下工作。

金属与半导体接触除可形成如上所述的整流接触外, 还可形成欧姆接

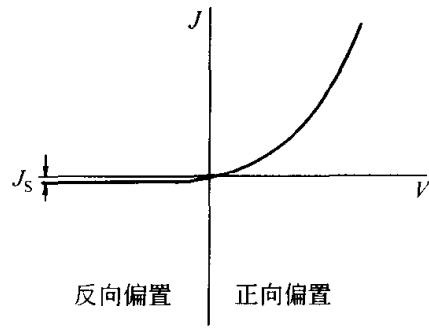


图 9-25 金属-半导体整流接触的伏-安特性

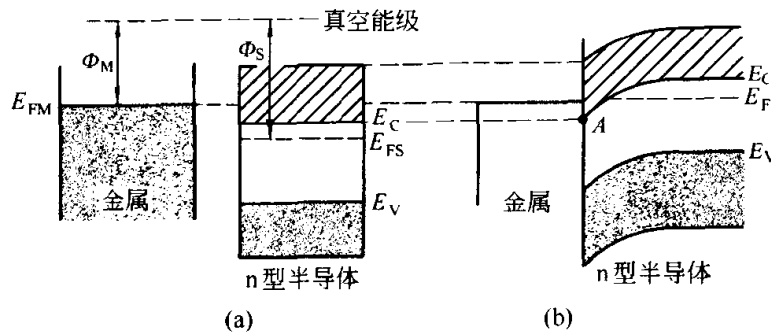


图 9-26 金属与 n 型半导体的欧姆接触

(a) 接触前; (b) 接触后

触,我们先讨论金属与 n 型半导体接触,且 $\Phi_M < \Phi_S$ 的情况,如图 9-26(a) 所示。当这两种材料接触在一起时,由于金属的费米能级较高,电子将从金属流向半导体,使金属表面缺少电子,带正电;而半导体表面积累电子,带负电,半导体体内电子势能升高,费米能级上升。当半导体和金属费米能级相平时,系统达到平衡,如图 9-26(b) 所示。此时半导体表面能带向下弯曲,表面形成高电导的电子积累层。金属与半导体界面没有势垒,没有载流子耗尽层。当加上偏压时,无论偏压的极性如何,电流都可以无阻挡地通过界面。如此金属与半导体接触称为欧姆接触。与此相似,当金属与 p 型半导体接触,如果 $\Phi_M > \Phi_S$,在半导体表面会形成高电导的空穴积累层,亦为欧姆接触。以上讨论的是不存在表面态时的理想情况。而 Ge、Si 和 GaAs 这些常用半导体材料的表面态密度很高,与金属接触都形成势垒,金属功函数的影响不大。在实际工作中,常通过重掺杂半导体与金属接触,其势垒宽度很薄,电子可以通过隧道效应穿过势垒,双向阻抗都很小,从而形成欧姆接触。通常的半导体元件都需要电极引线,双向低阻抗的欧姆接触正是作为电极引线所需要的。

9.2.7 光电导与光伏效应 (太阳能电池)

半导体中除热激发能产生电子空穴外,光激发也能产生电子空穴。当光子能量大于禁带宽度的光照射在半导体材料上时,价带电子吸收光子能量跃迁到导带,从而产生电子-空穴对。如此光生载流子是由光注入产生的非平衡载流子。光生载流子的产生使材料的电导率 σ 升高,

$$\sigma = \sigma_0 + \Delta\sigma = q(n\mu_n + p\mu_p) \quad (9-66)$$

$$n = n_0 + \Delta n, \quad p = p_0 + \Delta p \quad (9-67)$$

其中, σ_0 、 n_0 和 p_0 分别为无光照时半导体的电导率以及电子和空穴浓度, $\sigma_0 = q(n_0\mu_n + p_0\mu_p)$, Δn 和 Δp 为光照产生的非平衡电子和空穴浓度。光照引起的附加电导率

$$\Delta\sigma = q(\Delta n\mu_n + \Delta p\mu_p) \quad (9-68)$$

利用半导体的光电导效应制成的光敏电阻可作传感器用于自动控制等。

在稳定的光照下,光生电子和空穴以一定的平均速率不断产生,与此同时光生电子和空穴也以一定的平均速率不断地复合而消失,光生电子和空穴的产生率与复合率相等,光生电子和空穴浓度 Δn 和 Δp 保持恒定。按照统计规律,非平衡载流子(包括光注入和电注入非平衡载流子)具有一定的平均存在时间,即寿命。非平衡载流子寿命 τ 是半导体材料的一个重要参数。例如,对于用作光敏电阻的材料,寿命 τ 决定了当光照停止后半导体的电导恢复到平衡时电导所需要的时间,这直接影响到该光敏电阻作为传感器的

响应速度。

当光子能量大于禁带宽度的光照射到 p-n 结上时, 同样也会产生光生电子空穴。由于 p-n 结空间电荷层有自建电场存在, 在该电场的作用下, 光生电子被扫向 n 型区一边, 光生空穴被扫向 p 型区一边, 如此便产生光生电动势, 这被称为光生伏特效应, 简称光伏效应, 如图 9-27 (b) 所示。此光生电压的极性, p 型区为正, n 型区为负, 相当于对 p-n 结加了一个正向偏压,

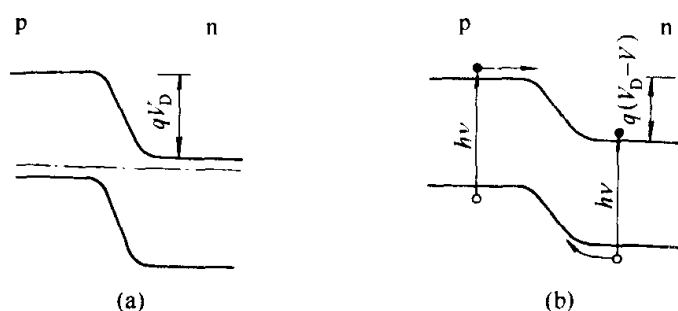


图 9-27 p-n 结光伏效应

(a) 无光照; (b) 光照激发

使 p-n 结势垒降低, 由 qV_D 降为 $q(V_D - V)$ 。在其他条件一定的情况下, 光生电动势的大小直接反映了入射光的强度, 由此可制成光度计, 例如照相机的曝光表。利用光生伏特效应可制作光电池, 根据此原理制成的太阳能电池已被广泛采用, 这是一种取之不尽、用之不竭的无污染能源。因为只有靠近势垒区在空穴扩散长度 L_p 范围内 n 型区中的光生空穴和在电子扩散长度 L_n 范围内 p 型区中的光生电子能扩散到势垒区, 所以光生电流密度

$$J_L = qS_{\text{平均}}(L_p + L_n) \quad (9-69)$$

其中, $S_{\text{平均}}$ 是扩散长度范围内光生载流子的平均产生率。

图 9-28 为实际太阳能电池的结构示意图和能带图, 为了提高光电转换效率, 将 p 型区设计得很薄, 只有 $1\mu\text{m}$ 。由单晶硅制成的太阳能电池的转换效率一般为 14%, 高的可达 20% ~ 24%, 但价格较高。而用多晶硅制成的太阳能电池的转换效率虽然只有 6% ~ 8%, 但因成本低、价格廉而更受欢迎。

9.2.8 晶体管^[6]

图 9-29(a) 为两个背靠背紧紧耦合在一起的 p-n 结, 它们之间的耦合是靠公共的 p 型区来实现的, 此称之为 n-p-n 晶体管。左右两个 n

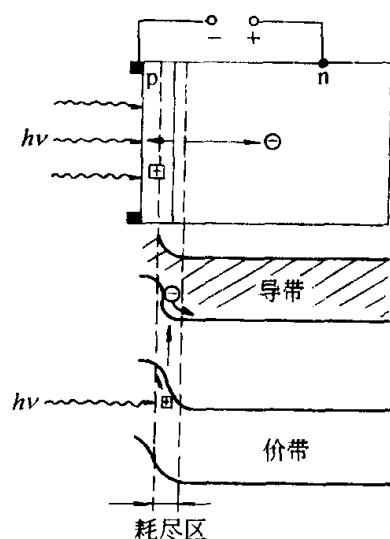


图 9-28 太阳能电池的结构示意图和能带图

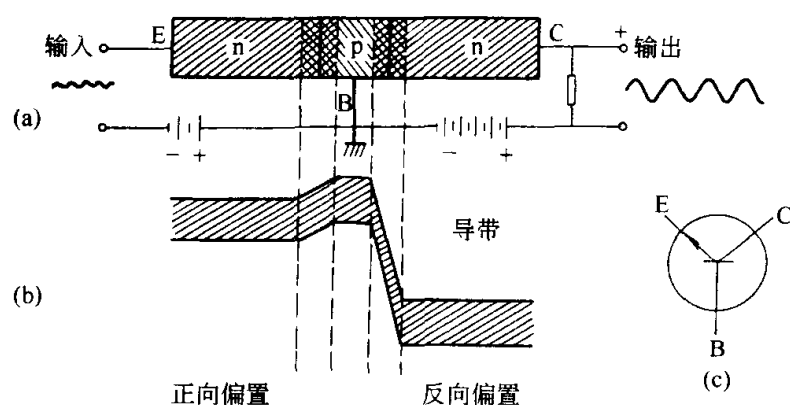


图 9-29 n-p-n 晶体管

(a) 结构示意图和偏置；(b) 能带图；(c) 符号

型区分别为发射区和集电区，中间的 p 型区为基区。三个引出电极 E、B、C 分别为发射极、基极和集电极。发射区与基区间的 p-n 结为发射结，基区与集电区间的 p-n 结为集电结。图 9-29(a) 给出了该晶体管工作时的一种偏置方式，B 极为公共地端（称之为共基极偏置方式），E 极接负电位，C 极接正电位，则发射结为正向偏置，集电结为反向偏置，此时的能带图如图 9-29(b) 所示。由于发射结为正向偏置，电子由发射区注入基区，从而在基区（p 型区）形成电子浓度梯度。电子因扩散运动进入集电结，并在反向偏置的集电结势垒区强电场的作用下被扫向集电区。在此过程中，除了因与空穴复合引起少量电子流损失外，绝大部分电子流由处于正向偏置的低阻发射结（输入端）流过处于反向偏置的高阻集电结（输出端），从而实现对输入电流信号的功率放大。该放大作用是靠输出端反向偏置集电结势垒电场对流入电子的漂移作用来实现的。在整个过程中，有两种载流子参与，除了上述电子的运动外，还伴随着空穴的运动（例如，正向偏置的发射结在电子由发射区注入基区的同时，也有空穴由基区注入发射区）故称之为双极型晶体管（bipolar transistor）。为了提高发射结注入电子流与空穴流之比（称之为注入比）和降低发射区和集电区的串联电阻，应提高发射区和集电区的电子浓度，降低其电阻率。为了减小基区（p 型区）中电子（少子）与空穴的复合而引起的损失，应提高基区的非平衡少子寿命，并尽可能减小基区的宽度。与上述 n-p-n 晶体管相似，可以制成 p-n-p 晶体管，其中，空穴流代替了上述电子流。

另一种类型的晶体管为场效应晶体管（field effect transistor-FET）。图 9-30(a) 给出的是具有金属（M）-氧化物（O）-半导体（S）结构的场效应晶体管（MOSFET）结构示意图。在金属栅与半导体衬底之间有一层起绝缘作用的氧化层，在栅电极两侧分别有源和漏两个电极。该二电极下面分别有

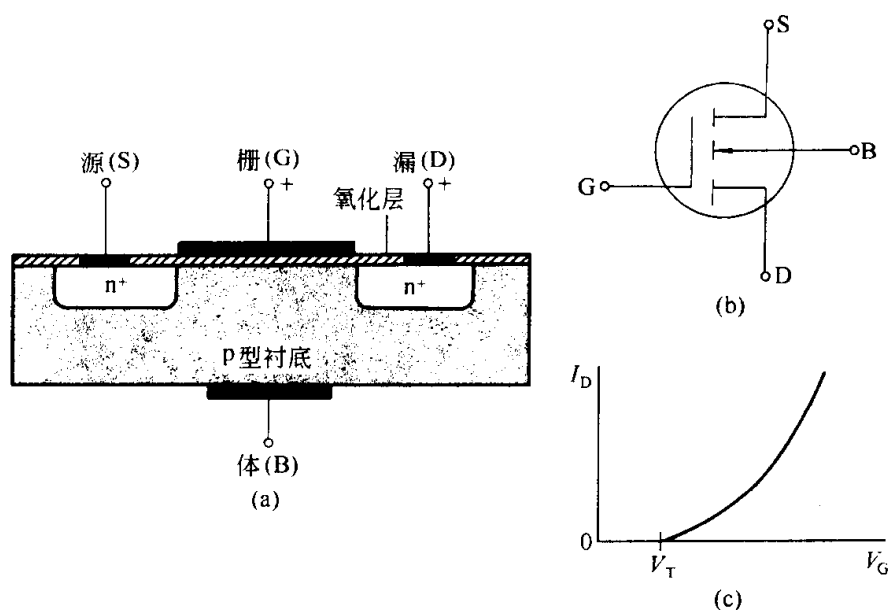


图 9-30 增强型 n 沟 MOSFET
(a) 结构示意图；(b) 符号；(c) I_D - V_G 曲线

与衬底掺杂类型 (p 型) 相反的重掺杂 (n^+ 型) 半导体区域, 分别称之为源区和漏区。当源漏间加上电压 V_{DS} 时, 如果栅与衬底间偏压 V_G 为零, n^+ 型的源漏之间有 p 型区相隔, 源漏间漏电流 I_D 为零。如果栅相对于衬底的偏压为正, 在紧靠氧化层的半导体衬底表面层中的空穴被排向体内, 使表面层带负电, 同时表面能带向下弯曲。随着栅压 V_G 的升高, 表面能带进一步下弯, 表面层的空穴逐步耗尽, 形成耗尽层。当栅压继续升高并达到某一值使表面开始强反型, 从而在表面形成一层 n 型导电沟道, 源漏间开始有一定强度的漏电流 I_D 通过, 此时的栅压 $V_G = V_T$ 称之为开启电压。此后, 漏电流将随栅压 V_G 的进一步升高而加大, 如图 9-30(c) 所示。这种通过改变栅压可以控制半导体表面电导率和导电类型的现象称之为场效应。对于上述的 MOSFET 器件, 当栅压为零时, 源漏间处于“关”态, 当栅压大于开启电压 V_T 时, 源漏间处于“开”态, 这种常态为“关”态、沟道为 n 型的 MOSFET 称之为增强型 n 沟 MOSFET。如果在 p 型半导体衬底表面与氧化层的交界处加一层 n 型层, 如图 9-31 所示, 这样在常态 (即栅压为零时) 源漏之间呈“开”态; 而当栅压 V_G 为某一负偏压时, 表面 n 层中的电子耗尽为高阻, 使源漏之间呈“关”态, 此种 MOSFET 称之为耗尽型 n 沟 MOSFET。与此相类似, 可设计出增强型 p 沟 MOSFET 和耗尽型 p 沟 MOSFET。p 沟 MOSFET 和 n 沟 MOSFET 可以结合在一起形成互补 MOSFET (互补 MOS 即 complementary MOS, 被称为 CMOS), 以实现低工作电压、低功耗。

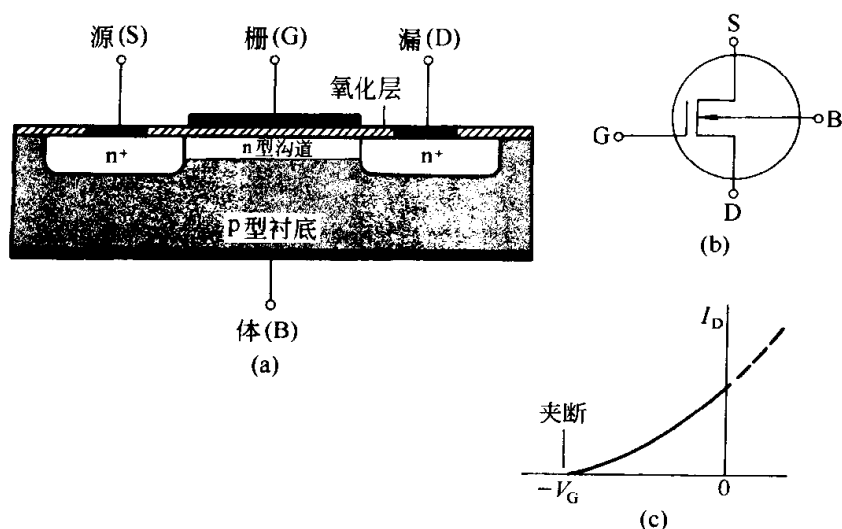


图 9-31 耗尽型 n 沟 MOSFET

(a) 结构示意图; (b) 符号; (c) I_D - V_G 曲线

9.2.9 量子半导体器件

随着电子科学技术的迅猛发展, 为了适应高速度、高集成度和低功耗的

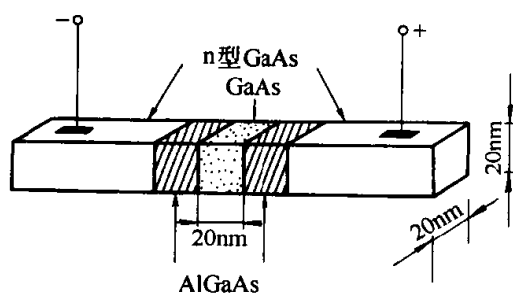


图 9-32 量子半导体器件

AlGaAs-GaAs-AlGaAs 的结构示意图

需要, 半导体器件的尺寸愈来愈小, 以至于达到老一代器件尺寸的极限, 因此必须寻求新型的器件。这些新型器件的尺寸已经可与电子的德布罗意波长相比拟 (例如, GaAs 中电子德布罗意波长约为 20nm), 表现出明显的量子化效应, 原来在大尺寸时连续分布的能级变成了分立能级, 这被称为尺寸量子化 (size quantization) 或量子限制效应 (quantum confinement effect)。

图 9-32 所示是一个典型的量子半导体器件, 两个宽带隙的 AlGaAs (AlGaAs 的化学分子式应为 $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$, 这里从简) 区中间夹一个线度为 20nm 的窄带隙 GaAs 区, 两边的电极通过重掺杂的低阻 n 型 GaAs 区与两个 AlGaAs 区相连接。由于 AlGaAs 材料的亲和能小于 GaAs 材料, 所以前者的导带底高于后者, 如图 9-33(a) 所示。在这样的三明治 (sandwich) 结构中, 导带在 GaAs 区形成势阱, 原来在小尺寸时能量准连续分布的导带在此小尺寸的势阱中由于量子化效应分裂成分立能级。如果 GaAs 区三个方向的尺寸都很小 (例如, 图 9-32 中所示, 皆为 20nm), 被称为量子点; 如果仅两个方向的尺寸很小, 另一个方向的尺寸很大, 则为量子线; 如果只有一个方向的尺寸很小, 其他两个方向的尺寸很大, 则为量子阱。若在器件两端加上偏压, 左端 A 区接负, 右

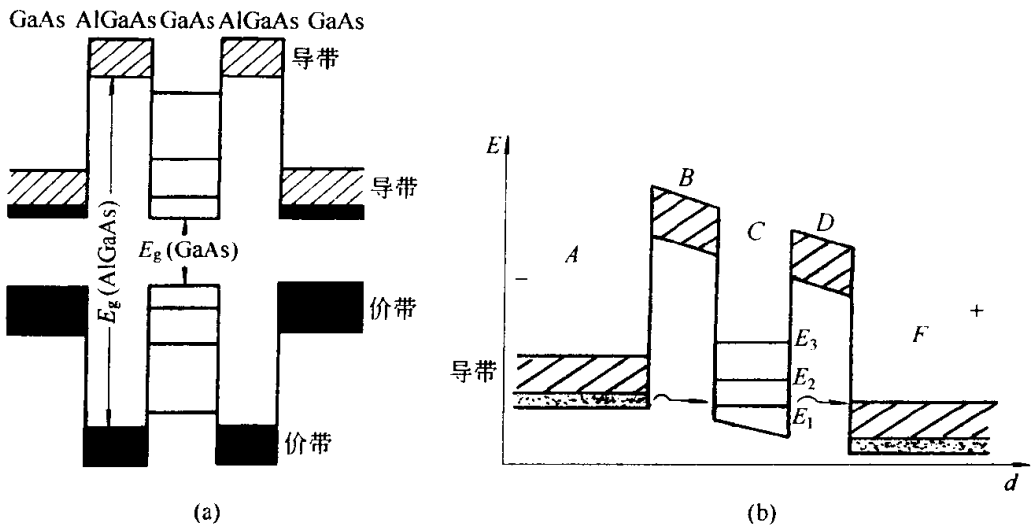


图 9-33 量子半导体器件 AlGaAs-GaAs-AlGaAs 的能带图
(a) 平衡时；(b) 偏压下

端 F 区接正，如图 9-33(b) 所示，由于 A 区电子受到 B 区高势垒的阻挡，起初电子流很小。随着偏压的加大，使得 B 、 C 、 D 区的导带底进一步倾斜， A 区的导带底部进一步升高，当升高到与 C 区量子阱中的 E_1 能级相平时，由于量子力学隧道效应， A 区电子隧穿过 B 区势垒进入 C 区，并可再隧穿通过 D 区进入 F 区，形成较大的电子流，称为共振隧穿。随着偏压的进一步加大， B 、 C 、 D 区的导带底更加倾斜， A 区导带底部继续升高，逐步高于 C 区中的 E_1 能级，由 A 区到 C 区的隧穿逐步被禁止，电流下降，出现如图 9-34 中所示的第一个电流峰。此时，随着偏压的增加，电流不是增加，而是减小，出现了负阻现象（负的微分电阻 dI/dV ），此被称为负阻效应，如图 9-34 所示（类似的隧道效应及负阻现象早在 20 世纪 60 年代就被日本科学家江崎在重掺杂的 p^+-n^+ 结二极管中发现）。随着偏压的继续加大，当 A 区的导带底部与 C 区的 E_2 能级相平时，隧穿再度发生，并出现如图 9-34 所示的第二个电流峰。另一个与量子限制效应相关的现象是发生在 C 区的共振现象。在某特殊偏压下，使得 C 区势阱内的电子波恰好在两边势阱壁间来回反射，发生相长干涉而加强，引起共振。此偏置电压为共振电压。量子阱的负阻效应和共振效应可以用来制作振荡器。负阻效应还可用于制作开关器件，开关器件是计算机的重要组成元件。量子开关器件的尺寸只有通常晶体管尺寸的 1%，这样就会使计算机集成电路芯片的集成度提高 1 万倍，如此诱人

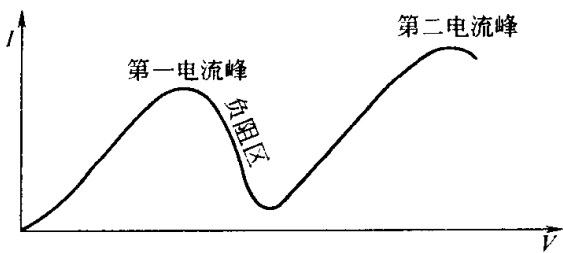


图 9-34 量子点的电流-电压特性

的前景是令人振奋的。不过，目前还有一些技术问题需要解决，例如器件设计和引线等。很容易想像，可以将 $B/C/D$ 三层结构的单量子阱发展成为多层结构的多量子阱，并且可以在半导体材料中人为地引入周期性，宽窄带隙的材料交替相间地有序排列，这样的结构称之为超晶格结构。

9.3 聚合物、非晶态材料等的输运性质

本章前两节所论及的金属和半导体材料基本上都是无机晶态材料，所讨论的载流子是电子和空穴，空穴是近填满能带的电子集体，其本质仍然是电子。而人们研究的材料除无机晶态材料外还有有机、聚合物及非晶态材料等，人们研究的载流子除了电子、空穴外还有离子、孤子和极化子等。近二三十年来人们在有机、非晶态、聚合物等导电材料（包括金属和半导体）方面的研究有长足的进展，取得了不少引人注目的研究成果。例如，孤子和极化子就是一类新型的载流子，是电子和晶格耦合体系的元激发，具有特殊的性质。如此多样的导电材料类型和载流子种类，使得各种金属和半导体材料的输运性质丰富多彩，这大大丰富了基础研究的内容，并为研究开发种类更多、性能更优越的材料提供了更加灵活多样的途径。

9.3.1 导电聚合物^[7,8]

聚乙炔 $(CH)_x$ 是研究得最多的导电聚合物材料。聚乙炔结构简单，仅含 C 和 H 两种原子。掺杂聚乙炔的电导率可以很高，甚至可以与铜媲美。下面以聚乙炔为例对导电聚合物作简单介绍。聚乙炔材料中的碳原子彼此键合形成准一维碳链。每一个氢原子分别与一个碳原子键合位于碳链的两侧，如图 9-35 所示。每个碳原子与两个碳原子和一个氢原子相邻。每个碳原子有四个价电子，在聚乙炔分子键中，碳原子以 sp^2 杂化轨道相互交叠与相邻碳原子形成碳碳 σ 键，并以 sp^2 杂化轨道与氢原子的 s 轨道交叠形成碳氢 σ 键。 σ 键中的电子称为 σ 电子。每个碳原子还有一个价电子（ p_z 轨道），其在分子键中形成 π 电子，见图 9-36。

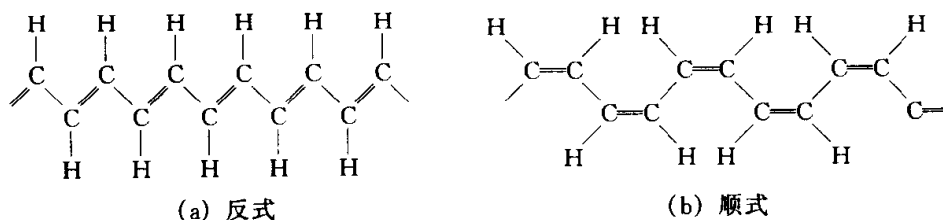


图 9-35 聚乙炔分子链

不妨设想碳原子等间距地排列成一维晶格（如图 9-37 的上部所示），晶格常数为 a （对于聚乙炔， $a = 0.122\text{nm}$ ）。在一维晶格中，碳原子提供的 π 电子的能态形成能带，能带中的能级数为碳原子数 N ，考虑到电子自旋，

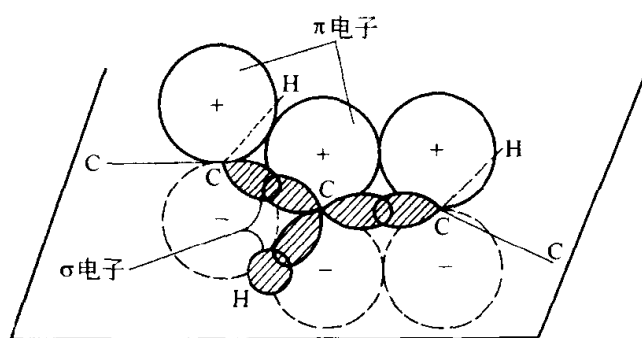
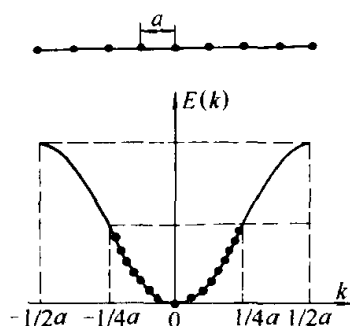
图 9-36 聚乙炔中碳原子的 σ 电子和 π 电子

图 9-37 等间距一维碳链的能带

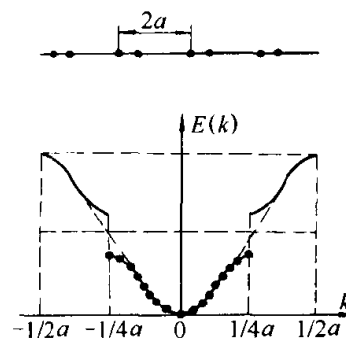


图 9-38 二聚化一维碳链的能带

该能带可容纳 $2N$ 个电子。每个碳原子提供一个 π 电子， N 个 π 电子占据该能带的一半（如图 9-37 所示），如此半填满的能带具有很强的导电性，材料应呈金属性。然而，实验测量结果表明，本征的聚乙炔并不是导体，上述推测与实际情况不符。通过分析研究发现，问题在于碳原子等间距排列并非是系统的最低能量状态。考虑碳原子两两相聚的情况，即二聚化状态（如图 9-38 的上部所示），与等间距排列相比，晶格常数增大一倍，由 a 增加为 $2a$ ；布里渊区缩小一倍，由 $-\frac{1}{2}a \sim \frac{1}{2}a$ 缩小为 $-\frac{1}{4}a \sim \frac{1}{4}a$ ，能带一分为二。能带中央出现的禁带将二聚化前半填满能带的电子能量下压，使系统的能量降低。由此可见，等间距排列的状态能量较高，是不稳定的，这被称之为派尔斯（Peierls）失稳性。二聚化引起的形变称之为派尔斯形变。形变将伴随弹性势能的升高，当弹性势能的升高与电子能量的降低达到平衡时，即

$$\frac{\partial E_{\text{弹性}}}{\partial \Delta} + \frac{\partial E_{\text{电子}}}{\partial \Delta} = 0 \quad (9-70)$$

形变达到稳定，此时二聚化位移量 $\Delta = \Delta_0$ 。在二聚化状态的分子链中，距离较近的碳原子间的相互作用较强，为双键；距离较远的碳原子间的相互作用较弱，为单键。如此单双键交替的结构是一维碳链能量较低的状态。具有单双键交替碳链结构的有机聚合物被称为共轭有机聚合物。聚乙炔有两种主

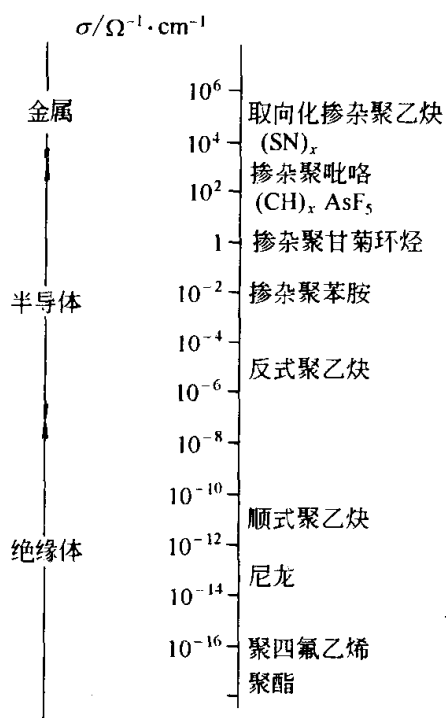


图 9-39 聚合物的
室温电导率

要的同分异构体：反式聚乙炔和顺式聚乙炔，分别如图 9-35(a) 和 (b) 所示。所谓反式 (*trans*) 是指与由双键连接的两个碳原子键合的两个氢原子分布在该双键的异侧；所谓顺式 (*cis*) 是指与由双键连接的两个碳原子键合的两个氢原子分布在该双键的同侧。测量结果表明，反式聚乙炔中二聚化位移量 Δ_0 为 0.004nm，禁带宽度约为 1.5eV，这与无机半导体材料的禁带宽度相当（硅的禁带宽度为 1.12eV）。如图 9-39 所示，反式聚乙炔的室温电导率在半导体电导率的范围内。这种聚合物被称为半导性聚合物 (semiconducting polymer)，也称为聚合物半导体。

与无机半导体相似，通过掺杂可以提高聚合物半导体的电导率。但在掺杂的反式聚乙炔中人们发现有异常的实验现象，即在一定的掺杂浓度范围内，反式聚乙炔具有高的电导率 σ ，但其磁化率 $\bar{\chi}$ 却为零，如图 9-40 所示。这意味着其中的载流子虽然荷电却没有自旋。W.P.Su (苏武沛)、J.R.Schrieffer 和 A.J.Heeger 三人提出孤子模型成功地解释了这一异常的实验现象。反式聚乙炔有两个能量最低的二聚化状态 (基态) A 相和 B 相，这两个状态的结构对称，能量相同，如图 9-41 的两侧所示，这样的体系称之为基态简并的准一维导电体系。如果 A 相和 B 相在同一反式聚乙炔分子链上共存，两者之间就会形成畴壁，畴壁处单、双键交替的结构被破坏，此畴壁被称为孤子 (soliton)，如图 9-41 的中间所示。若图中孤子左侧的单键与其相邻的双键交换位置，则孤子向左发生了移动。如此移动可以连续向左进行，孤子便在分子链中向左运动，此运动的结果使 A 相缩短，使 B 相增长。由于 A 相和 B 相的能量相等，系统的总能量不变，孤子在分子链中可以自由地运动，向左向右皆可。孤子如果带电，在电场的作用下将会发生迁移运动，因此，荷电的孤子是载流子。理论计算结果表明，

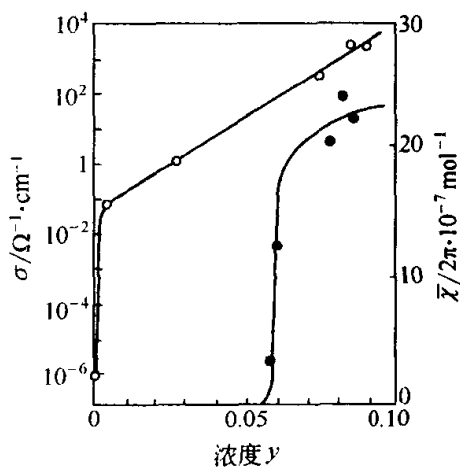


图 9-40 反式聚乙炔的
电导率 σ 和磁化率 $\bar{\chi}$

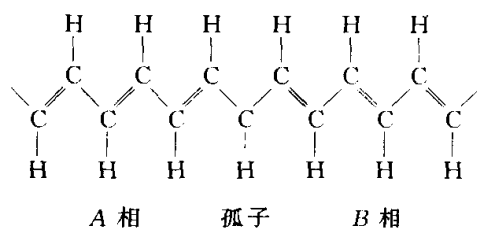


图 9-41 反式聚乙炔的 A 相、B 相和孤子

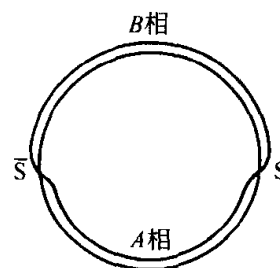


图 9-42 孤子对示意图

孤子沿着分子链方向的迁移率约为 $200\text{cm}/(\text{V}\cdot\text{s})$ 。聚乙炔分子链通常很长，我们不妨以环形的分子链来表示无限长的分子链。如果在环形的 A 相反式聚乙炔分子链中出现一段 B 相的反式聚乙炔分子链，便有一对孤子 (S) 和反孤子 (\bar{S}) 出现，如图 9-42 所示。因此，通常认为孤子总是成对产生的。

反式聚乙炔分子链中出现孤子会在禁带中央形成局域的孤子能级 E_S ，如图 9-43(a) 所示。对于中性孤子，孤子能级 E_S 上占有一个电子，自旋为 $\pm 1/2$ ；而正电孤子，孤子能级 E_S 是空的，自旋为 0；负电孤子，孤子能级 E_S 上占有两个电子，自旋也为 0。如此荷电与自旋的反常关系（荷电却无自旋，不荷电反而有自旋）是孤子的一个重要特征。正是由于孤子具这样的特征，才能很好地解释上述如图 9-40 所示的异常实验现象。

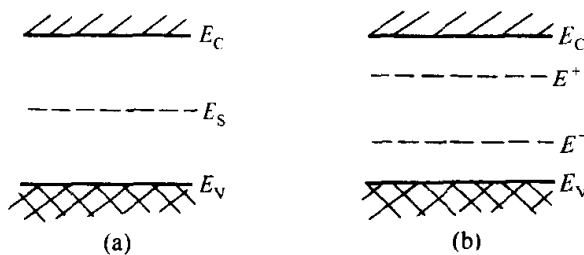


图 9-43 孤子(a)和极化子(b)的电子能谱

聚乙炔分子链的电荷可以由掺杂产生。当反式聚乙炔材料中掺入施主杂质（如 I_2 、 AsF_5 ）时，这些位于链间的施主杂质向反式聚乙炔分子链施放电子，使之带负电，从而可以形成负电孤子。同样，掺入受主杂质（如 Na 、 K ），可以使反式聚乙炔分子链带正电，可以形成正电孤子。随着掺杂浓度的增加，有更多的分子链荷电，荷电孤子浓度增加，电导率上升。当掺杂浓度进一步增大时，分子链的线电荷密度升高，单位长度分子链的荷电孤子数增加，众多荷电孤子和反孤子在分子链上交替地周期性排列，形成孤子晶格。聚乙炔分子链的电荷也可以由光照产生。光照可以将反式聚乙炔中某分子链的电子激发到另一个分子链，从而使前者带正电，后者带负电，并形成正电孤子和负电孤子。

导电聚合物中，除了反式聚乙炔是基态简并外，绝大多数是基态非简并的，例如，聚苯胺(polyaniline)和聚吡咯(polypyrrole)等，连顺式聚乙炔也

是基态非简并的。在基态非简并的导电聚合物中， B 相与 A 相能量不等 ($E_B > E_A$)。正因为如此，倘若如图 9-42 所示，在 A 相的顺式聚乙炔分子链中出现 B 相， B 相将有缩短的趋势，以降低系统的总能量。因此，在基态非简并的导电聚合物中一对孤子总是束缚在一起的，如此束缚在一起的孤子对被称为极化子。极化子也可以在分子链中运动，荷电的极化子也是载流子。导电聚合物分子链中出现极化子会在禁带中形成一对能量对称分布的局域能级 E^+ 和 E^- ，如图 9-43(b) 所示。在中性极化子中，下能级 E^- 填满两个电子，上能级 E^+ 全空；在正电极化子中，上能级 E^+ 全空，下能级 E^- 占据一个电子（为正电单极化子，带 $+q$ 电荷）或全空（为正电双极化子，带 $+2q$ 电荷）；在负电极化子中，下能级 E^- 填满两个电子，上能级 E^+ 占据一个电子（为负电单极化子，带 $-q$ 电荷）或填满两个电子（为负电双极化子，带 $-2q$ 电荷）。研究表明，在基态简并的一维体系中，亦可有稳定的单极化子存在。作为载流子的荷电孤子和荷电极化子与无机金属和半导体中的电子与空穴不同，它们是电子和晶格耦合体系的元激发。在一维分子链中形成孤子或极化子时，电子能量状态的改变与晶格状态的改变相伴发生，这是一类新的载流子。

9.3.2 非晶态半导体^[9]

非晶态半导体的能带结构与晶态半导体的能带结构有显著的不同。晶态半导体中由于晶格排列的周期性，电子波函数是布洛赫波，电子可在整个晶体中运动（称为共有化运动）。而非晶态半导体中的电子态有两类：一类电

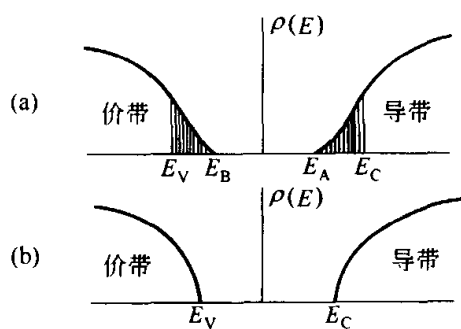


图 9-44 导带和价带的态密度分布
(a) 非晶态半导体；(b) 晶态半导体

子态的波函数可延伸到整个晶体，与晶态半导体中的共有化运动状态相类似，称为扩展态；另一类电子态的波函数局限在一些中心附近，随着与中心距离的增大，其强度按指数衰减，称为定域态。晶态半导体和非晶态半导体中导带和价带的态密度函数 $\rho(E)$ 如图 9-44 所示。对于非晶态半导体，导带中除了在 $E > E_C$ 区域有扩展态外，在 $E_A < E < E_C$ 区域还有定域态，称之为带尾态。同样，价带中除了在 $E < E_V$ 区域有扩展态外，在 $E_B > E > E_V$ 区域也有定域态，亦为带尾态。 E_C 和 E_V 是扩展态与定域态的分界，因为在温度 $T \rightarrow 0$ 时，定域态中的载流子迁移率为零，而扩展态中载流子的迁移率为有限值，故人们将此分界称为迁移率边。

由于非晶态半导体中缺陷密度大，因此在禁带中存在大量定域态，这些

定域态能级形成窄的能带，其能态密度分布如图 9-45 所示。带尾态和缺陷态是两种不同类型的定域态。带尾定域态是由非晶半导体的本征性质决定的，它源于无规网络中键的畸变。而缺陷定域态对制备方法和工艺条件有很强的依赖性，一般认为缺陷定域态与悬挂键相关。例如非晶态硅，在四价硅原子以共价结合形成的无序网络中存在大量的悬挂键，如图 9-46(b) 和(c)所示，这些悬挂键类似于晶态硅中的空位，见图 9-46(a)。

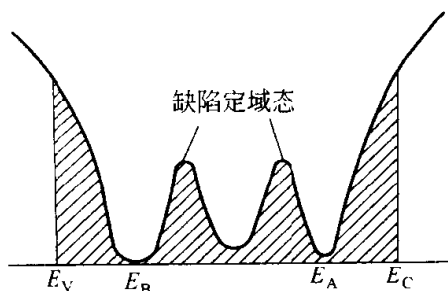


图 9-45 禁带中的缺陷
定域态密度分布

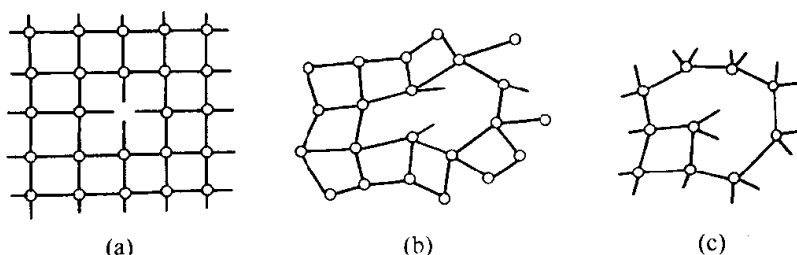


图 9-46 材料中的缺陷
(a) 晶态硅中；(b)、(c) 非晶态硅中

非晶态半导体的导电机制与晶态半导体的导电机制亦有很大的不同，晶态半导体的电导靠导带中的电子和价带中的空穴；而在非晶态半导体中，导带和价带中的扩展态和带尾定域态以及禁带中的缺陷定域态都可对电导有贡献，需分别加以考虑。非晶态半导体导带和价带中扩展态的导电机制和晶态半导体相似，但是由于非晶态半导体中结构畸变和缺陷的散射作用很强，使得电子和空穴的迁移率较低。以 n 型非晶态半导体为例，其扩展态的电导率 σ_0 与扩展态中的电子浓度 n_0 (取决于 $E_C - E_V$) 成正比，因此

$$\sigma_0 = \sigma'_0 \exp\left[\frac{E_0}{k_B T}\right] \quad (9-71)$$

式中， $E_0 = E_C - E_F$ ， E_F 为费米能级。由于禁带中的缺陷定域态密度很高，所以费米能级 E_F 被“钉扎”在禁带中，几乎不随温度改变。 σ'_0 为一系数，与扩展态的有效能级密度及其中载流子的迁移率相关。

电子在定域态之间的运动是跳跃式的，依赖于与晶格的相互作用，这是一个热激活过程。跳跃的几率为

$$P = \nu \exp(-2\alpha R) \exp\left(\frac{-W}{k_B T}\right) \quad (9-72)$$

其中， ν 为声子的振动频率，约为 10^{12} /s； R 为两定域态之间的平均距离，

R 愈大, 两定域态之间波函数的重叠愈小, 跳跃的几率也愈小; α 为一系数; W 为跳跃过程所需的平均能量, 称为激活能。不妨把电子在定域态之间的跳跃运动视为布朗运动, 根据布朗运动的统计理论可知扩散系数为

$$D = PR^2 \quad (9-73)$$

在这里, R 为每次跳跃在空间移动的平均距离。根据爱因斯坦关系 $\mu/D = q/k_B T$, 再利用式 (9-73) 和式 (9-72) 可得迁移率

$$\mu = \left(\frac{qR^2}{k_B T} \right) \nu \exp(-2\alpha R) \exp\left(\frac{-W}{k_B T}\right) \quad (9-74)$$

如前所述, 定域态有带尾态和缺陷态两种。当带尾定域态起导电作用时, 电子间的跳跃通常只在相邻带尾定域态间进行, 因此跳跃几率

$$P_1 \propto \exp\left(\frac{-W_1}{k_B T}\right) \quad (9-75)$$

W_1 为电子在相邻带尾定域态间跳跃的激活能。考虑到其电导率 σ_1 应与带尾中电子浓度 n_1 (取决于 $E_A - E_F$) 成正比, 因此有

$$\sigma_1 = \sigma'_1 \exp\left(\frac{-W_1}{k_B T}\right) \exp\left[\frac{-(E_A - E_F)}{k_B T}\right] = \sigma'_1 \exp\left(\frac{-E_1}{k_B T}\right) \quad (9-76)$$

式中, $E_1 = E_A - E_F + W_1$; σ'_1 为一系数, 与带尾态的有效能级密度及其中载流子的迁移率相关。当禁带中缺陷定域态起导电作用时, 由于费米能级 E_F 处于缺陷定域态的能级之中, 因此就像金属中的电导那样, 只是能量在 E_F 附近的电子才对电导有贡献, 并认为这些电子仅限于 E_F 附近 $k_B T$ 范围内, 其浓度 n_2 为 $\rho(E_F) k_B T$ 。这里 $\rho(E_F)$ 为 E_F 处的能态密度。利用式 (9-74) 可得其电导率

$$\sigma_2 = n_2 q \mu = \rho(E_F) q^2 R^2 \nu \exp(-2\alpha R) \exp\left(\frac{-W}{k_B T}\right) \quad (9-77)$$

在不太低的温度下, 电子的跳跃在相邻的缺陷定域态之间进行, 其激活能 $W = W_2$ 。故有

$$\sigma_2 = \sigma'_2 \exp\left(\frac{-E_2}{k_B T}\right) \quad (9-78)$$

其中, $E_2 = W_2$, $\sigma'_2 = \rho(E_F) q^2 R^2 \nu \exp(-2\alpha R)$ 。综合上述三种不同情况的电导, 利用式 (9-71)、式 (9-76) 和式 (9-78), 可得总电导率为

$$\sigma = \sigma_0 + \sigma_1 + \sigma_2 = \sigma'_0 \exp\left(\frac{-E_0}{k_B T}\right) + \sigma'_1 \exp\left(\frac{-E_1}{k_B T}\right) + \sigma'_2 \exp\left(\frac{-E_2}{k_B T}\right) \quad (9-79)$$

其中, 第一项为扩展态的电导, 第二项为带尾定域态的电导, 第三项为缺陷定域态的电导。图 9-47 给出非晶半导体电导率随温度变化的曲线, 在各个温度区域, 往往以某种导电机制为主, 其中(a)、(b)和(c)三段分别对应于

上述三种不同导电机制。

用辉光放电方法制备的非晶硅，可以将缺陷定域态密度控制得很低。这是因为在此法的制备过程中，硅烷（ SiH_4 ）分解时产生大量的氢进入非晶硅中，或饱和了非晶硅中的悬挂键，或与硅形成非晶态硅氢合金。低缺陷定域态密度的氢化非晶硅具有重要的实用价值，其室温电导以扩展态电导为主，性能可以与晶态硅相比，用氢化非晶硅制成的太阳能电池的效率可达

8%，虽然低于单晶硅太阳能电池的效率，但价格低廉，因此被广泛采用。并且，由非晶态硅制成的薄膜晶体管阵列被大量应用于液晶显示中。

9.3.3 离子晶体^[10]

在氯化钠晶体中，钠原子将电子转移给氯原子，形成带正电的钠离子和带负电的氯离子。这种由正负离子构成的晶体称之为离子晶体。离子晶体的禁带宽度很宽，本征激发产生的电子空穴浓度极小，由此引起的导电作用完全可忽略。离子晶体中的载流子主要是正负离子。可是由于正负离子间有很强的静电束缚，因此在室温下，离子晶体的电导率要比典型金属的电导率低很多（约低 22 个量级，见图 9-1）。像氯化钠这样的离子晶体，有两种荷电相反的离子，但往往只有一种离子在电导中起主导作用。例如，在氯化钠晶体中，钠离子起主要导电作用。因此，离子晶体的电导率可写为

$$\sigma_{\text{离子}} = q_{\text{离子}} n_{\text{离子}} \mu_{\text{离子}} \quad (9-80)$$

其中， $n_{\text{离子}}$ 和 $\mu_{\text{离子}}$ 为起主导作用的离子的浓度和迁移率， $q_{\text{离子}}$ 为离子的电量。离子晶体中离子的浓度虽不小，但迁移率很低。离子在晶格中运动需要具备足够的能量，以克服周围的势垒跳出原格点，这是一个热激活过程，离子具备此能量的几率将随温度的升高而增加。在没有外加电场的情况下，晶体中离子的运动是相反方向彼此平衡、互相抵消的，见图 9-48(a)。当有外加电场存在时，离子便产生了定向的漂移运动，见图 9-48(b)。离子的迁移率依赖于离子的热激活几率。根据统计理论，可将离子的迁移率写成如下形式

$$\mu_{\text{离子}} = \mu'_{\text{离子}} \exp\left(\frac{-Q}{k_B T}\right) \quad (9-81)$$

式中， $\mu'_{\text{离子}}$ 和 Q 皆为系数， Q 被称为离子的激活能。将式 (9-81) 带入式 (9-80) 得

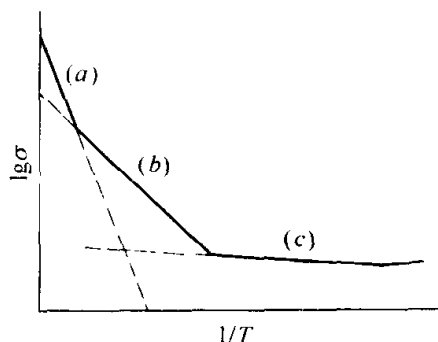


图 9-47 非晶态半导体
电导率与温度的关系

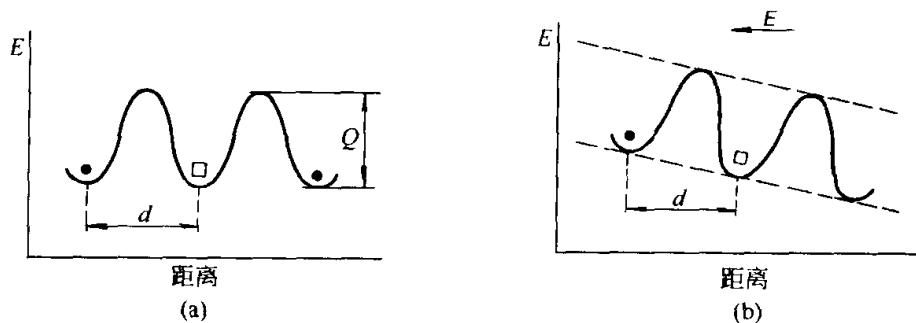


图 9-48 离子晶体中离子运动需要克服的势垒

(a) 无外电场; (b) 有外电场

$$\sigma_{\text{离子}} = \sigma'_{\text{离子}} \exp\left(\frac{-Q}{k_B T}\right) \quad (9-82)$$

式中, $\sigma'_{\text{离子}} = q_{\text{离子}} n_{\text{离子}} \mu'_{\text{离子}}$ 。对式(9-82) 取自然对数得

$$\ln \sigma_{\text{离子}} = \ln \sigma'_{\text{离子}} - \left(\frac{Q}{k_B}\right) \left(\frac{1}{T}\right) \quad (9-83)$$

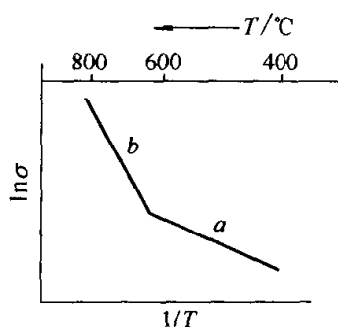


图 9-49 氯化钠中钠离子的电导率随温度的变化

由此式可知, $\ln \sigma_{\text{离子}} \sim 1/T$ 曲线应是斜率为某一负数的直线。该理论分析结果得到实验的证实, 图 9-49 给出了氯化钠晶体中钠离子的 $\ln \sigma \sim 1/T$ 关系, 确实呈线性, 且斜率为负, 由直线的斜率可以求出其激活能。然而, 图 9-49 中呈现的却不是一条直线, 而是由两条直线构成的折线。较低温度区域的直线 a 所对应的激活能较小, 较高温度区域的直线 b 所对应的激活能较大, 其原因如下。因为格点间隙处势能很高, 所以离子位于格点间隙处的几率很小, 也很难在格点的间隙中运动。因此, 离子在晶体中运动不但要有足够的能量克服势垒跳出所在的格点, 而且附近要有空位能接纳该离子。离子晶体中的空位有两类。一类是本征的, 是离子晶体本身因热激发引起的。离子晶体由热激发可以产生两种缺陷, 一种是离子由格点被热激发到格点的间隙处, 从而形成空位-间隙离子对, 此为费仑克尔缺陷; 另一种是离子离开原格点跑到晶体表面的正常格点上, 此为肖特基缺陷。正如前面所说, 由于填隙离子要挤在格点的间隙中, 能量较高, 产生的几率较小, 因此本征空位的产生以肖特基缺陷型为主。另一类空位是非本征的, 是由杂质引起的。例如, 在氯化钠晶体中掺入镁, 二价的镁离子代替了一价的钠离子的格点位置。为了保持电中性, 镁离子附近必然会有一个钠离子被移开, 如此便形成一个钠离子空位。前一类由热激发引起的本征空位的产生伴随激活过程, 空

小, 也很难在格点的间隙中运动。因此, 离子在晶体中运动不但要有足够的能量克服势垒跳出所在的格点, 而且附近要有空位能接纳该离子。离子晶体中的空位有两类。一类是本征的, 是离子晶体本身因热激发引起的。离子晶体由热激发可以产生两种缺陷, 一种是离子由格点被热激发到格点的间隙处, 从而形成空位-间隙离子对, 此为费仑克尔缺陷; 另一种是离子离开原格点跑到晶体表面的正常格点上, 此为肖特基缺陷。正如前面所说, 由于填隙离子要挤在格点的间隙中, 能量较高, 产生的几率较小, 因此本征空位的产生以肖特基缺陷型为主。另一类空位是非本征的, 是由杂质引起的。例如, 在氯化钠晶体中掺入镁, 二价的镁离子代替了一价的钠离子的格点位置。为了保持电中性, 镁离子附近必然会有一个钠离子被移开, 如此便形成一个钠离子空位。前一类由热激发引起的本征空位的产生伴随激活过程, 空

位的数目随温度的升高而指数上升；后一类由杂质引起的非本征空位的数目取决于杂质的多少。在图 9-49 的低温区(a 区)内，热产生的本征空位很少，应以杂质引起的非本征空位为主，它不是由热激发引起的，因此，直线 a 所对应的激活能只是离子克服势垒跳出格点所需的能量。此区域称之为非本征区。随着温度的升高，到了高温区(b 区)，由热激发引起的本征空位数大大增加，超过了由杂质引起的非本征空位。空位热激发所需的能量加上离子克服势垒跳出格点所需的能量，使得离子迁移率（或电导率）的激活能变大。这就是直线 b 所对应的激活能。如果材料不是单晶，而是多晶，多晶的晶粒间界处有大量的空位，为离子的运动提供了条件，从而提高了材料的电导率。

在外加偏压的作用下，离子晶体中正负离子沿着相反的方向朝正负电极移动，这将会导致物质的分凝。事实上，在实验中可以观察到组成物质在电极附近的集聚，当然，这只是在阻塞电极的情况下。对于非阻塞电极，因已为物质迁移提供了通道，则不会有物质的堆积。

参 考 文 献

- 1 Hummel R.E. "Electronic Properties of Materials", Second Edition, Springer, Springer International Student Edition, 1994
- 2 C. 基泰尔著，杨顺华等译，固体物理导论，北京：科学出版社，1979
- 3 E.H. 普特来著，傅德中译，霍尔效应及有关现象，上海：上海科学技术出版社，1964
- 4 刘恩科，朱秉升，罗晋生等编，半导体物理学，北京：国防工业出版社，2001
- 5 叶良修编著，半导体物理学，上册，北京：高等教育出版社，1987
- 6 A.S. 格罗夫著，齐建译，半导体器件物理与工艺，北京：科学出版社，1976
- 7 孙鑫著，高聚物中的孤子和极化子，成都：四川教育出版社，1987
- 8 Su W.P, Schrieffer J.R. and Heeger A.J. Phys. Rev. Lett., 1979, 42: 1698; Phys. Rev. B, 1980, 22: 2099
- 9 郭贻诚，王震西主编，非晶态物理学，北京：科学出版社，1984
- 10 方俊鑫，陆栋主编，固体物理学，上册，上海：上海科学技术出版社，1989

第 10 章 磁 学 性 质

撰稿人 翟宏如

一切物质均有磁性。通常所谓的磁性材料与非磁性材料，实际上是指强磁性及弱磁性材料。前者的磁化率比后者大 $10^4 \sim 10^{12}$ 倍，因而获得巨大应用，成为电工技术的基础材料之一，并广泛应用于电工、电子和计算机等技术中。

最早发现的磁性为磁铁矿的强磁性。我国于公元前 2500 年已开始将磁铁矿应用于指南^[1]。在公元前 3~4 世纪已有“磁石之取针”及“磁石召铁”的记载^[2]。在西方，希腊 Thales 的著作中指出公元前 600 年有磁铁矿吸铁的概念。然而，现代磁学的建立及磁性材料的发展则是近 100 多年的事，是与近代工业技术发展相联系的。

人们对于弱磁性的认识发生在 18~19 世纪，因为弱磁性的测量需借助于仪器。1845 年法拉第 (M. Faraday) 建立了对抗磁性与顺磁性的认识，19 世纪末居里 (P. Curie) 进一步发现了抗磁磁化率不随温度变化及一些顺磁物质磁化率与温度成反比的居里定律。20 世纪初郎之万 (P. Langevin) 用经典统计理论证明了居里定律，外斯 (P. Weiss) 提出分子场自发磁化的假说与磁畴的概念，用以解释铁磁性，并建立居里-外斯定律。

量子力学的出现使人们开始把对物质磁性的认识建立在原子及电子的基础上，海森伯 (W. K. Heisenberg) 用静电性的交换作用对铁磁体中的分子场的性质给出了正确的解释，揭开了现代磁学的篇章。1932 年 J. H. Van Vleck 的著作^[3]已对物质磁性作出相当全面的量子理论阐述。与此同时，磁畴及磁化过程的实验研究和宏观理论逐渐发展。它们与固体结构、特别是微结构密切相关。20 世纪 40 年代末及 50 年代，以铁氧体为代表的亚铁磁体的发现、研究及其应用形成一个热潮，改变了 100 多年来金属铁磁性材料独占强磁体领域的局面。强磁材料的研究及应用发展到高频和微波领域^[4]。20 世纪 50 年代后的两个重大突破为稀土化合物及其合金和非晶态磁性材料的研究及应用^[5,6]。强磁材料所含元素从 $3d$ 发展到 $4f$ 族。非晶合金磁性的发展不仅打破了晶态统一强磁物质的局面，而且开拓了一个优质软磁材料的领域。

近 20 年，一个新的应用领域正在迅猛发展，即各种磁记录材料和磁光记录材料，后者将磁性材料的研究和应用发展到光频领域^[7]。同时，薄膜、超薄膜、多层膜、超微粒材料和纳米结构的研究和应用成为又一热点。1988 年

在多层膜中发现巨磁电阻以来,自旋相关导电等新的材料及器件不断发展,开始形成一个新的学科——磁电子学^[8]。此外,有机铁磁体、C₆₀化合物铁磁性和室温下铁磁半导体的发现预示了磁学与磁性材料引人入胜的发展前景。

目前,磁学的研究领域十分广泛^[1,2],遍及了各种元素及其合金和化合物,包括金属、半导体和绝缘体,有机和无机。多数材料为弱磁性,但材料磁性的研究重点集中在与应用有关的强磁性方面。这也将是本章的重点。

10.1 材料磁性来源与分类

10.1.1 原子磁矩

材料由原子组成,为了理解材料的磁性,应该对原子磁性有基本了解。这里,我们首先讨论孤立原子的磁矩,然后再简介原子在材料中的磁性表现。

原子磁矩有三个来源,即电子轨道运动磁矩、电子自旋磁矩及原子核磁矩。

原子核磁矩比前二者约小 2000 倍,对于宏观磁性而言,可以忽略。

(1) 孤立原子磁性

① 电子轨道磁矩。电子轨道运动的角动量为

$$p_l = \sqrt{l(l+1)}\hbar \quad (10-1)$$

式中, l 为角量子数,若主量子数为 $n=1, 2, 3, \dots$, 则 $l=n-1, n-2, \dots, 0$ 。

$$\hbar = \frac{h}{2\pi} = 1.05450 \times 10^{-34} \text{J}\cdot\text{s}$$

相应的磁偶极矩为

$$\mu_l = \sqrt{l(l+1)}\mu_B \quad (10-2)$$

μ_B 称为波尔磁子。SI 制: $\mu_B = \frac{\mu_0 e \hbar}{2m} = 1.165 \times 10^{-29} \text{Wb}\cdot\text{m}$

若取磁矩, 则 $\mu_B = \frac{e \hbar}{2m} = 0.9273 \times 10^{-23} \text{J/T}$

或 CGS 制: $\mu_B = \frac{e \hbar}{2mc} = 0.9273 \times 10^{-20} \text{erg/G}^\bullet$

磁矩与角动量的关系为

$$\mu_l = -\gamma_l p_l \quad (10-3)$$

式中, γ_l 称为旋磁比或磁力比, $\gamma_l = \frac{\mu_0 e}{2m}$ (SI 制) $= \frac{e}{2mc}$ (CGS 制)。

● 1erg = 10^{-7} J; 1T = 10^4 G。

角动量及磁矩沿磁场方向 Z 的分量决定于轨道磁量子数 m_l ，其值为：

$$p_{lz} = m_l \hbar \quad (10-4)$$

$$\mu_{lz} = m_l \mu_B \quad (10-5)$$

式中 $m_l = \pm l, \pm (l-1), \dots, \pm l, 0$

② 电子自旋磁矩。电子自旋的固有角动量及磁矩分别决定于自旋量子数 s ， $s = \frac{1}{2}$ ，

$$p_s = \sqrt{s(s+1)} \hbar = \sqrt{\frac{3}{4}} \hbar \quad (10-6)$$

$$\mu_s = 2 \sqrt{s(s+1)} \mu_B = \sqrt{3} \mu_B \quad (10-7)$$

自旋的磁力比为

$$\gamma_s = \frac{\mu_0 e}{m} (\text{SI 制}) = g_s \frac{e}{2mc} \quad (10-8)$$

$\gamma_s = 2\gamma_l$ ，或 $g_s = 2$ （精确值为 $g_s = 2.003$ ），自旋的磁量子数只有两个， $m_s = \pm \frac{1}{2}$ 则

$$p_{sz} = m_s \hbar = \pm \frac{1}{2} \hbar \quad (10-9)$$

$$\mu_{sz} = 2 m_s \mu_B = \pm \mu_B \quad (10-10)$$

③ 原子的总磁矩。原子的总角动量与总磁矩由所含电子的轨道与自旋的角动量耦合而成，多数原子取 Russell-Saunders 耦合。各电子的轨道与自旋角动量先分别合成总轨道角动量 P_L 和总自旋角动量 P_S ，然后两者再合成总角动量 P_J 。

总轨道角动量由总轨道量子数 L 决定：

$$P_L = \sqrt{L(L+1)} \hbar \quad (10-11)$$

$L = \sum m_{li}$ ，总的轨道磁矩 μ_L 为

$$\mu_L = \sqrt{L(L+1)} \mu_B \quad (10-12)$$

其在磁场方向的分量为

$$\mu_{Lz} = m_L \mu_B \quad (10-13)$$

$m_L = \pm L, \pm (L-1), \dots, 0$ 相应于 $(2L+1)$ 个取向。

总自旋角动量为

$$P_S = \sqrt{S(S+1)} \hbar \quad (10-14)$$

总自旋磁矩的数值为

$$\mu_S = 2 \sqrt{S(S+1)} \mu_B \quad (10-15)$$

$S = \sum_i m_{si}$ 。总自旋磁矩在磁场方向的分量为

$$\mu_{sz} = 2 m_S \mu_B \quad (10-16)$$

其中 $m_S = \pm S, \pm(S-1), \dots, 0$ 共 $(2S+1)$ 个取向。

原子总角动量决定于总角量子数 J ：

$$P_J = \sqrt{J(J+1)} \hbar \quad (10-17)$$

J 由 L 及 S 合成，依赖于 P_L 与 P_S 的相对取向， $J = |L - S|, |L - S| + 1, \dots, |L + S|$ 。

原子的总磁矩与总动量矩方向相反。

$$\mu_J = g_J \sqrt{J(J+1)} \mu_B \quad (10-18)$$

g_J 称为朗德劈裂因子

$$g_J = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} \quad (10-19)$$

原子的磁力比为

$$\gamma_J = g_J \frac{\mu_0 e}{2m} \text{ (SI 制)} = g_J \frac{e}{2mc} \text{ (CGS 制)} \quad (10-20)$$

g_J 的数值反映出电子轨道及自旋在原子总磁矩中的比例。当 $S=0$ 而 $L \neq 0$

时， $g_J = 1$ ， $\gamma_J = \gamma_l = \frac{\mu_0 e}{2m}$ ，当 $S \neq 0$ 而 $L=0$ 时， $g_J = 2$ ， $\gamma_J = \gamma_s = \frac{\mu_0 e}{m}$ ，当 S 和 L 都不为零时，孤立原子或离子的 g_J 值可以大于或小于 2。

原子总角动量和总磁矩沿磁场的分量及与磁场作用能分别为：

$$P_{Jz} = m_J \hbar \quad (10-21)$$

$$\mu_{Jz} = g_J m_J \mu_B \quad (10-22)$$

$$E_H = - g_J m_J \mu_B H \quad (10-23)$$

$m_J = \pm J, \pm(J-1), \dots, 0$ 共 $(2J+1)$ 个可能的值。

以上关于孤立原子磁矩的各个表示式亦都适用于孤立离子。

当原子的 $J=0$ 时，原子磁矩为零，当原子中的电子壳层均被填满时即属此情况。当原子中有的电子壳层未被填满时，其 $J \neq 0$ 时，原子具有不为零的磁矩，称为原子的固有磁矩。

④ 洪德法则。洪德 (Hund) 法则是用以确定含有未满壳层的原子或离子基态的电子组态及其总角动量的常用规则。简述如下：

a. 未满壳层中各电子的自旋取向 (m_s) 使总自旋量子数 S 最大时能量最低。

b. 在满足法则 a 的条件下，以总轨道角量子数 L 最大的电子组态能量最低。

c. 当未满壳层中的电子数少于状态数一半时, $J = |L - S|$ 的能量最低; 超过一半时, $J = |L + S|$ 。

在应用上述定则时, 以服从泡利 (Pauli) 原理为前提。每个电子状态中只能有一个电子。

(2) 材料中的原子和磁性^[1,2] 材料中原子的电子态和磁性与孤立原子相比发生了变化。材料中电子态的变化首先是由于键合使外层电子发生变化。共价结合常使价电子配对甚至杂化形成总磁矩为零的电子结构, 如氢分子。但也有例外, 如氧分子, 虽然其总电子数为偶数, 但电子组态为 $L=0$, $S=1$ 。离子化合物中的价电子在原子间转移使原子变为正负离子。如为简单元素, 可使有磁矩的原子变为无磁矩的离子。在金属中, 原子的价电子起着金属键的作用, 成为公有的传导电子。于是金属的磁性来自其正离子实和传导电子的磁性。在过渡金属中, d 电子能级也变为能带。 d 带 s 带的重叠使 s 及 d 带中的电子数与孤立原子时不同。例如孤立的钪原子外层电子组态为 $3d^1 4s^2$, 没有磁矩, 但金属钪中为 $3d^{0.6} 4s^{1.4}$ 。

晶体电场效应是引起材料中电子态发生变化的另一原因。局域在离子中的电子运动受邻近离子产生的静电场的作用而发生变化, 称为晶体电场效应。其一为使电子轨道简并态发生分裂; 其二为轨道磁矩对总磁矩的贡献减少甚至消失, 即为轨道矩的部分或全部淬灭。例如, 孤立的铁离子 Fe^{2+} 的 $L=2$, 而在化合物或金属中轨道矩几乎全部淬灭, 铁磁共振测得的 $J \sim S$, 其 g 因子比 2 略大。但稀土离子的 $4f$ 电子半径很小, 受到外部电子的屏蔽, 晶场效应弱, 其轨道矩基本未被淬灭, J 值与孤立原子相似。有的材料中, 晶场效应特别强, 简并能级分裂后的能差很大, 以致破坏洪德法则第 1 条—— S 最大的要求。如 $\text{Fe}^{3+} \text{CN}_6$ 中强晶场使分裂后的基态为三重简并, 高能态为二重简并, 于是 5 个 $3d$ 电子均进入到三重简并的轨道, 使自旋 $S=1/2$, 称为 Fe^{3+} 的低自旋态。而在较弱晶场中 Fe^{3+} 的 $S=5/2$, 称为高自旋态, $4d$ 和 $5d$ 族亦常因强的晶场效应而呈低自旋状态。

10.1.2 材料磁性概述和分类^[1-3,6,9,10]

材料在外磁场中恒被磁化而获得磁矩。单位体积的磁矩称为磁化强度 M 。磁化强度与磁场强度的比值被定义为磁化率 χ , 即 $\chi = \frac{M}{H}$ 。磁感应强度 B 及磁导率 μ 则由下二式决定:

在 SI 制中: $B = \mu_0 (H + M) = \mu_0 \mu H$, $\mu = 1 + \chi$

在 CGS 制中: $B = H + 4\pi M = \mu H$, $\mu = 1 + 4\pi\chi$

两个单位制中 μ 的数值相同, SI 制中的 χ 比 CGS 制中的 χ 值大 4π 倍, 在 SI 制中, H 的单位为 A/m, M 的单位与 H 相同。 B 的单位为特斯拉

(T)。在 CGS 制中, H 的单位为 Oe, B 和 M 的单位为 Gs, 量纲相同。1A/m(H) = $4\pi \times 10^{-3}$ Oe, 1T = 10^4 Gs, 1A/m(M) = 10^{-3} Gs, $\mu_0 = 4\pi \times 10^{-7}$ H/m, 为真空磁导率。

物质的磁性大体可分为五类, 即抗磁性、顺磁性、反铁磁性、铁磁性及亚铁磁性。前三种磁性很弱, χ 约为 $10^{-3} \sim 10^{-6}$ 。后两种为强磁性, χ 约为 $10 \sim 10^6$ 。强磁性广泛用于工程技术中。

(1) 抗磁性 大多数有机和无机材料均呈现抗磁性, 抗磁性物质的磁化率为负, 其磁化强度 M 与磁场 H 反向。

抗磁性的主要机制为局域电子抗磁性。可用经典理论描述。电子的轨道矩在磁场作用下作以磁场方向为轴的进动, 如图 10-1。这种磁场感应的附加电子运动产生的磁矩总是与磁场反向, 因而产生抗磁性。

所有材料中的局域电子都有这种抗磁性。经计算, 局域电子抗磁磁化率为

$$\chi_d = -\frac{\mu_0 n e^2}{6 m c} \sum_i \overline{r_i^2} \quad (10-24)$$

式中, n 为单位体积的原子数, 将电子的 e 和 m 值代入, 轨道半径 r_i 的量级约为 10^{-10} m, 可得 $\chi_d \approx 10^{-6} Z$ 。可以看出, 磁化率随局域电子数 Z 和电子轨道半径而增大, 且基本上不随温度变化。

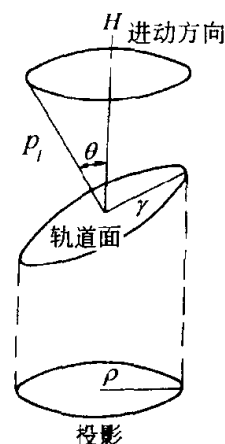


图 10-1 电子轨道矩的进动

抗磁性的另一种机制是传导电子抗磁性, 普遍存在于金属中, 它来源于传导电子在磁场作用下回旋运动的量子效应——朗道能级分裂。经计算传导电子抗磁磁化率为

$$\chi_{de} = -\frac{1}{2} \frac{n \mu_B^2}{k_B T_F} \left(\frac{m}{m^*} \right)^2 \quad (10-25)$$

式中, T_F 为费密温度, 费密面能量为 $E_F = k_B T_F$, k_B 为波尔兹曼常数, m 和 m^* 为自由电子质量和能带电子有效质量。对于正常金属, 如铜, m/m^* 约等于 1, 可估计出 χ_{de} 约为 10^{-6} 。但对于能带将满的金属, 如铋, m/m^* 很大, 因而得到很大的抗磁磁化率。在低温高场下, 传导电子抗磁磁化率随磁场强度而振荡, 称为 de Haas-Van Alphen (DHVA) 效应, 这是因为磁场引起朗道能级分裂后, 费米面能量随磁场而振荡之故。因此 DHVA 效应成为研究金属中费密面形貌的重要手段之一。

虽然抗磁性普遍存在于所有材料中, 但由于其磁化率很小, 当材料的原子、离子或分子有固有磁矩时, 顺磁磁化率掩盖了抗磁磁化率。只有无固有磁矩或固有磁矩很少的材料抗磁性才能表现出来。抗磁磁化率基本上不随磁

场和温度变化。

(2) 顺磁性 顺磁性物质的磁化率为正, M 与 H 同向。顺磁性有三种不同的来源。

① 局域电子顺磁性。来源于磁场对原子或离子固有磁矩的取向作用。由于磁场作用能近似为 $H\mu_B$, 远小于常温下的热运动能 $k_B T$, 因而顺磁磁化率 χ 也很小, 约 $10^{-3} \sim 10^{-5}$ 。在常温 and 一般磁场下不随磁场变化, 而随温度上升而减少。当材料中含有的磁性离子少、相互作用微弱时, 常服从居里定律:

$$\chi = \frac{C}{T} \quad (10-26)$$

当相互作用不可忽略时, 常服从居里-外斯定律,

$$\chi = \frac{C}{T - T_p} \quad (10-27)$$

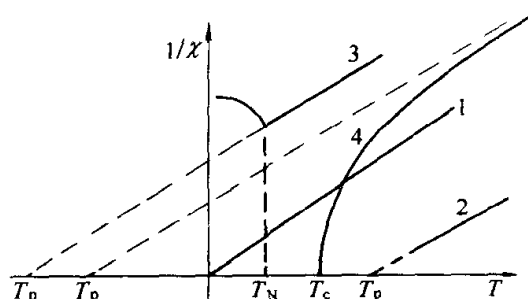


图 10-2 几种局域电子磁化率与温度的关系示意图

1—居里定律; 2—居里-外斯律;

3、4—反铁磁体和亚铁磁体

在相变点以上的 $\frac{1}{\chi} - T$ 关系

T_p 可能来源于交换作用、偶极子相互作用或晶体电场作用。图 10-2 为上二式代表的 $\frac{1}{\chi} - T$ 关系和反铁磁体及亚铁磁体在相变点以上的 $\frac{1}{\chi} - T$ 关系的示意图。

根据统计物理计算,

$C = ng_J^2 J(J+1) \mu_B^2 / 3k_B$ 。其中 n 为磁性离子密度, $g_J \sqrt{J(J+1)} \mu_B$ 为离子固有磁矩的绝对值。

② 传导电子顺磁性。普遍存在于金属材料中, 金属材料导带中正负自旋的电子在磁场作用下能量不一样, 相当于正负自旋能带发生劈裂。平衡时正负自旋电子数不等, 从而出现磁化, 根据自由电子近似的计算, 泡利给出 0K 时的磁化率的表示式为

$$\chi_{ep} = \frac{3}{2} \frac{n \mu_B^2}{k_B T_F} \quad (10-28)$$

与式 (10-25) 比较, 自由电子顺磁磁化率为自由电子抗磁性磁化率绝对值的 3 倍。因此, 自由电子顺磁性比抗磁性强, 总的表现为顺磁性。一般情况, 传导电子的总磁化率近似地可写为

$$\chi_e = \chi_{ep} + \chi_{ed} = \frac{3}{2} \frac{n \mu_B^2}{k_B T_F} \left[1 - \frac{1}{3} \left(\frac{m}{m^*} \right)^2 \right] \quad (10-29)$$

当电子的有效质量 $m^* < m/\sqrt{3}$ 时, 传导电子磁化率为负, 表现为抗磁性,

反之为顺磁性。半金属铋、锑等属于前者，正常金属（包括碱金属，碱土金属和铝）和过渡金属属于后者。但铜、金、银等金属的传导电子磁化率虽然也表现为顺磁性，由于内层局域电子抗磁性更大，因而它们为抗磁性金属。自由电子顺磁磁化率几乎不随温度变化，但过渡金属的顺磁磁化率由于能带的复杂影响，具有复杂的温度依赖性，有的随温度上升而减小，有的相反，如图 10-3 所示。

③ Van Vleck（范弗来克）顺磁性。它来源于磁场对电子云的形变，即二级微扰使激发态混入基态，使电子态发生微小的变化所致。Van Vleck 顺磁性存在于局域电子及能带电子中。它常是对顺磁性及抗磁性的一个修正，且基本不依赖于温度。

无磁序材料的磁性常为上述五种磁性中几种的叠加。非金属性材料的磁性中可能包含局域电子抗磁性和 Van Vleck 顺磁性。如组成中包含了具有固有磁矩的离子，如过渡元素离子，则还叠加有局域电子顺磁性。对于金属材料，还有传导电子的顺磁及抗磁性的叠加。

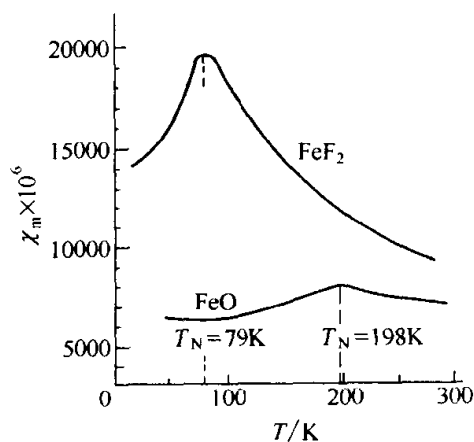


图 10-4 FeF_2 和 FeO 的摩尔磁化率与温度的关系

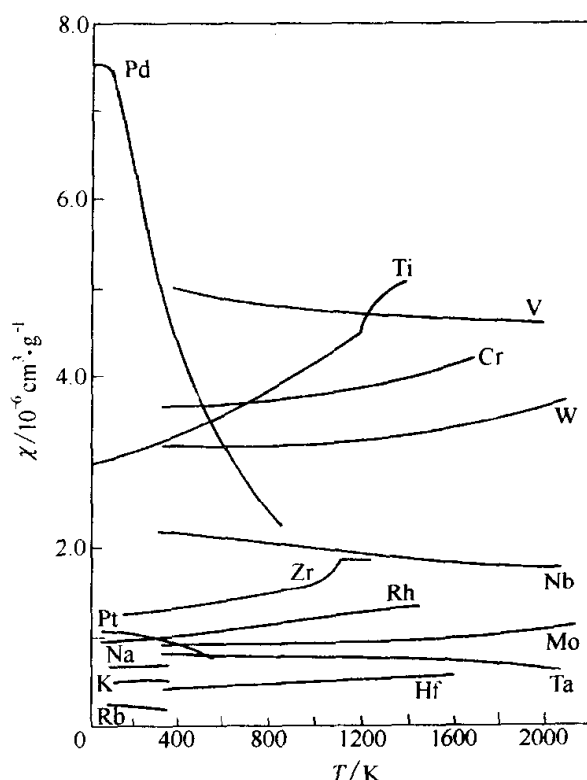


图 10-3 一些金属单位质量顺磁磁化率的温度依赖性

(3) 反铁磁性 反铁磁性也属于弱磁性，其磁化率为正，约为 10^{-3} 。反铁磁材料具有磁有序相变点，即奈耳点 T_N 。在 T_N 以上，呈现顺磁性，在 T_N 以下反铁磁体的原子或电子磁矩自发有序排列，或作反平行排列，或取特殊的排列，如螺旋形排列（Eu, MnAu_2 等），电子自旋密度在空间振动（Cr 中的自旋密度波）。故不加磁场时，其磁化强度为零。在磁场作用下其磁化率仍很小且随温度下降而减小。在 T_N 处出现极大值，如图 10-4

所示。

上述三种磁性均为弱磁性，在通常温度和磁场下磁化曲线为直线， χ 为常数，故可用磁化率 χ 表征其磁性。但在低温和强场下，顺磁和反铁磁材料可有非线性磁化曲线，甚至趋于饱和。

(4) 铁磁性和亚铁磁性 铁磁性材料和亚铁磁性材料都有磁有序相变点，称为居里点 T_c 。当温度高于 T_c 时，呈顺磁性，当温度低于 T_c 时呈铁磁性或亚铁磁性，均为强磁性。强磁性的特点是其磁化率远高于弱磁性的磁化率， χ 约为 $10^0 \sim 10^5$ 。其磁化曲线呈非线性，较易于达到磁饱和，磁化率随磁场而变，具有磁滞现象。图 10-5 及图 10-6 是代表性的强磁性磁化曲线、磁滞回线及磁导率与磁场的关系。强磁性的上述特征来源于其自发磁化。在温度低于 T_c 时，铁磁物质中的原子磁矩自发平行，或自发地使能带中正负自旋的电子数不相等，故不加磁场时，出现了不等于零的自发磁化强度 M_s 。亚铁磁性物质常为化合物或合金，在 T_c 温度以下时，原子磁矩呈反铁磁性排列，即自发的反平行排列或其他类型的自旋排列，但是或由于原子磁矩大小不等，或由于排列方式的特殊性，它们的磁矩没有抵消为零。其未抵消的净磁矩便导致了自发磁化强度。自发磁化常出现在大块强磁体内的微小区域中，称为磁畴。未加磁场而强磁体呈磁中性时，各磁畴中的自发磁化矢量 M_s 的分布使合成的总磁化强度为零。加入磁场后，磁畴重新调整，导致了强磁物质的磁化及磁化曲线、磁滞回线等特征。

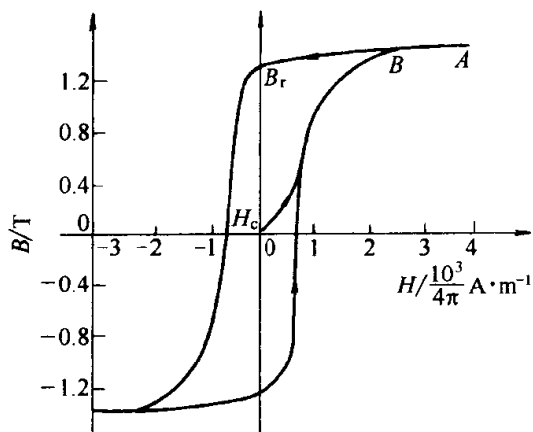


图 10-5 铁的磁化曲线和磁滞回线

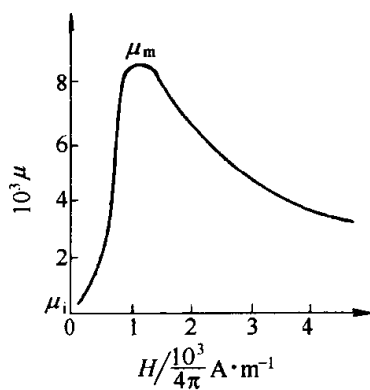


图 10-6 铁的磁导率对磁场的依赖性
(μ_i 为初始磁导率； μ_m 为最大磁导率)

(5) 周期表中各元素的磁性。表 10-1 给出周期表中各元素室温下的克原子磁化率。

各元素的磁性可以根据上述材料各种磁性的原理作定性的理解和讨论^[2,10]。

表 10-1 元素原子磁化率实验数据

	I A	II A	III B	IV B	V B	VI B	VII B	VIII B		
1	H ¹ -3.94									
2	Li ³ 25.2	Be ⁴ -9.02								
3	Na ¹¹ 15.6	Mg ¹² 6								
4	K ¹⁹ 21.5	Ca ²⁰ 44	Sc ²¹ 315	Ti ²² 150	V ²³ 230	Cr ²⁴ 160	Mn ²⁵ 527	Fe ²⁶ —		
5	Rb ³⁷ 19.2	Sr ³⁸ 92	Y ³⁹ 191	Zr ⁴⁰ 120	Nb ⁴¹ 120	Mo ⁴² 54	Tc ⁴³ 270	Ru ⁴⁴ 44		
6	Cs ⁵⁵ 29.9	Ba ⁵⁶ 20	La 系 ^① (57~71)	Hf ⁷² 70	Ta ⁷³ 145	W ⁷⁴ 40	Re ⁷⁵ 68.7	Os ⁷⁶ 7.6		
7	Fr ⁸⁷	Ra ⁸⁸	Ac 系 ^② (89~103)	Ku ¹⁰⁴	Ha ¹⁰⁵					
	VII B	VIII B	I B	II B	III A	IV A	V A	VI A	VII A	VIII A
1										He ² -1.9
2					B ⁵ -6.7	C ⁶ -42	N ⁷ -12	O ⁸ 3450	F ⁹	Ne ¹⁰ -10.7
3					Al ¹³ 16.7	Si ¹⁴ -3.6	P ¹⁵ -20, -27	S ¹⁶ -16, -15	Cl ¹⁷ -40	Ar ¹⁸ -18
4	Co ²⁷ —	Ni ²⁸ —	Cu ²⁹ -5.4	Zn ³⁰ -10.3	Ga ³¹ -16.8	Ge ³² -8.9	As ³³ -5.5	Se ³⁴ -26.5	Br ³⁵ -56	Kr ³⁶ -37
5	Rh ⁴⁵ 113	Pd ⁴⁶ 580	Ag ⁴⁷ -21.56	Cd ⁴⁸ -19.6	In ⁴⁹ -12.6	β-Sn ^{50③} 4.4	Sb ⁵¹ -107	Te ⁵² -40.8	I ⁵³ -89	Xe ⁵⁴ -59
6	Ir ⁷⁷ 25	Pt ⁷⁸ 200	Au ⁷⁹ -29.6	Hg ⁸⁰ -33.8	α-Tl ⁸¹ -44.0	Pb ⁸² -24.9	Bi ⁸³ -285	Po ⁸⁴	At ⁸⁵	Rn ⁸⁶
7										

① La 系

La ⁵⁷	Ce ⁵⁸	Pr ⁵⁹	Nd ⁶⁰	Pm ⁶¹	Sm ⁶²	Eu ⁶³	Gd ⁶⁴	Tb ⁶⁵	Dy ⁶⁶	Ho ⁶⁷	Er ⁶⁸	Tm ⁶⁹	Yb ⁷⁰	Lu ⁷¹
130	2100	3530	5180	—	1320	3340	—	1720	70700	77100	50100	26700	81	17

② Ac 系

Ac ⁸⁹	Th ⁹⁰	Pa ⁹¹	U ⁹²	Np ⁹³	Pu ⁹⁴	Am ⁹⁵	Cm ⁹⁶	Bk ⁹⁷	Cf ⁹⁸	Es ⁹⁹	Fm ¹⁰⁰	Md ¹⁰¹	No ¹⁰²	Lr ¹⁰³
—	130	601	110	—	600	—	—	—	—	—	—	—	—	—

③ α-Sn 的 $\chi_a = -31.5 \times 10^{-6}$ 。

注：表中为原子磁化率 $\chi_a \times 10^6$ 。

从周期表的左侧 I A 和 II A 族开始，跨过过渡族 III B 至 VIII B，直至 IV A 和 V A 的一部分，为金属材料及半金属。它们的磁性来源于正离子满壳层的局域电子抗磁性 χ_{id} 和传导电子顺磁性 χ_{ep} 与抗磁性 χ_{ed} 的竞争， $\chi = \chi_{id} + \chi_{ep} + \chi_{ed}$ 。当 $\chi_{ep} > |\chi_{id} + \chi_{ed}|$ 时，呈现顺磁性，反之呈抗磁性。碱金属、碱土金属和过渡金属除铍外均为顺磁性。前两族 χ 与温度的依赖性微弱，近于

自由电子顺磁性。过渡金属中 d 带电子状态随原子序数作周期变化, 使其 χ_{ep} 的数值作周期变化, 温度系数或正或负, 如图 10-3 所示。I B 族铜、银、金以右的金属中, 正离子局域抗磁性占据优势, $|\chi_{id}| > \chi_{ed} + \chi_{ed}$, 除铝和白锡(β)外均呈抗磁性。半金属、铋、铊则由于有效质量 m^* 特别小, 有大的传导电子抗磁性。IV A 和 V A 中的半导体及其右侧的非金属元素直至惰性气体, 或为磁矩为零的共价键固体、液体或气体, 或为满壳层的原子, 均呈现局域电子抗磁性。在过渡族中, 铬与锰为反铁磁元素。铁、钴、镍为铁磁元素。稀土族中多数在低温下有复杂的磁有序相变及复杂的磁结构。含有这些元素的多种铁磁与亚铁磁合金和化合物在工业技术上有重要应用, 其磁有序的成因和强磁性原理将在下面进一步讨论。

10.2 磁有序与交换作用^[2,6]

10.2.1 铁磁有序

铁磁材料均有一相变温度, 即居里点 T_c 。当温度高于 T_c 时, 铁磁材料中的原子或电子磁矩常为无序状态, 材料呈顺磁性。当 $T < T_c$ 时, 出现磁有序, 呈铁磁性。

铁磁材料的顺磁磁化率近似地服从居里-外斯定律。实际上只有温度远较 T_p 高时, $\frac{1}{\chi} \sim T$ 关系才大体为直线性。几个例子示于图 10-7。

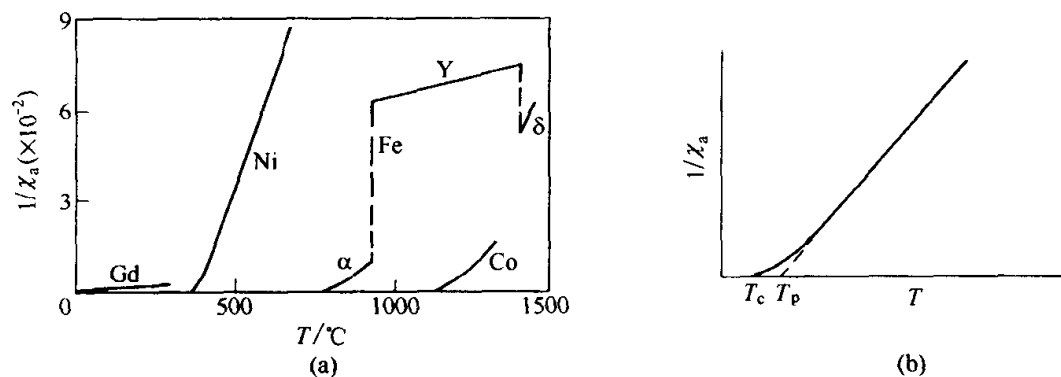


图 10-7 一些铁磁金属的 $\frac{1}{\chi} \sim T$ 关系曲线

(a) Fe、Co、Ni 和 Gd 的 $\frac{1}{\chi} \sim T$ 的关系 (χ_a 为克原子磁化率);

(b) $\frac{1}{\chi_a} \sim T$ 关系在居里点附近示意图

从图中还可看到, 结构相变使 χ 发出跳跃。当温度低于 T_c 时, 铁磁状态下的磁化率远高于顺磁状态, 且易于饱和, 如图 10-5 和图 10-6 所示。饱和磁化强度 M_s 随材料而异, 但均为温度的函数, M_s 随温度上升而下降。在绝对零度达最大值 M_0 。趋于居里温度时, M_s 趋于零。示于图 10-8。

早在 1907 年外斯 (P. Weiss) 提出了分子场理论以解释铁磁性。此理论

假设铁磁体内存在着分子场，在 T_c 以下引起自发磁化。通常自发磁化发生在微小的磁畴内。外磁场使磁畴结构调整从而发生技术磁化直至饱和。

(1) 局域磁矩铁磁性 外斯的理论是局域磁矩顺磁性模型的发展，设原子磁矩间存在着的相互作用等效为一个分子场 H_λ ，且

$$H_\lambda = \lambda M \quad (10-30)$$

式中， λ 称为分子场常数， $\lambda > 0$ ，意味着分子场的方向与 M 平行，使原子磁矩平行。当外磁场 H 不为零时，作用在原子磁矩上的有效场大于外场，

$$H_{\text{eff}} = H + \lambda M \quad (10-31)$$

与处于基态 J 的原子磁矩的作用能为

$$E_H = -g_J m_J \mu_B H_{\text{eff}} \quad (10-32)$$

材料的宏观磁化强度 M 决定于磁场取向作用与热运动的平衡。根据玻尔兹曼统计，配分函数为

$$Z = \left[\sum_{-J}^J e^{m_J x} \right]^N$$

式中， $x = g_J \mu_B H_{\text{eff}} / k_B T$ ， N 为单位体积中原子磁矩数。

$$M = k_B T \frac{\partial}{\partial H} \ln Z$$

$$M = M_0 B_J(\alpha_J) \quad (10-33)$$

式中， $M_0 = N g_J \mu_B$ ， $\alpha_J = \frac{g_J \mu_B H_{\text{eff}}}{k_B T}$ ，

$$B_J(\alpha_J) = \frac{M}{M_0} = \frac{2J+1}{2J} \text{cth} \left(\frac{2J+1}{2J} \right) \alpha_J + \frac{1}{2J} \text{cth} \frac{\alpha_J}{2J} \quad (10-34)$$

称为布里渊 (Brillouin) 函数。对布里渊函数的分析可得到以下结果。

① $T > T_c$ ，为顺磁状态。这时只有 $H \neq 0$ 时，式 (10-33) 才不等于零。当 $\alpha_J \ll 1$ 时，取 $B_J(\alpha_J)$ 的近似表示，可得到顺磁磁化率的居里-外斯定律，即式 (10-27)

$$\chi = \frac{C}{T - T_p}$$

其中 $C = N g^2 J(J+1) \mu_B^2 / 3 k_B$ ， $T_p = \lambda C$ (10-35)

可以从 $\chi \sim T$ 的实验关系来确定居里常数 C 和顺磁居里点 T_p 和分子场

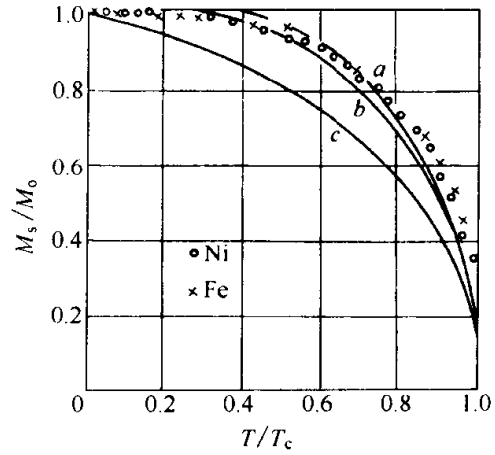


图 10-8 局域分子场理论给出的约化 $M_s \sim T$ 关系

$$a - J = \frac{1}{2}; \quad b - J = 1; \quad c - J = \infty$$

常数 λ , 从居里常数可以按下式确定原子磁矩的有效玻耳磁子数

$$p_{\text{eff}}^2 = \frac{3k_B C}{N\mu_B^2} = g^2 J(J+1) \quad (10-36)$$

② $T < T_c$, 为铁磁状态。这时没有外场, $H = 0$ 时, 式 (10-33) 表示的 $M \neq 0$ 。分子场引起了自发磁化。自发磁化强度即是铁磁体的饱和磁化强度 M_s 。而铁磁居里点 T_c 与 T_p 相同, 将 M_o 和下式

$$T_c = T_p = \lambda N g^2 J(J+1) \mu_B^2 / 3k_B \quad (10-37)$$

代入 α_J 和式 (10-33), 可得

$$\alpha_J = \frac{3J}{J+1} \left(\frac{M_s}{M_o} \right) \left(\frac{T}{T_c} \right) \quad (10-38)$$

$$\frac{M_s}{M_o} = B_J \left(\frac{3J}{J+1} \frac{M_s}{M_o} \right) \left(\frac{T}{T_c} \right) \quad (10-39)$$

图 10-8 给出了不同 J 值时的约化 $\frac{M_s}{M_o} \sim \frac{T}{T_c}$ 的关系及铁和镍的实验数据。实验与理论大体相符。当 $T = 0\text{K}$ 时,

$$M_{s(T=0)} = M_o = N g J \mu_B = N p_s \mu_B \quad (10-40)$$

p_s 称为原子饱和玻耳磁子数, 表 10-2 给出了 Fe、Co、Ni 和 Gd 的实验数据。

表 10-2 几种铁磁材料的数据

材料	T_c/K	T_p/K	$M_o/10^{-3}\text{A}\cdot\text{m}^{-1}$	$p_s(\mu_B)$	$p_{\text{eff}}(\mu_B)$	p_s/H	p_{eff}/H
铁	1043	1101	1743	2.216	3.13(4.2)	4	4.90
钴	1395	1428	1447	1.715	3.13	3	3.87
镍	631	650	521	0.606	1.62	2	2.83
钆	293.2	317	1980	7.55	7.98	7	7.94

对表中数据有以下几点值得注意。a. $T_c < T_p$, 与外斯分子场理论不一致。可以用居里点附近近程有序形成有序团簇的模型给予解释^[1]。b. 铁、钴、镍的 p_s 和 p_{eff} 值远小于按洪德法则并假设轨道完全被淬灭的计算值 $p_s(H)$, 且 $p_s(H)$ 为整数而实验值 p_s 远非整数, 此外从 p_{eff} 估算出的 gJ 大于从 p_s 估算的数值。这些是用局域磁矩理论描述过渡族金属铁磁性时出现的严重问题。许多事实说明, 过渡族金属中 $3d$ 电子为非局域的。在巡游电子铁磁性理论中上述问题将很自然地得到解决。多数稀土族元素的 $4f$ 壳层不满, 因而有固有原子磁矩。稀土金属的 $4f$ 电子轨道半径远小于原子间距的一半, 故 $4f$ 电子为典型的局域电子。其中钆具有简单的铁磁/顺磁转

变, 其他具有复杂的磁序转变。在表 10-2 中给出的钐的 p_s 和 p_{eff} 与洪德法则的计算值接近但略高。这来源于 $4f$ 外部的 $5d$ 和 $6s$ 电子的微小贡献。钐的 $M_s \sim T$ 和 $\chi \sim T$ 的实验数据与局域理论相符得很好。从居里点和居里常数的实验值可以求出分子场常数 λ 和分子场的数值, 分子场的估算值约为 $10^7 \times \frac{10^3}{4\pi} \text{A/m}$ 或 10^7Oe , 远高于实验室产生的磁场。分子场并非真正的磁场而是静电性交换作用的等效磁场。局域磁矩分子场理论给出的铁磁性出现的条件如下:

- 材料中含有磁矩不等于零的原子、离子或分子, 即 $S \neq 0$;
- 存在着正的分子场, 亦即正的交换作用。

(2) 巡游电子铁磁性, Stoner 模型 过渡族元素原子性的 $3d$ 和 $4s$ 电子能级在金属中成为共有的 $3d$ 和 $4s$ 能带, 且互相重叠。电子按泡利原理在其中重新分布。在 T_c 以下铁、钴、镍的分子场或交换作用足以使正负自旋的 $3d$ 次带自发分裂, 电子进一步重新分布, 正负自旋电子数不等导致了自发磁化。图 10-9 为镍的 $3d$ 和 $4s$ 能带交换劈裂的示意图, 其中 $4s$ 带的劈裂被忽略。几个铁族金属能带电子的分布示于表 10-3。铁、钴、镍的饱和玻尔磁子数 p_s 的非整数性等问题自然地得到解释。

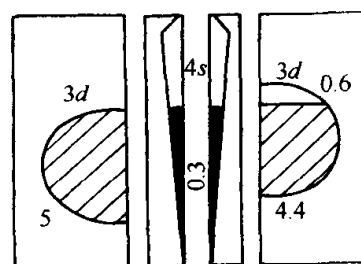


图 10-9 镍的能带
交换劈裂的示意图

表 10-3 几个铁族金属中 $3d$ 和 $4s$ 带电子分布

元素	孤立原子 电子组态	按能带模型电子分布				空穴数		未抵消 自旋数	p_s 实验值
		$3d^+$	$3d^-$	$4s^+$	$4s^-$	$3d^+$	$3d^-$		
Cr	$3d^4 4s^2$	2.7	2.7	0.3	0.3	2.3	2.3	0	0
Mn	$3d^5 4s^2$	3.2	3.2	0.3	0.3	1.8	1.8	0	0
Fe	$3d^6 4s^2$	4.8	2.6	0.3	0.3	0.2	2.4	2.2	2.216
Co	$3d^7 4s^2$	5.0	3.3	0.35	0.35	0	1.7	1.7	1.715
Ni	$3d^8 4s^2$	5.0	4.4	0.3	0.3	0	0.6	0.6	0.606
Cu	$3d^{10} 4s^1$	5.0	5.0	0.5	0.5	0	0	0	0

按 Stoner 模型, 过渡金属铁磁性出现的条件, Stoner 判据为

$$\xi = \lambda \mu_B^2 N(E_F) = \epsilon_{\text{ex}} N(E_F) > 1 \quad (10-41)$$

式中, λ 为分子场常数, ϵ_{ex} 为交换常数。因此巡游电子铁磁性要求①足够大的正分子场常数或交换常数; ②费密面处电子态密度足够大。这是由于自

发磁化后, 交换能降低, 但费密面能提高。只有满足式 (10-41) 才能使总能量降低。对周期表中一些金属的交换常数和费密面态密度的理论计算指出铁、钴、镍具有不太高的交换常数, 但 $N(E_F)$ 高, 因而为铁磁金属, 碱金属和碱土金属具有更高的交换作用, 但 $N(E_F)$ 很小, 因而不出现铁磁性^[2]。

10.2.2 反铁磁和亚铁磁性

(1) 概况 反铁磁材料在 T_N 以上为顺磁态, 在 T_N 以下呈现反铁磁有序状态。反铁磁状态下, 原子或电子磁矩的空间分布呈反平行或其他类型排列, 但宏观自发磁化强度为零。磁化率很小, 甚至小于顺磁态下的数值, 在 T_N 点时形成极大。反铁磁有序大体出现在三类材料中。①过渡元素化合物。②过渡金属与合金和③稀土金属。表 10-4 给出一些代表性的反铁磁材料及其数据。稀土金属多具有不只一个相变及复杂的磁结构, 将另行讨论。

表 10-4 一些反铁磁体及其数据

物 质	奈耳点 T_N/K	顺磁居里点 T_p/K	$\chi(0)/\chi(T_N)$	顺磁离子的晶格
MnO	122	-610	2/3	面心立方
MnS	165	-528	0.82	面心立方
MnSe	247	-748		面心立方
MnTe	307			六角
MnF ₂	72	-113	0.76	体心正方
FeF ₂	79	-117	0.72	体心正方
FeCl ₂	23.5	48	<0.2	六角
FeO	198	-570	0.8	面心立方
CoCl ₂	24.9	38.1	约 0.6	六角
CoO	291	-280, -330		面心立方
NiO	520, 523, 647	-1310, -2470; -3000	0.67	面心立方
NiCl ₂	49.6	68.2		六角
CrSb	690	-200		六角
Cr-Al	600			
Cr-Ga	640			
Cr ₂ O ₃	310	-485	0.76	六角
CrCl ₂	约 40	-14.9		
VCl ₃	30	-30.1		
TiCl ₂	约 100			
CuCl ₂	约 70	-109		
α -Fe ₂ O ₃	950	-2000		六角
α -Mn	95			复杂

续表

物 质	奈耳点 T_N/K	顺磁居里点 T_p/K	$\chi(0)/\chi(T_N)$	顺磁离子的晶格
γ -Mn	约 390			面心四方
Cr	312			体心立方, 单晶
Cr	约 500			体心立方, 粉末
γ -Fe	约 40			面心立方
MnAu ₂	363			
Pt ₃ Fe	约 100			
MnFe	490			

亚铁磁材料在其 T_c 以上常呈现顺磁性, 在 T_c 以下为亚铁磁有序状态。亚铁磁状态下, 磁矩空间分布与反铁磁相似, 呈反平行或其他类型排列, 但相互未抵消, 宏观自发磁强度不等于零, 因而呈现强磁性。其技术磁性与铁磁性相似。亚铁磁有序出现在两大类材料中。其一为过渡金属化合物, 如氧化物、硫化物等。其中铁氧体在工程技术中有广泛的用途。另一类为稀土-过渡族合金。在高性能永磁及磁光存储上有重要应用。

(2) 反铁磁性和亚铁磁性的分子场解释 反铁磁性和亚铁磁性及其分子场理论是由奈耳首先给出的。在相变温度 T_N 或 T_c 以下, 原子或电子磁矩自发地呈现反铁磁排列。最常见的是反平行排列。具体的排列方式依赖于原子间交换作用和晶体结构。根据反铁磁排列的特点可将晶格分成两个以上次晶格。将局域铁磁性分子场理论推广, 同一次晶格中的磁矩平行排列, 有似于铁磁性的自发磁化。不同次晶格中的磁性自发平行于不同方向, 例如, 可以相互反平行, 亦可平行, 甚至呈某个角度。次晶格内及不同次晶格间的交换作用可用不同的分子场表示, 取 $-\lambda_{ij}M_j$ 的形式。即 j 次晶格产生的作用于 i 次晶格原子磁矩的分子场。总等效场为外场 H 与分子场的总和。

$$H_i = H - \sum_j \lambda_{ij} M_j \quad (10-42)$$

式中, λ_{ij} 为分子场常数, $i=j$ 表示同一次晶格内交换作用的等效分子场, $i \neq j$ 表示不同次晶格间的分子场, $\lambda_{ij} > 0$ 表示导致反平行的分子场, $\lambda_{ij} < 0$ 表示导致相互平行或铁磁耦合的分子场。与铁磁性分子场相似, 每个次晶格 i 的磁化强度 M_i 为

$$M_i = M_{i0} B_{J_i}(\alpha_i) \quad (10-43)$$

其中

$$M_{i0} = n_i g_{J_i} J_i \mu_B \quad (10-44)$$

$$\alpha_i = g_{J_i} J_i \mu_B \frac{H - \sum_j \lambda_{ij} M_j}{k_B T} \quad (10-45)$$

总磁化强度为

$$\mathbf{M} = \sum_i \mathbf{M}_i \quad (10-46)$$

从式(10-43)可见,每个次晶格的磁化强度 M_{i0} 以与铁磁体相似的规律随温度而下降,如图 10-8 所示。总磁化强度为各次晶格 M_i 的矢量和。

在相变温度以上,反铁磁性或亚铁磁性均转变为顺磁状态。根据式(10-43)和式(10-46),考虑到 $\alpha_i \ll 1$, 取近似,可得到顺磁磁化率。反铁磁材料的顺磁磁化率多近似为居里-外斯定律(式 10-27)

$$\chi = \frac{C}{T - T_p}$$

但 T_p 常为负值,反映负的分子场常数。如图 10-2 中曲线 3 所示。对亚铁磁材料,奈耳给出了二次晶格亚铁磁体的表达式

$$\frac{1}{\chi} = \frac{T}{C} + \frac{1}{\chi_0} - \frac{\sigma}{T - \theta} \quad (10-47)$$

如图 10-2 中曲线 4,在高温下略去最后一项,得到渐近线,趋于居里-外斯定律。

在奈耳温度以下,当无外场作用时,反铁磁材料各次晶格的 M_i 相互抵消, $M_s = 0$, 无自发磁化。在外场作用下,式(10-43)给出各次晶格的 M_i 。因 H 常远小于 $\lambda_{ij}M_j$, 磁化强度及磁化率很小。对单晶反铁磁体,当 H 与反铁磁磁矩排列的方向垂直时,引起磁矩转动的磁化,其磁化率 χ_{\perp} 大于 H 平行于磁矩方向的磁化率 χ_{\parallel} 。图 10-10 为 MnF_2 单晶在自发反铁磁排列方向附近的磁化曲线。在低场时 χ_{\parallel} 很小,当 H 达到自旋突转 (spinflip) 场时,磁矩

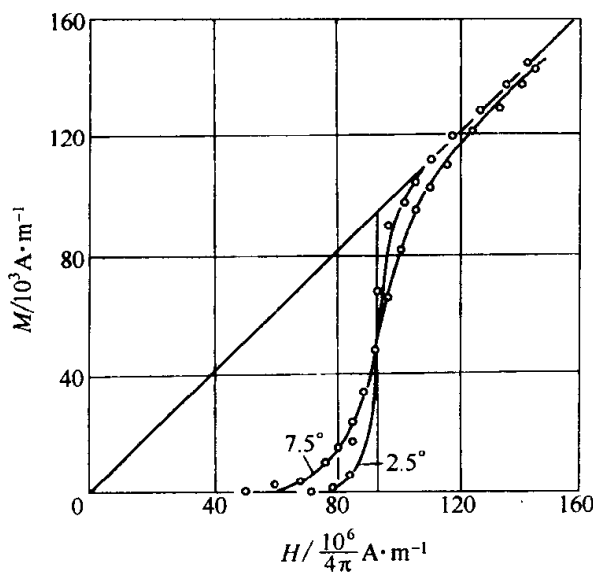


图 10-10 MnF_2 单晶磁化曲线

突转近 90° 。这时 χ_{\perp} 较大,使 M 陡升,然后逐渐转向磁场。

对于亚铁磁材料,在 $T < T_c$ 。且无外场时,各次晶格的自发磁化强度 M_i 不相互抵消,故有总自发磁化, $M_s \neq 0$ 。但由于不同温度下 M_i 的数值不一,抵消情况不一,总的 $M_s \sim T$ 可有多种类型,例如,除正常的 M_s 随温度上升而单调下降的 $M_s \sim T$ 关系外,还有具有抵消点的 N 型及反常的 P 型 $M_s \sim T$ 曲线。图 10-11 给出了

N 型的情况, 两反平行次晶格的 M_A 和 M_B 在抵消点 T_d 相互抵消, 使 $M_s(T_d)=0$; P 型的 M_B 随 T 上升而下降的速度比 M_A 慢。这类的 $M_s \sim T$ 在磁性氧化物或稀土过渡族合金中均被观察到, 后者用于磁光存储介质。除反平行磁结构外, 还有非共线磁结构, 反映了复杂的相互作用。稀土金属中有多种磁结构, 而且还有多种磁相变, 分别在图 10-12 和表 10-5 中给出。

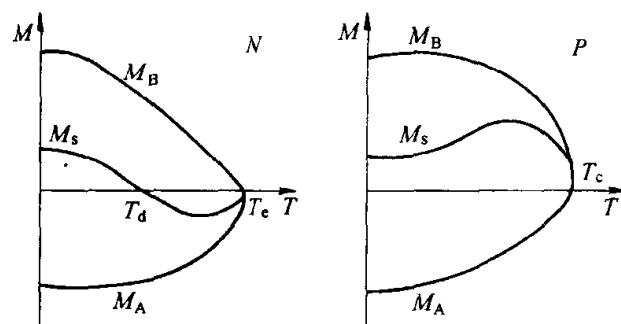


图 10-11 N 型及 P 型 $M_s \sim T$ 曲线

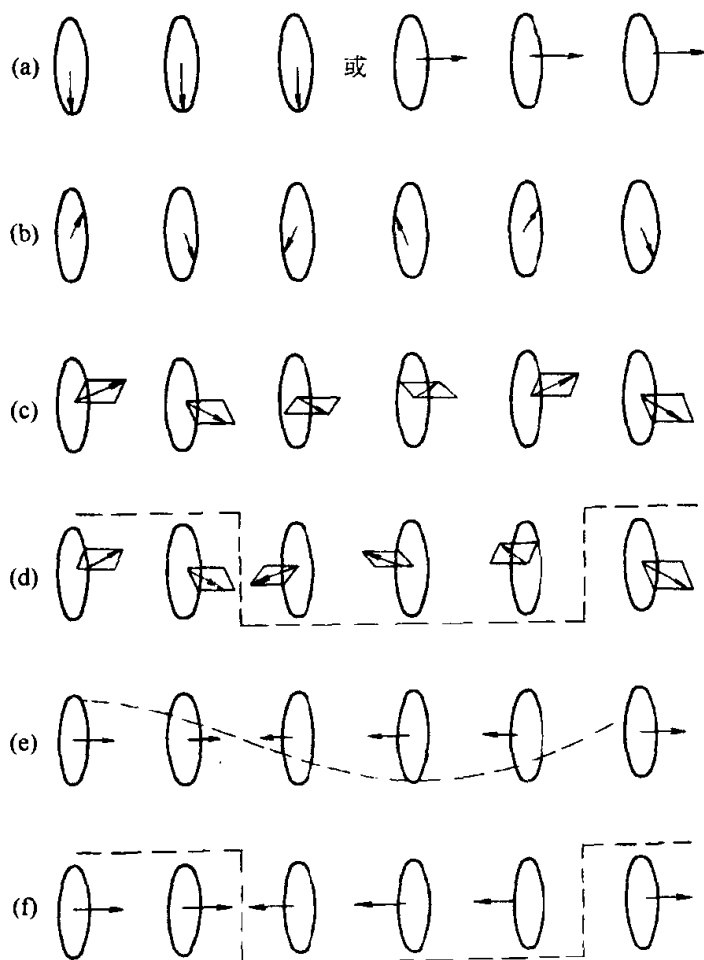


图 10-12 各种复杂的磁结构示意图

表 10-5 稀土金属的磁结构及转变温度^[15,16]

金属	磁结构和转变点 T_c 及 T_N (K)
Ce	反铁磁(c 平面内亚铁磁, 面间反平行) $< T_N = 12.5K >$ 顺磁
Pr	反铁磁(可能与 Nd 相似) $< T_N \approx 20K >$ 顺磁
Nd	铁磁或亚铁磁 $< T_c \approx 7.5K >$ 反铁磁(面内平行, 面间反铁磁) $< T_N = 19K >$ 顺磁
Sm	可能铁磁 $< T_c = 14K >$ 反铁磁(与 Nd 相似) $< T_N \sim 106K >$ 顺磁
Eu	反铁磁性(SSS) $< T_N = 90K >$ 顺磁
Gd	铁磁(平行 c 轴, $T > 248K$ 偏离 c 轴) $< T_c = 289 >$ 顺磁
Tb	铁磁(平行 c 面) $< T_c = 218K >$ 反铁磁(SSS) $< T_N = 230K >$ 顺磁
Dy	铁磁(平行 c 面) $< T_c = 90K >$ 反铁磁(SSS) $< T_N = 179K >$ 顺磁
Ho	铁磁(FS) $< T_c = 20K >$ 反铁磁(SSS) $< T_N = 133K >$ 顺磁
Er	铁磁(FS) $< T_c = 20K >$ 反铁磁(CS) $< T_{N1} = 52K >$ 反铁磁(SLSW) $< T_{N2} = 80K >$ 顺磁
Tm	铁磁(APD) $< T_c = 22K >$ 反铁磁(SLSW) $< T_N = 56K >$ 顺磁

注: 表中的符号分别为: SSS—简单螺旋结构; FS—铁磁螺旋结构或锥面结构; CS—复杂螺旋结构或反相锥面结构; SLSW—静态纵向自旋波结构或简谐波调制结构; APD—反相畴或方波调制结构, Tm 的结构是四层原子磁矩向上, 三层向下, 平均净磁矩为 $1\mu_B$ /原子。

局域分子场理论不适用于过渡金属的反铁磁性, 后者应考虑磁性电子的巡游性。典型的例子是金属铬, 在 $T_N = 311^\circ\text{C}$ 以下, 其 d 电子的自旋不是局域在晶体格点上的反平行结构, 而是反平行自旋密度作连续简谐变化的自旋密度波。

10.2.3 交换作用^[2]

交换作用是分子场的来源和本质。尽管分子场是以磁场的形式出现的, 但其本质不是磁的相互作用。原子间磁的相互作用的数量级为 $\frac{\mu_B}{a^3} \sim 10^3 \text{Oe}$, 远小于 Fe, Co, Ni 高达 10^7Oe 的分子场。分子场的本质为静电作用的量子效应。

(1) 交换作用的物理图像, 局域直接交换作用 经典的静电性相互作用不影响磁矩的取向。但量子力学证明, 静电性的交换作用与电子自旋的取向相关。这首先在氢分子中被发现。海森伯将其用于解释铁磁性的来源。因它来自相邻原子间的直接相互作用, 故称为海森伯直接交换作用。

根据泡利原理, 两个以上电子的波函数包含了电子在不同单电子态的交换, 因而引进了电子间的相关——两个电子不能处于同一状态。于是, 自旋相同的两个电子靠近的几率比反平行时的几率小, 称为泡利排斥或交换相关。

氢分子中两电子系的波函数用氢原子波函数的组合来表示, 两电子自旋

反平行及平行的波函数 Ψ_s 和 Ψ_t 分别为

$$\Psi_s = N_s [\phi_a(r_1)\phi_b(r_2) + \phi_a(r_2)\phi_b(r_1)] [\alpha(1)\beta(2) - \beta(1)\alpha(2)] \quad (10-48)$$

$$\Psi_t = N_t [\phi_a(r_1)\phi_b(r_2) - \phi_a(r_2)\phi_b(r_1)] \begin{cases} \alpha(1)\alpha(2) \\ \alpha(1)\beta(2) + \beta(1)\alpha(2) \\ \beta(1)\beta(2) \end{cases} \quad (10-49)$$

式中, $\phi_a(r_1)$ 为电子 1 处于原子态 ϕ_a 中, $\alpha(1)$ 为电子 1 的自旋向上, 其他可

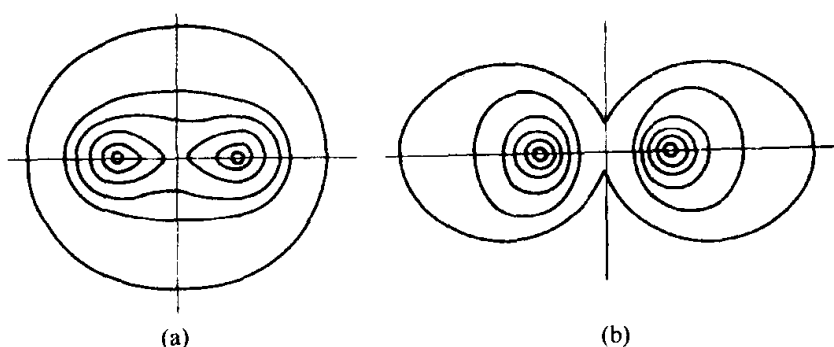


图 10-13 氢分子的两电子云分布

(a) 自旋反平行 Ψ_s^2 ; (b) 自旋平行 Ψ_t^2

类推。 Ψ_s 和 Ψ_t 相应的电子云空间分布示于图 10-13, 泡利排斥明显可见。因而使自旋态不同时电子间和电子与原子核间的静电作用能不同。它们分别是

$$E_s = 2E_0 + 2N_s^2(C + J) \quad (10-50)$$

$$E_t = 2E_0 + 2N_t^2(C - J) \quad (10-51)$$

右边 $2E_0$ 为两个氢原子的能量, 后两项为氢原子间相互作用能, 其中 C 相当于经典的库仑作用能, J 来自电子的交换项, 称为交换积分。

$$C = \int |\phi_a(r_1)|^2 \left[-\frac{e^2}{r_{a2}} - \frac{e^2}{r_{b1}} + \frac{e^2}{r_{12}} \right] |\phi_b(r_2)|^2 dv_1 dv_2 \quad (10-52)$$

$$J = \int \phi_a(r_1)\phi_b(r_2) \left[-\frac{e^2}{r_{a2}} - \frac{e^2}{r_{b1}} + \frac{e^2}{r_{12}} \right] \phi_a^*(r_2)\phi_b^*(r_1) dv_1 dv_2 \quad (10-53)$$

J 代表的交换作用能在经典理论中没有对应, 是静电作用的量子效应。以上采用原子波函数, 故为局域电子模型, J 依赖于原子波函数的重叠。

若 $J > 0$, 则式(10-51)表示的能量低, 基态的自旋平行, $J > 0$ 为铁磁性的必要条件。

若 $J < 0$, 则式 (10-50) 表示的能量低, 基态的自旋反平行。 $J < 0$ 为反铁磁性的必要条件。

在处理磁有序时, 常将上两式统一在一个表示式中而略去 E_0 及 C , 将交换能写为算符形式, 称为交换哈密顿量,

$$H_{\text{ex}} = -2 \sum_{i < j} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j \quad (10-54)$$

上式为电子间的交换哈密顿, 对整个多电子系求和。如不考虑原子内交换作用, 设 \mathbf{S}_i 和 \mathbf{S}_j 为 i, j 原子的自旋矢量, 则

$$H_{\text{ex}} = -2 \sum_{i < j} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j \quad (10-55)$$

设固体中由一种原子组成, $\mathbf{S}_i = \mathbf{S}_j = \mathbf{S}$, 单位体积中有 N 个原子, 原子的最近邻数为 Z , 且最近邻的交换积分均为 J , 非最近邻的原子间交换积分为零, 则

$$E_{\text{ex}} = -2JZ \frac{N}{2} S^2 = -NZJS^2$$

而局域磁矩分子场能量为 $E_{\lambda} = -\lambda M_0^2/2$ 。令 $E_{\text{ex}} = E_{\lambda}$, 若轨道完全淬灭, 并考虑 $M_0 = NgS\mu_B$ 及 $T_c = \lambda Ng^2 S(S+1) \mu_B^2 / 3k_B$, 可有

$$\lambda = \frac{2ZJ}{Ng^2 \mu_B^2} \quad (10-56)$$

$$T_c = \frac{2ZS(S+1)J}{3k_B} \quad (10-57)$$

上述局域电子直接交换作用是最早为解释过渡族金属铁磁性而提出的, 但以后的研究说明它并不适用于 Fe、Co、Ni 等的铁磁性。Fe、Co、Ni 中的直接交换作用的理论值远小于 T_c 的要求。而且尚未发现任何材料的铁磁性可以简单地用局域电子直接交换作用来解释。

(2) 过渡族化合物中的局域电子间接交换作用 许多过渡族化合物, 如铁氧体及 MnO 等氧化物或硫化物, 为亚铁磁及反铁磁材料。其磁性阳离子被基态无磁矩的阴离子隔开。近邻阳离子的 d 电子叠交很小, 直接交换作用很微弱。与阳离子电子云有叠交的阴离子中的电子为满壳层, 如 O^{2-} 为 $2p^6$ 。式 (10-55) 的交换能为零, 因此不能用直接交换作用来解释这类材料的磁有序。

这类材料中的磁有序是通过局域电子间接交换作用实现的, 对于不同的材料有超交换作用, 双交换作用及共价键交换作用等。它们的共同特点是近邻的阳离子与阴离子间的相互作用使阴离子出现激发态的混入, 出现了电子的非局域化或共价键, 从而出现了较大的交换作用。图 10-14 为 MnO 中四电子系统的超交换作用的示意图。图 10-14(a) 为基态, 交换作用为零。图

10-14 (b) 为一种激发态, 氧离子的一个 p 电子跃迁到相邻的锰中成为 d' , 形成 $\text{Mn}^+ - \text{O}^- - \text{Mn}^{2+}$ 。由于洪德法则 Mn^+ 中的 d' 电子与原有的 5 个 d 电子处于低能的自旋反平行态。 O^- 中剩余的一个不成对的 p 电子与另侧 Mn^{2+} 的直接交换作用为负, 这个 Mn^{2+} 的 5 个 d 电子 ($S = \frac{5}{2}$) 与 p 电子的自旋反平行时为低能。总的结果是锰离子自旋反平行时能量低, 或总交换常数 $J_i < 0$, 经二级微扰计算, 可得到与式 (10-55) 相似的表示

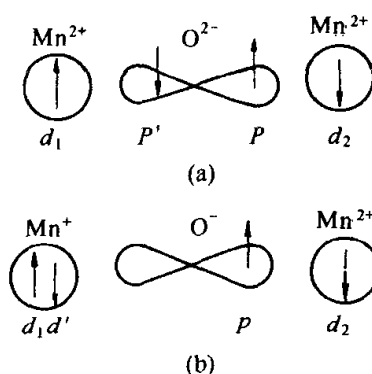


图 10-14 MnO 四电子系的超交换作用示意图

$$H_{\text{ex}} = -J_i \mathbf{S}_i \cdot \mathbf{S}_j \quad (10-58)$$

式中, J_i 为等效交换积分

$$J_i = \frac{1}{2} \rho^2 J_d \left[\frac{1}{(\Delta E_1)^2} - \frac{1}{(\Delta E_3)^2} \right] \quad (10-59)$$

式中, ρ 为 p 电子激发为 d 电子的跃迁积分, J_d 为 O^- 与 Mn^{2+} 的直接交换积分, ΔE_1 和 ΔE_3 为 p 电子跃迁为 d' 电子而其自旋与该锰离子中的 5 个 d 电子的自旋反平行 (单重) 和平时能量的增加。

超交换作用给出了以下与实验符合的规律。

① 由于通常 $J_d < 0$, 故当磁性离子中 d 电子数 < 5 时, $\Delta E_3 < \Delta E_1$, 于是 $J_i > 0$, 为铁磁耦合。当 d 电子数 ≥ 5 时为反铁磁耦合。

② J_i 的绝对值正比于 ρ^2 。 ρ 的大小依赖于阳离子的低价倾向和阴离子的负电离性。 ρ 又依赖于阳离子是否处于具有哑铃形 p 轨道的阴离子的两侧, 成 180° 。

许多反铁磁、亚铁磁和铁磁化合物相变温度 T_N 或 T_c 的规律可用超交换作用解释。

有一类氧化物, 如 $\text{La}_{1-x}\text{Ca}_x\text{MnO}_3$, 有铁磁性相变, 并在 T_c 附近伴有金属-半导体相变和庞磁电阻 (colossal magnetoresistance)。其铁磁相变可用双交换作用定性解释。这类材料中存在着变价离子 Mn^{3+} 和 Mn^{4+} 。在相邻的 $\text{Mn}^{3+} - \text{O}^{2-} - \text{Mn}^{4+}$ 中, Mn^{3+} 中的一个 d 电子可通过中间的 O^{2-} 转移到另侧的 Mn^{4+} , 使 Mn^{3+} 与 Mn^{4+} 的位置互换而总能量不变。这个过程的存在要求 Mn^{3+} 及 Mn^{4+} 的电子自旋同方向, 即具有正的双交换作用, 导致铁磁耦合, 同时使电导增加。因而当 $T < T_c$ 时, 该材料呈铁磁性及金属性导电, 而且这个导电为自旋极化电子的导电。 $T > T_c$ 时呈顺磁态且电导大大下降。

(3) 金属中的间接交换作用, RKKY 交换作用^[6] RKKY 交换作用是

由 Ruderman, Kittel, Kasuya 和 Yosida 先后根据 Ag 和 Cu-Mn 合金的实验提出的。它是通过金属中的传导电子作为中间媒介,使局域磁矩发生间接交换作用。对局域磁矩与传导电子间的相互作用进行微扰计算证明,传导电子的自旋被局域磁矩极化,且极化密度随与局域磁矩的距离而呈振荡型衰减。被极化的传导电子使相邻的局域磁矩发生间接交换耦合,其耦合亦随距离呈振荡型衰减。其等效交换哈密顿为

$$H = -2 \sum_{ij} J(R_{ij}) \mathbf{S}_i \cdot \mathbf{S}_j \quad (10-60)$$

$$J(R_{ij}) = -\frac{3n}{4\pi E_F} |J_{im}|^2 k_F^2 F(2k_F R_{ij}) \quad (10-61)$$

$$F(2k_F R_{ij}) = \frac{2k_F R_{ij} \cos(2k_F R_{ij}) - \sin(2k_F R_{ij})}{(2k_F R_{ij})^4} \quad (10-62)$$

式中, i, j 对局域磁矩求和, R_{ij} 为两个局域磁矩的距离; J_{im} 为局域矩与传导电子间的交换积分; \mathbf{S}_i 和 \mathbf{S}_j 为两个局域矩的自旋矢量; E_F 及 k_F 为传导电子的费密能及波矢。RKKY 交换常数 $J(R_{ij})$ 为 R_{ij} 的振荡型衰减函数如式 (10-62) 所示。它是一种长程交换作用,且强度和符号随距离而变。在局域电子附近传导电子的自旋极化密度亦有相似的空间分布。

$$\langle S_z \rangle = CF(2k_F R_{im}) \quad (10-63)$$

R_{im} 为电子 m 与局域磁矩 i 的距离。图 10-15 为交换常数 J 和自旋极化对距离 R 的依赖关系。这种关系很适合于稀土金属。4f 电子半径远小于原子半径,局域性很强。4f 间的交换耦合只能是通过传导电子来实现。稀土

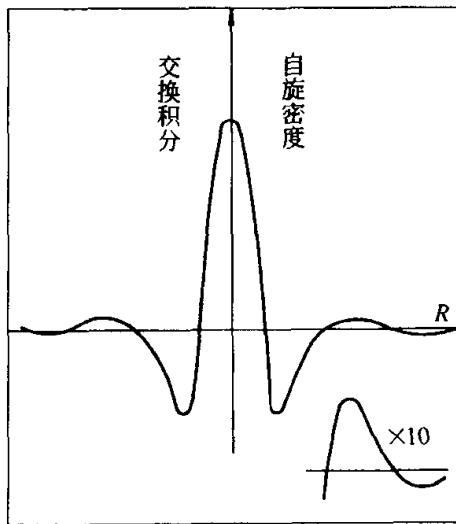


图 10-15 局域矩附近传导电子
自旋极化密度和交换常数
与距离的关系

金属复杂的磁结构反映出其最近邻、次近邻存在着分子场且符号可能不同。理论预期的传导电子振荡型分布则是用核磁共振在 Cu-Fe-Co-Mn 稀释合金和多层膜中直接观察到的^[2]。

(4) 巡游电子交换作用^[2] 铁族合金中 3d 电子的非局域性要求用巡游电子交换作用来解释其分子场的来源。d 电子处于窄能带中,更近似于紧束缚巡游电子。虽然电子属于金属公有,但在巡游于晶格中时有更多的几率陷入到正离子的位阱中。基于原子轨道线性组合(LCAO)方法的理论,计算出的多电子系统的总能量中包含了库仑作用能 C 及交换能 J 。由于电子的巡游性与局域性结

合, 其相互作用能中包含了多种电子组态的相互作用能, 其中尤以当 d 电子巡游至同一个原子时的原子内库仑作用能和原子内交换能最大, 远大于原子间的交换能。因而可解释其高的总交换常数和居里温度, 它们可表为

$$I_{dd'}^c = \int \frac{e^2}{r_{nm}} \phi_d^*(r_n) \phi_{d'}^*(r_m) \phi_d(r_n) \phi_{d'}(r_m) dv_n dv_m \quad (10-64)$$

$$I_{dd'}^x = \int \frac{e^2}{r_{nm}} \phi_d^*(r_n) \phi_{d'}^*(r_m) \phi_{d'}(r_n) \phi_d(r_m) dv_n dv_m \quad (10-65)$$

它们分别是电子 n , m 在同一原子 d 及 d' 轨道时由于电子间排斥能引起的库仑能和交换能。

10.2.4 强磁材料的本征性能^[2]

强磁材料的本征性能包括居里温度、饱和磁化强度和饱和磁化强度的温度依赖关系。

(1) 居里温度 实验上常用磁性测量方法来确定居里温度。将 $M_s \sim T$ 关系在居里点附近外插到 M_s 趋于零以确定 T_c , 如图 10-16。其他物性在居里点处的反常变化也可被用来确定居里点, 如磁导率、电阻率等。

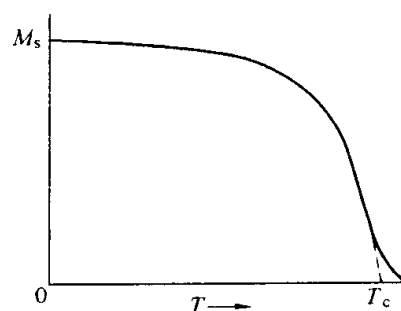


图 10-16 外插法求居里温度

根据局域模型的式(10-57), T_c 正比于分子场常数或交换积分以及 S^2 。但 S 变化的幅度有限, 因而 J 的调控更加重要。不同的交换作用机理有不同的规律。对于 $A-B$ 二元合金, 将式(10-57)推广, 得到以下关系^[9]

$$T_c = \frac{Z}{2k_B} [n_A^2 J_{AA} + n_B^2 J_{BB} + 2n_A n_B J_{AB}] \quad (10-66)$$

式中, n_A 、 n_B 为合金元素的百分比, Z 为配位数, J_{ij} 为合金中不同近邻的交换积分。公式对 Fe-Ni 合金和 Co-Ni 合金的实验结果可很好地拟合。一些二元稀土-过渡族合金的 T_c 与稀土元素的关系示于图 10-17。 R_2Co_{17} 的 T_c 对稀土元素的依赖不大, 说明其居里点主要由 J_{Co-Co} 决定, 但 R_2Fe_{17} 的 T_c 表现出明显的对稀土元素的依赖性, 意味着由于 J_{Fe-Fe} 较小, J_{Fe-R} 和 J_{R-R} 的影响不可忽略, 而且其变化与稀土原子 S 的变化相似。

(2) 饱和磁化强度 测量得到的高场磁化强度是磁场和温度的函数, $M = M(T, H)$ 。为了精确地测定 0K 下的 M_0 和 p_s , 一般需通过两次外插。首先要从高场磁化曲线 $M(T, H) - H$, 用外插方法确定该温度下的自发磁化强度 $M_s(T, 0)$, 见 10.3 节。然后按低温下 $M_s(T) \sim T$ 的关系外插到 $T = 0$, 从而得到 M_0 , 表 10-6 列出一些强磁物质的饱和磁化强度和居里温度。按式(10-40), $M_0 = N p_s \mu_B$ 。

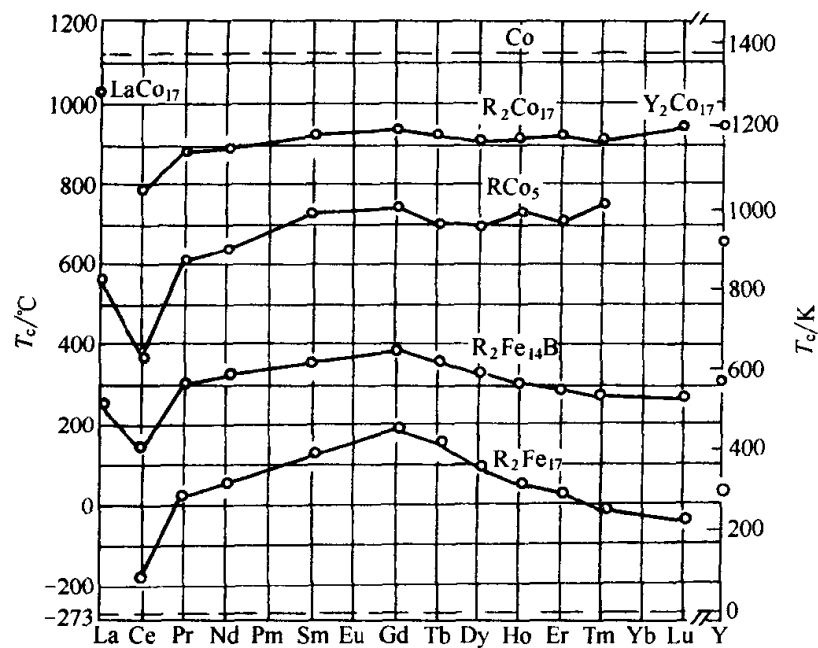


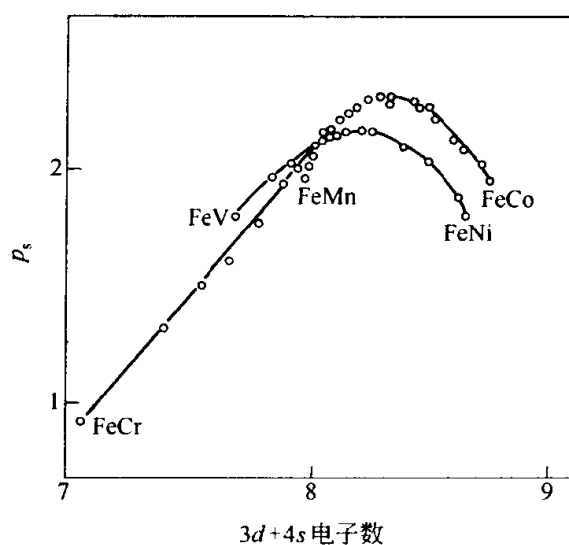
图 10-17 一些稀土过渡族金属合金的居里点

表 10-6 一些强磁物质的饱和磁化强度和居里温度 (M_s 的单位为 G 或 10^3 A/m)

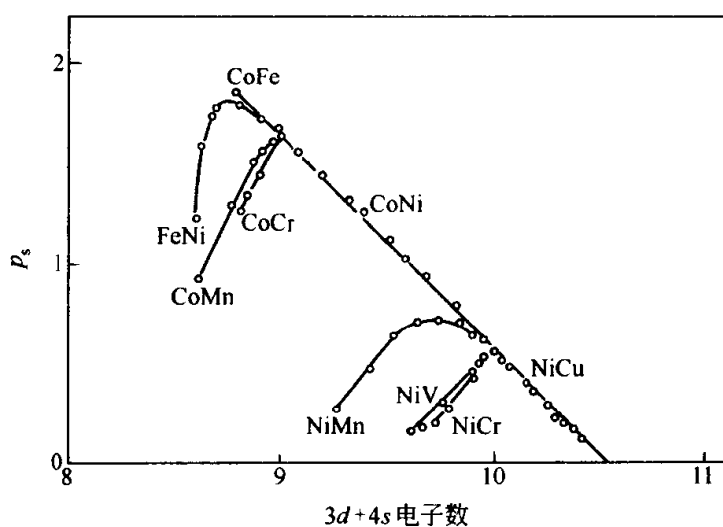
物 质	M_s (室温)	M_s (0K)	p_s (μ_B /原子或分子)	T_c /K
Fe	1707	1743	2.216	1043
Co	1400	1447	1.715	1395
Ni	485	521	0.606	631
Gd	—	1980	7.55	293
Dy	—	2920	10.2	88
MnBi	620	680	3.52	630
Cu_2MnAl	430	(580)	(4.0)	603
CuMnIn	500	(600)	(4.0)	506
MnAs	670	870	3.4	318
MnB	147	—	—	533
Mn_4N	183	—	0.24	745
MnSb	710	—	3.53	587
CrTe	240	—	2.39	336
CrO_2	515	—	2.03	386
EuO	—	1920	6.8	69
MnFe_2O_4	410	—	5.0	573
Fe_3O_4	480	—	4.1	858
$\text{Y}_3\text{Fe}_5\text{O}_2$	130	200	5.0	560

3d 族合金的 p_s 多服从斯莱特-泡林曲线，示于图10-18(a)、(b)。图中两条直线可用刚带模型解释，即假设能带不随合金元素含量而变化，合金成分的改变只是改变了 $3d + 4s$ 电子数，从而得到 p_s 对 $3d + 4s$ 电子数的线性性

依赖关系。其中加入 Mn, V, Cr 元素的非线性偏离则是由于反铁磁耦合或杂质束缚态所致。



(a)



(b)

图 10-18 斯莱特-泡林曲线

(a) bcc; (b) fcc

(3) 饱和磁化强度对温度的依赖关系 强磁材料的分子场理论给出了饱和磁化强度对温度依赖关系的描述, 但尚不能与实验精确地一致。为确定 0K 下饱和磁化强度和居里点均需要进行外插。因此需要进一步确定居里点及 0K 附近 $M_s \sim T$ 的行为, 见图 10-19。

当 $T \leq T_c$ 时, 自发磁化强度 $M_s \ll M_0$, 在局域分子场理论的式 (10-

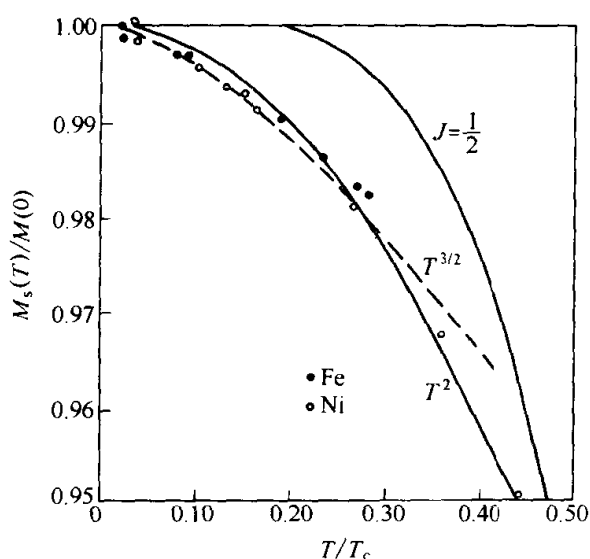


图 10-19 0K 附近自发磁化强度随温度的变化

33) 中的 $\alpha_J \ll 1$ 时, 可有近似表示:

$$M_s(T)/M_0 = C(T_c - T)^\beta \quad (10-67)$$

式中, 指数 $\beta = 0.5$ 。但是 Fe, Co, Ni, Gd 的 β 测量值分别为 0.38, 0.42, 0.38, 0.5, 且系数 C 不一样。

在低温下当 $T \ll T_c$ 时, 局域分子场理论的 $M_s \sim T$ 的关系明显比实验下降的慢。原因是激发单个不平行的原子磁矩, 需消耗较高的能量以克服分子场作用。低温下实验曲线与 $T^{3/2}$, 或稍高

温度下与 T^2 较为相符。理论指出, 激发自旋波需要的能量较小。自旋波是反向自旋在原子间的传播。经计算, 低温段 M_s 随 $T^{3/2}$ 的规律下降。

$$M_s(T) = M_0(1 - CT^{3/2}) \quad (10-68)$$

对立方晶体, $C = \frac{0.0587}{\xi S} \left(\frac{k_B}{2JS} \right)^{3/2}$ 。对简立方、体心和面心立方, ξ 分别为 1, 2, 4。与实验较为一致。对低温下的 $M_s \sim T$ 的拟合可用来确定磁体的交换常数 J 或劲度系数 $D = 2JSa^2$ 。

10.3 强磁材料的磁畴与磁化^[2,6]

强磁材料多种多样的磁化曲线、磁滞回线、磁谱等技术性能基于其多种多样的磁畴结构和磁场使磁畴结构的调整过程。本节将对磁畴的形成、影响磁畴结构的因素及磁化过程作简要讨论。磁畴与磁化通常采用宏观理论处理, 即按自由能极小的原理来处理。

10.3.1 影响磁畴结构的能量

(1) 交换能 强磁体在 T_c 以下的自发磁化来源于交换作用。它是形成磁畴的首要原因。将式 (10-55) 的自旋算符看作经典矢量, 若 $S_i = S_j = S$, 则交换能可写为

$$E_{ex} = -2JS^2 \sum'_{ij} \cos \phi_{ij} \quad (10-69)$$

i 对所有原子求和, j 对最近邻求和, \sum' 表示求和不重复, ϕ_{ij} 为 i, j 原子自旋的夹角, 有时取 $\cos \phi_{ij}$ 的近似式, 只保留含 ϕ_{ij} 的项, 则

$$E_{\text{ex}} = JS^2 \sum_{ij}' \phi_{ij}^2 \quad (10-70)$$

E_{ex} 的极小要求 $\phi_{ij} = 0$, 即相邻原子自旋平行。在磁畴结构中 ϕ_{ij} 常常很小, 因而可将 ϕ_{ij} 看作是坐标的连续函数, 经推导可得到单位体积的交换能为

$$E_{\text{ex}} = A [(\nabla \alpha_1)^2 + (\nabla \alpha_2)^2 + (\nabla \alpha_3)^2] \quad (10-71)$$

式中, $A = \frac{\xi JS^2}{\alpha}$, 对于简立方、体心立方和面心立方, $\xi = 1, 2, 4$ 。 α_1, α_2 和 α_3 为 \mathbf{S} 或 \mathbf{M}_s 对于直角坐标的方向余弦。

有时用等效交换场 \mathbf{H}_{ex} 更为方便。 $E_{\text{ex}} = -\frac{1}{2} \mathbf{H} \cdot \mathbf{M}_s$

$$\mathbf{H}_{\text{ex}} = \frac{2A}{M_s^2} \nabla^2 \mathbf{M}_s \quad (10-72)$$

(2) 磁各向异性性能 磁各向异性性能为磁性材料中的重要性能之一。磁畴中自发磁化在易磁化方向。晶体中原子排列的各向异性性能导致磁晶各向异性性能。自发磁化 \mathbf{M}_s 沿晶体不同方向时能量密度不同, 称为磁晶各向异性性能 E_K , 立方晶体的 E_K 有立方对称性, 其表达式为:

$$E_K = K_1(\alpha_1^2 \alpha_2^2 + \alpha_2^2 \alpha_3^2 + \alpha_3^2 \alpha_1^2) + K_2 \alpha_1^2 \alpha_2^2 \alpha_3^2 \quad (10-73)$$

α_1, α_2 及 α_3 为 \mathbf{M}_s 相对于以立方边为直角坐标的方向余弦。对六角晶系

$$E_K = K_{u1} \sin^2 \theta + K_{u2} \sin^4 \theta + K_{u3} \sin^6 \theta + K_3 \sin^6 \theta \cos^6 \phi \quad (10-74)$$

式中, θ 为 \mathbf{M}_s 对六角轴 c 的夹角, ϕ 为方向角。对正方晶系

$$E_K = K_{u1} \sin^2 \theta + K_{u2} \sin^4 \theta + K_3 \cos^2 \alpha \cos^2 \beta \quad (10-75)$$

θ 为 \mathbf{M}_s 与正方轴 c 的夹角, α 和 β 为与其他两轴的夹角。

磁晶各向异性性能最小的方向称为易磁化方向, E_K 最高的方向为难磁化方向, 难易磁化方向决定于磁晶各向异性性能常数 K_1, K_2 等的符号及绝对值的大小。沿单晶体不同方向的磁化曲线不同, 如图 10-20 所示。没有结晶组织的多晶材料的磁化曲线没有各向异性性能, 但其晶粒中的磁各向异性性能对多晶材料的技术磁性有很大影响, 因此对磁晶各向异性性能的实验及理论研究以控制磁晶各向异性性能及其对温度的依赖关系十分重要。

除磁晶各向异性性能外, 在单晶和多晶体中均可出现的感生各向异性性能也很重要, 经过磁场热处理或应力热处理, 或由于薄膜的生长工艺, 由于原子的特殊排列, 均可产生单易磁化轴的感生磁各向异性性能, 多晶中的

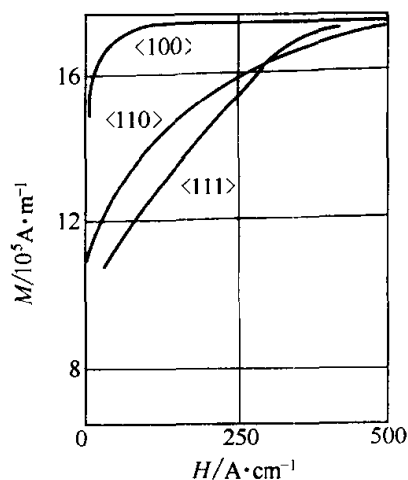


图 10-20 铁单晶的 $\langle 100 \rangle$ 方向为易向, $\langle 111 \rangle$ 为难向

能量表达式为

$$E_{\text{in}} = K_{\text{in}} \sin^2 \theta \quad (10-76)$$

在铁磁与反铁磁材料组成的包裹微粒和双层薄膜中由于界面交换耦合可能出现非偶数性的单向各向异性性能, 表示为

$$E_{\text{unid}} = -K_{\text{ex}} \cos \theta \quad (10-77)$$

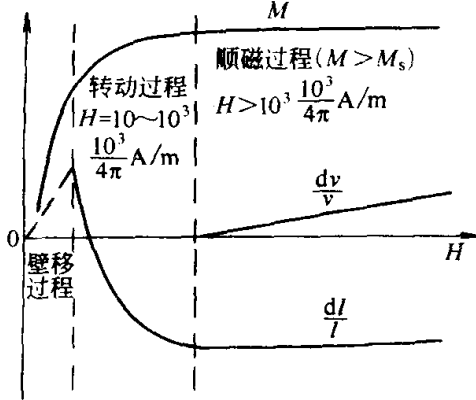


图 10-21 铁的磁化曲线与磁致伸缩的示意图

(3) 磁致伸缩与磁弹性能

① 磁致伸缩。当强磁体被磁场磁化时, 它的形状和体积随之变化, 称为磁致伸缩。其长度的相对变化称为线磁致伸缩, 常用 $\frac{\Delta l}{l}$ 表示; 体积相对变化称为体积磁致伸缩。它们来源于与磁畴中自发磁化相应的自发形变。因此磁致伸缩随磁场增加而渐趋饱和。图 10-21 为铁的磁致伸缩示意图, 在磁场较低的未饱和区以线伸缩为主。

线伸缩既依赖于磁场方向又依赖于测量线伸缩的方向。立方晶体的饱和线伸缩为

$$\left(\frac{\Delta l}{l} \right)_{\alpha\beta} = \frac{3}{2} \lambda_{100} \left(\alpha_1^2 \beta_1^2 + \alpha_2^2 \beta_2^2 + \alpha_3^2 \beta_3^2 - \frac{1}{3} \right) + \lambda_{111} (\alpha_1 \beta_1 \alpha_2 \beta_2 + \alpha_2 \beta_2 \alpha_3 \beta_3 + \alpha_3 \beta_3 \alpha_1 \beta_1) \quad (10-78)$$

α_i 和 β_i 分别为 M_s 和 Δl 的方向余弦。 λ_{100} 和 λ_{111} 分别为饱和磁化在 $[100]$ 和 $[111]$ 方向并沿该方向测量的线伸缩。无织构的多晶的饱和线伸缩为

$$\frac{\Delta l}{l} = \frac{3}{2} \lambda_s \left(\cos^2 \theta - \frac{1}{3} \right) \quad (10-79)$$

$$\lambda_s = (2\lambda_{100} + 3\lambda_{111})/5$$

② 磁弹性能。磁弹性能为磁致伸缩的逆效应——应力及应变对磁化的影响, 导致附加的磁各向异性, 又称应力磁各向异性。立方晶体的磁弹性 E_σ 的表达式为

$$E_\sigma = -\frac{3}{2} \sigma [\lambda_{100} (\alpha_1^2 \gamma_1^2 + \alpha_2^2 \gamma_2^2 + \alpha_3^2 \gamma_3^2) + 2\lambda_{111} (\alpha_1 \alpha_2 \gamma_1 \gamma_2 + \alpha_2 \alpha_3 \gamma_2 \gamma_3 + \alpha_3 \alpha_1 \gamma_3 \gamma_1)] \quad (10-80)$$

式中, σ 为应力的数值, α_i 和 γ_i 为 M_s 和 σ 的方向余弦。 $\sigma > 0$ 为张应力, $\sigma < 0$ 为压应力。若材料的 $\lambda_{100} = \lambda_{111} = \lambda_s$, 则

$$E_{\sigma} = -\frac{3}{2}\lambda_s\sigma\cos^2\theta \quad (10-81)$$

式中, θ 为应力 σ 与磁化矢量 M_s 的夹角。上式为单轴各向异性, 当 $\lambda_s\sigma > 0$ 时, 应力方向为易磁化方向。当 $\lambda_s\sigma < 0$ 时, 易向垂直于 σ 的方向。

(4) 磁场作用能和退磁场能 宏观磁化强度 M 与磁场作用能为

$$E_H = -\mu_0 MH\cos\theta \quad (10-82)$$

式中, θ 为 M 与 H 的夹角, 若在磁畴中, $M = M_s$ 。

当强磁体被磁化后, 它本身的磁化亦产生磁场。这个磁场在强磁体中常与磁化的方向相反, 称为退磁场, 以 H_M 表示。磁矩与退磁场的作用产生退磁场能。

$$E_M = -\frac{1}{2}\mu_0 M \cdot H_M \quad (10-83)$$

退磁场的一般表达式为

$$H_M = -\nabla \left(\int_S \frac{M \cdot n}{R} dS + \int_v \frac{-\nabla \cdot M}{R} dv \right) \quad (10-84)$$

第一项为表面磁荷 $M \cdot n$ 产生的退磁场, 第二项为体磁荷 $-\nabla \cdot M$ 产生的退磁场。

当强磁体为均匀单相且为椭球形状时, 在均匀外磁场下可被均匀磁化, $\nabla \cdot M = 0$, 于是退磁场为由表面磁荷产生的均匀退磁场, 可以证明

$$H_M = -NM \quad (10-85)$$

N 称为退磁因子, 只依赖于磁体的形状及坐标。当坐标为椭球的主轴时

$$H_M = -(iN_x M_x + jN_y M_y + kN_z M_z) \quad (10-86)$$

而 $N_x + N_y + N_z = 1$ (SI 制), 或 $= 4\pi$ (CGS 制)。饱和磁化在椭球不同方向时, 退磁场能不同, 称为形状磁各向异性。但非椭球磁体中不可能有均匀的退磁场和 M 。这时式(10-85)及退磁因子可表示为坐标的函数。

应该指出, 一个有多畴结构的强磁体的表面, 晶界和缺陷处均可有表面磁荷或体磁荷, 从而产生不均匀的退磁场, 亦称散磁场。

10.3.2 强磁体中的磁畴结构

在居里温度以下大块强磁体中自发磁化的分布形成多畴结构。各个磁畴的 M_s 沿各自的易磁化方向。相邻磁畴间为若干原子厚的过渡层, 称为畴壁。在退磁状态下, 磁畴的分布使强磁体的总磁矩或平均磁化强度为零。在磁场作用下磁畴调整导致技术磁化, 最后趋于饱和。

(1) 畴壁 畴壁为一个有限厚度的过渡层。其中的 M_s 逐步从一侧磁畴的 M_s 方向转动到另一侧的方向。图 10-22 为一个单易磁化晶体中的 180° 畴壁结构。其两侧磁畴的 M_s 相差 180° , 在畴壁中从正 z 转到 $-z$ 。 θ 为 M_s

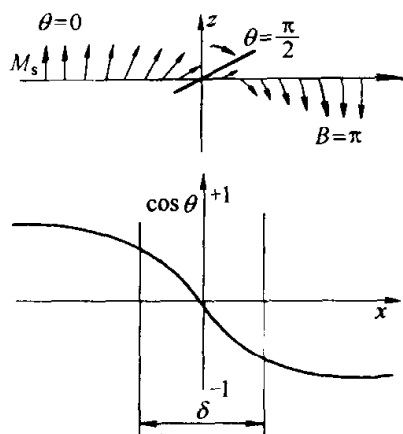


图 10-22 单轴晶体中 180° 畴壁中的磁化分布示意

与 z 的夹角，畴壁中的 M_s 始终在 yz 平面中转动，使 $\nabla \cdot \mathbf{M} = 0$ ，不产生磁荷，称为布洛赫畴壁。变分计算证明，畴壁中 M_s 的转动不是均匀的，而服从 $E_{ex} = E_K$ ，难磁化方向转动最快，在畴壁两侧接近易磁化方向时，转动很慢。根据公认的定义，单轴晶体中 180° 畴壁的厚度为 $\delta = \pi\delta_0 = \pi\sqrt{A/K_{ui}}$ 。畴壁中能量比磁畴高。 180° 畴壁单位面积的能量为 $\gamma = 4\gamma_0 = 4\sqrt{AK_{ui}}$ 。立方晶体中有各种取向的 180° 及 90° 畴壁，其中 M_s 绕畴壁法线转动。其厚度及能量均以 δ_0 及 γ_0 为单位，但系数不同。经计算，铁和钴的 δ_{180} 分别为 140nm

和 15nm。 γ_{180} 分别为 1.25mJ/m^2 和 8.2mJ/m^2 。

当晶体中有不均匀内应力时

$$\delta = \pi \sqrt{A/K_{\text{eff}}} \quad (10-87)$$

$$\gamma = 4 \sqrt{AK_{\text{eff}}} \quad (10-88)$$

$K_{\text{eff}} = bK_{ui} + \frac{3}{2}\lambda_s\sigma$ 。 σ 为坐标的函数， γ 亦随坐标而变。

在厚度仅数十纳米的薄膜中，布洛赫畴壁在薄膜表面上出现的磁荷使畴壁能量升高，这时出现了所谓奈耳畴壁。这种畴壁内 M_s 以薄膜法线为轴进行转动，避免了薄膜表面上的磁荷，但畴壁中出现磁荷。图 10-23 中曲线 B 和 N 为 80Ni-Fe 合金膜中两种畴壁能随膜厚 d 的关系。

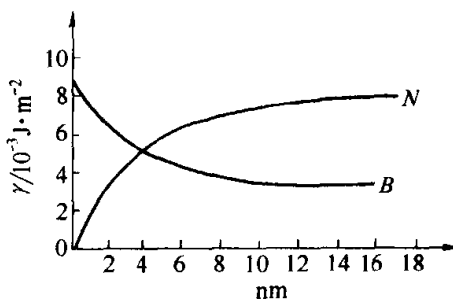


图 10-23 布洛赫壁和奈耳壁能密度对膜厚的依赖关系

畴壁中 M_s 转动可以是左旋也可右旋，当同一个畴壁中一部分为左旋而另一部分为右旋时，则在它们连接处出现一个过渡区，称为布洛赫线，布洛赫线首先在作为磁存储的磁畴中出现，并正在被深入研究以实现新的磁存储技术。

(2) 磁畴结构 大块材料常具有多畴结构，其首要原因是因为多畴有利于降低退磁场能。但多畴带来了畴壁能，稳定的多畴结构常决定于体内畴壁能与表面退磁场或表面畴能量的平衡，相应于总自由能的极小。多畴结构的特点通常是内部为粗大整齐的磁畴，磁化在易磁化方向，表面或晶粒间界处有细小的表面磁畴花样以降低退磁场。实验观察及理论计算证明，主畴的宽

度 $D \sim L^n$, L 为晶体或晶粒沿 M_s 方向的尺寸, $n \sim \frac{1}{2}$ 。

固体缺陷对磁畴结构有重要影响, 内应力既影响磁化方向又影响畴壁位置, 掺杂、空泡等宏观缺陷及其产生的局部退磁场会引起局部的精细畴影响畴壁位置。

已发展了多种观察磁畴结构的实验方法。包括最早的与磁粉检疵的原理相似的粉纹法, 磁光效应法, 透射电镜及扫描电镜法, X 光成像法, 中子衍射法以及新近的磁力显微镜法和 X 光磁二色性显微技术等。

(3) 单畴 当强磁体的尺寸很小时, 多畴结构中畴壁能的提高大于退磁能的降低, 因而使单畴成为能量低的状态。这种单畴状态的单畴颗粒常具有高矫顽力。

用能量比较的方法可以粗略地估计单畴颗粒的临界尺寸。一个球状单轴晶体处于单畴状态时, 退磁场能为 $E_M = \frac{\mu_0}{6} \frac{4\pi}{3} R^3 M_s^2$, 而成为含 180° 的双畴结构时, 能量约为 $E_d = \frac{1}{2} \frac{\mu_0}{6} \frac{4\pi}{3} R^3 M_s^2 + \pi R^2 \gamma_{180^\circ}$, 令两者相等可得单畴临界尺寸

$$R_c = \frac{9}{\mu_0} \frac{\gamma_{180^\circ}}{M_s^2} \quad (10-89)$$

将 Co 的数据代入, 得 $R_c = 29\text{nm}$ 。但这种方法常不一定合理, 如将铁的数据代入, 得到的 R_c 远小于铁的畴壁厚度, 显然不合理。严格的方法应采用微磁学理论, 根据自由能极小的原理来计算^[2]。Fe, Ni, Co, SmCo₅, MnBi 球形颗粒的 R_c 分别为 17nm, 39nm, 12nm, 17nm, 17nm。

10.3.3 磁化曲线与磁滞回线

(1) 磁化过程 强磁体的磁化有 3 种机制, 其一为畴壁位移, 使磁场作用能低的磁畴的体积增大而能量高的磁畴缩小, 磁畴内 M_s 的方向不变。其二为磁畴转动, 畴中 M_s 转向磁场。分别示意于图 10-24(b)和(c)。它们是磁畴结构调整的两个过程, 其三为内禀磁化, M_s 略有提高, 示意于图 10-24(e)。一般地说, 在磁化曲线的任一部分这 3 种过程都可存在。但由于内禀磁化的量很小, χ 约为 10^{-4} , 只有在高场下, 磁畴调整基本结束而趋近饱和时, 才能被测出, 低场下 M 增加较快的阶段以壁移及转动为主。

磁滞与磁滞回线说明磁化有不可逆过程。存在于相当大的低磁场范围中, 直到达到磁化曲线的膝点 (K 点) 以上时才逐渐消失。图 10-25 中最大的称为极限磁滞回线, 其他为在较低的 $\pm H_m$ 间的局部磁滞回线。其端点连线为基本磁化曲线。不可逆磁化是一种不连续的跳跃式的磁化, 称为巴克豪森跳跃, 在磁化曲线和磁滞回线的陡峻部分最为集中。

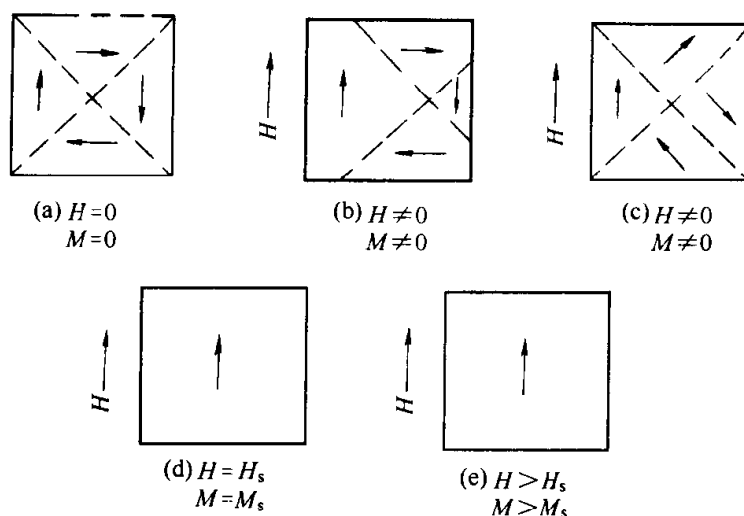


图 10-24 3 种磁化过程的示意图

(a) 未加磁场; (b) 畴壁位移; (c) 磁畴转动; (d) 磁化饱和; (e) 内禀磁化

(2) 畴壁位移 畴壁位移的动力为磁场作用能的降低, 在图 10-26 中, 一块 180° 畴壁 S ($=1$) 在 H 作用下向 x 方向移动。

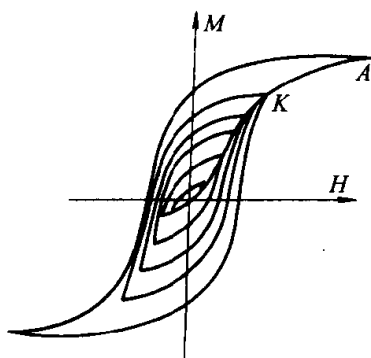


图 10-25 局部及极限磁滞回线

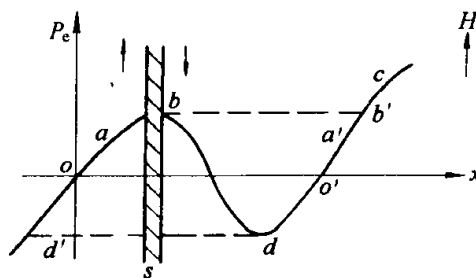


图 10-26 内力起伏及可逆和不可逆壁移的示意图

磁场作用能的下降可等效为外压力 $P_e = -\frac{\delta E_H}{\delta x} = 2\mu_0 HM_s \cos\theta$ 。畴壁位移使有关的内部自由能 E_i 提高, 因而出现了畴壁位移的阻滞力, $P_i = -\frac{\delta E_i(x)}{\delta x}$ 。 P_i 与 P_e 的方向相反。当两者相等时畴壁达到新的平衡位置。故准静态畴壁位移遵守的平衡条件:

$$P_e(H) = -P_i(x) \quad (10-90)$$

式中, $P_i(x)$ 来自 $E_i(x)$, 它们来自固体缺陷与畴壁的相互作用。例如缺陷产生的内应力起伏使畴壁能密度随畴壁位置而变, $\gamma(x)$ 。各种缺陷使局部地区的交换能、磁各向异性及退磁场能起伏, 使 E_i 随畴壁位置而变。当磁

场为零时, $P_e = 0$, 畴壁应处于 $P_i = 0$, 即 E_i 极小的位置。即图 10-26 中的 o 点。当 $H \neq 0$ 时, P_e 使畴壁位移至一个平衡位置, $P_e = -P_i$ 。当畴壁在 o 至 b 间移动时, 为可逆位移。但当磁场 H 增大至一个临界场 H_0 使畴壁移至 b 点时, 畴壁处于一不稳平衡位置。这时只要当外力或内力发生微小涨落导致了瞬时的 $P_e > P_i$, 使畴壁超过 b 点, 畴壁将发生跳跃式位移, 跨过 P_i 的极小 d , 直至达到另一个与 b 点等高的 b' 点, 才能稳定下来。这就是巴克豪森跳跃。到达 b' 点后, H 再增大, 又可继续沿 c 作可逆位移直至到达第二个极大。但若 H 下降, 畴壁不可能沿原路线跳回 b 点, 而是沿 a' 进行。 H 降到 o 时, 畴壁可逆地回到 o' 点, 只有当 H 变为负值使畴壁到达 d 点时才发生反向的不可逆跳跃到达 d' 点, 这一封闭曲线 $oabb'a'o'dd'o$ 正相应于这片畴壁位移导致的局部磁滞回线。出现不可逆跳跃式壁移的临界条件为 $P_i = P_{i\max}$ 相应的临界场为

$$H_0 = P_{i\max} / 2\mu_0 M_s \cos\theta \quad (10-91)$$

因此, 整个畴壁位移过程是连续的可逆位移和不连续的跳跃过程相间进行, 不连续跳跃是不可逆过程, 导致磁滞。 H_0 可有多, 其平均值相应于磁化曲线的陡峻部分及矫顽力 $\bar{H}_0 \sim H_c$ 。有时临界场又称为钉扎场。

根据各种固体缺陷与畴壁相互作用能的模型, 可以进一步得到磁化率和临界场的具体表达式。例如, 按内应力起伏模型, 初始磁化率和临界场分别为

$$\chi_i \approx \frac{\mu_0 M_s^2 \cos^2\theta}{\lambda_s \Delta\sigma} \frac{l}{\delta}, \quad l \gg \delta \quad (10-92)$$

$$H_0 \approx \frac{\lambda_s \Delta\sigma}{\mu_0 M_s \cos\theta} \frac{\delta}{l}, \quad l \gg \delta \quad (10-93)$$

按掺杂引起退磁场对畴壁作用的模型,

$$\chi_i \approx \frac{\mu_0 M_s^2 \cos^2\theta}{K_1} \frac{d^2}{\delta} \beta^{-2/3}, \quad d \gg \delta \quad (10-94)$$

$$H_0 \approx \frac{K_1}{\mu_0 M_s \cos\theta} \frac{\delta}{d} \beta^{2/3}, \quad d \gg \delta \quad (10-95)$$

上式为早期的壁移理论的结果, 以后根据缺陷理论的发展, 对各种缺陷包括点缺陷、位错以及面缺陷等的相关理论得到进一步的发展^[2,6]。但上式引申出的原则性结论还是适用的。那就是畴壁位移的阻滞来源于缺陷, 如应力 $\Delta\sigma$ 和掺杂体积分数 β , 其大小与磁各向异性和磁致伸缩密切相关, 应力起伏波长 l 和掺杂颗粒尺寸 d 代表缺陷作用的尺度。当它们远大于或小于畴壁厚度 δ 时, 使阻滞变小, 当它们与 δ 相当时, 阻滞最大。此外 χ_i 及 H_0 均有各向异性。

强磁体饱和磁化后, 磁场降低并反向, 便得到磁滞回线的下降支或上升支, 称为反磁化。反磁化可能通过畴壁位移或磁畴转动进行。饱和磁化的强磁体应为一个单畴, 若通过畴壁位移反磁化时, 首先应出现反磁化畴。这是一个成核长大的过程, 当磁场达到成核场 H_n 时, 出现了反磁化畴, 然后通过不可逆壁移实现大规模的反磁化。反磁化畴中 M_s 的方向转了 180° , 因此 H_n 应该比转动 180° 需要的磁场更高, 但实际上, 由于固体缺陷出现了局部反向退磁场或局部各向异性很低, 因而 H_n 可以很低, 甚至为负值, 意味着在第一象限就已成核。当材料的 $H_0 > H_n$ 时, 首先成核, 然后在更高的反向场下 $H \sim H_0$ 实现反磁化。当 $H_n > \bar{H}_0$ 时, 一旦出现了反磁化核, 这时外场已超过了 \bar{H}_0 , 故立即就可进行大规模不可逆壁移实现反磁化, 这种材料的回线常呈矩形, 具有高剩磁比 M_R/M_s 。

(3) 磁畴转动 磁畴中 M_s 转动的动力也是磁场作用能的降低, 磁场从零增大时, 磁畴中的 M_s 从易磁化方向转向磁场, 整个过程决定于与 M_s 方向有关的总自由能极小。即对角度的微分, $\delta(E_H + E_K + E_\sigma) = 0$, 于是

$$\mu_0 H M_s \sin \theta (M_s, H) = -\nabla (E_K + E_\sigma) \quad (10-96)$$

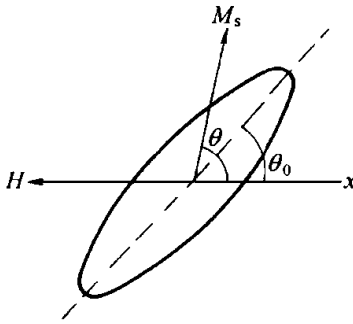


图 10-27 单轴单畴微粒的转动过程

左侧为磁场作用的外力矩 L_e , 右侧为由于磁各向异性与磁弹性的提高引起的内力矩 L_i 。两者方向相反。当满足上式时, 为可逆转动。当转动达到一临界角度时, $E_K + E_\sigma$ 的难磁化方向的位垒被磁场作用能抵消, 将发生 M_s 的不可逆跳跃式转动, 相应的磁场为临界场 H_0 , 图 10-27 为一个具有单易轴磁各向异性的单畴微粒的转动过程。总自由能为

$$E = E_K + E_H = K \sin^2(\theta - \theta_0) - \mu_0 H M_s \cos \theta \quad (10-97)$$

K 可以是形状各向异性, $K = \frac{1}{2}(N_{//} - N_{\perp})M_s^2$; 磁晶或感生各向异性, K_{ui} 或 K_{in} ; 或应力磁各向异性, $K = \frac{3}{2}\lambda_s\sigma$, 经计算当磁场沿不同方向 θ_0 时, 磁滞回线的形状不同。示于图 10-28, 当 $\theta_0 = 0$ 时, 得到最大的临界场和矩形回线, 没有可逆转动, 只有一个不可逆转动。

$$H_0 = \frac{2K}{\mu_0 M_s} \quad (10-98)$$

$\theta_0 = 45^\circ$ 时, H_0 为最小, 回线的曲线部分为可逆转动, 竖直线为不可逆转动。当 $\theta = 90^\circ$, 磁场与易向垂直时, 得一条无磁滞的直线形磁化曲线。

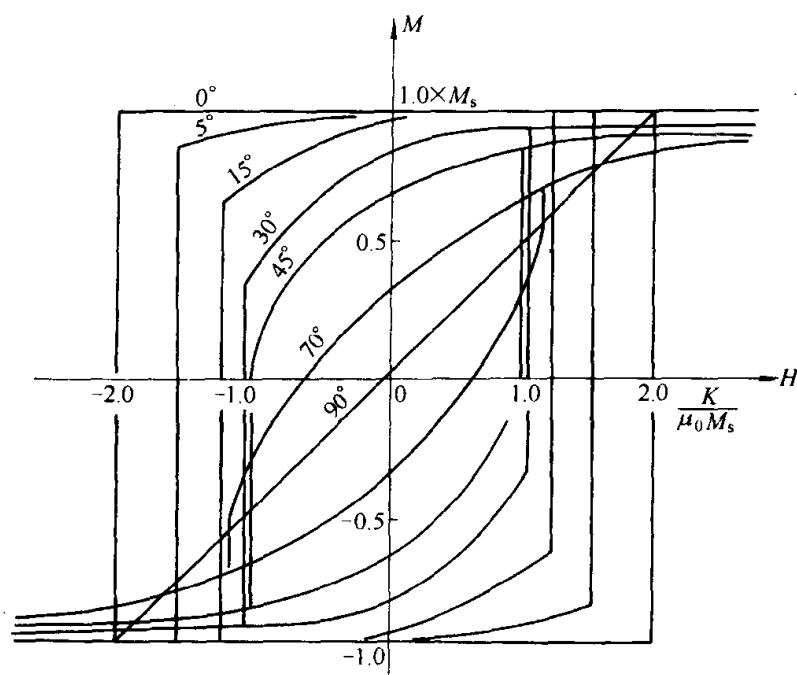


图 10-28 单轴单畴微粒的转动磁滞回线

将式(10-93)、式(10-95)与上式比较, 可以看到, 当材料的 K 和 M_s 等性能给定时, 转动过程的 H_0 远大于壁移。故材料的磁化和反磁化常以畴壁位移为主。一般, 只有在较高磁场下, 在磁化曲线的膝点以上, 转动过程才会成为主要的机制, 单畴微粒则全部为转动磁化过程。故有高 H_0 值。

除以上讨论的均匀转动外, 在单畴微粒中还可有不均匀转动: 在转动过程中微粒各处的 M_s 不保持始终平行, 因而降低了转动过程中的位垒和 H_0 值。

(4) 磁化率, 矫顽力和剩磁 软磁材料要求高磁化率, 低矫顽力。软磁材料的磁化和反磁化常以畴壁位移为主, 获得低 H_c 和高 χ 的方向首先是材料的 $K_1 \sim \lambda_s \sim 0$, 同时要求单相, 杂质少, $\beta \sim 0$, 良好的热处理使 $\Delta\sigma$ 小。当材料有织构时, 沿易向可得到较难向大的 χ_i 及小的 H_c 。对于不可避免的掺杂, 如铁氧体中的空泡, 可通过热处理使其凝聚为 $d \gg \delta$ 以减少阻滞效应。

硬磁材料要求具有高的矫顽力 H_c , 寻求高 H_c 的途径有二, 其一是提高成核场 H_n 和壁移的临界场 H_0 , 或钉扎场; 另一途径是制备高磁各向异性的单畴微粒的集合体, 使其反磁化过程为磁化矢量的转动。这两种途径都要求高的磁各向异性能。

磁滞回线的另一特征为剩磁 M_R 或剩磁比 $\frac{M_R}{M_s}$ 。设多晶的一个晶粒的易向与磁场夹角为 θ , 当磁场从饱和值降到零时, M_s 将从磁场方向可逆转动

到易向，故此晶粒的剩磁 $M_R = M_s \cos \theta$ 。多晶体的剩磁为各晶粒剩磁的平均值，故 $M_R = M_s \overline{\cos \theta}$ ，若无结晶织构，晶轴混乱取向，经计算，对镍型晶体，易向为四个 $\langle 111 \rangle$ 轴， $M_R = 0.866 M_s$ ；铁型晶体，易向为三个 $\langle 100 \rangle$ 轴， $M_R = 0.832$ ；对钴型，易向为单轴， $M_R = 0.5 M_s$ 。若有结晶织构或具有占主导的感生各向异性，当 H 沿易向时， $M_R = M_s$ 回线常为矩形。当 H 与易向垂直时， $M_R \sim 0$ 。若材料的成核场 $H_n < 0$ ，由于在第一象限中形成反磁化畴，将降低 M_R ，故高 M_R/M_s 还要求 $H_n > 0$ 。

(5) 趋近饱和律 在高场下，远在磁化曲线的膝点之上，强磁体趋于饱和。这时强磁体的磁化曲线常服从一个普适的公式，称为趋近饱和律，如下式，

$$M = M_s \left(1 - \frac{\alpha_1}{H} - \frac{\alpha_2}{H^2} - \dots \right) + x_p H \quad (10-99)$$

当磁场很高使 α_1 和 α_2 项均可略去时，磁化曲线为一条斜率为 x_p 的直线，外插到零磁场，可确定 M_s 的数值。 x_p 称为强磁状态下的顺磁磁化率，室温下约为 10^{-4} 。在磁场较低时， α_1 和 α_2 项不可忽略。可以在不同磁场范围中用作图法近似地求 M_s 。亦可用数值计算对高磁化曲线与式 (10-99) 拟合，同时确定 α_1 ， α_2 ， x_p 及 M_s 的数值。研究证明， α_2 项相应于转动过程的最后阶段。 α_1 项为固体缺陷引起的局部磁化不均匀性在高场下逐步缩小的贡献。

10.4 磁化动力学与铁磁共振^[4,6]

10.4.1 磁化进动及转动

当 M 与磁场 H 夹角为 θ 时，磁场力矩为

$$L = \mu_0 M \times H \quad (10-100)$$

M 来自电子自旋及轨道磁矩，它们与动量矩联系： $\mu_0 M = -\gamma P$ ，动量矩的运动方程为 $\dot{P} = L$ ，故

$$\dot{M} = \gamma H \times M \quad (10-101)$$

为磁化矢量 M 的运动方程：它是绕 H 以张角 θ 的右旋进动，如图 10-29。

进动频率为

$$\omega_0 = \gamma H$$

γ 为旋磁比， $\gamma = g \frac{\mu_0 e}{2m}$ (SI) 或 $g \frac{e}{2me}$ ，(CGS)。当 M 仅来自电子自旋时， $g = 2.003$ ， $\gamma = 2.21 \times 10^5 \text{ (A/m)}^{-1} \cdot \text{s}^{-1}$

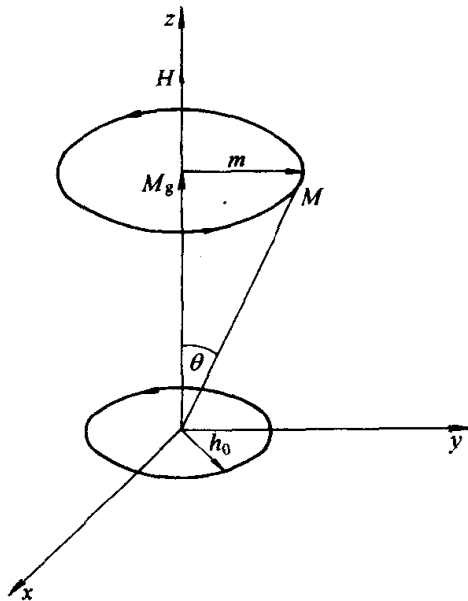


图 10-29 磁化矢量的进动

(SI) 或 $1.76 \times 10^7 \text{Oe}^{-1} \cdot \text{s}^{-1}$, 当电子轨道参与磁化时, $g \neq 2.003$ 。

由于阻尼效应, M 的进动张角 θ 会逐渐减小, 最终趋于零, 如图 10-30 所示。朗道-利夫希茨阻尼进动方程为

$$\dot{M} = \gamma(H \times M) + \frac{\lambda}{M^2} [M \times (H \times M)] \quad (10-102)$$

阻尼项是一个与 M 垂直, 使 θ 变小的矢量。因此, 转动的动态过程是一个阻尼进动过程。阻尼进动的另一表达式为朗道-利夫希茨-吉尔伯特方程, 如下式:

$$\dot{M} = \gamma^* (H \times M) + \frac{\alpha_0}{M} (M \times \dot{M}) \quad (10-103)$$

式中的 $\gamma^* = \gamma(1 + \alpha_0^2)$, $\alpha_0 = \lambda / \gamma M$, λ 和 α_0 为阻尼因子。若 H 为一直流场, 磁化转动为阻尼进动过程, 若阻尼进动在铁磁体各处同频率、同振幅、同相位, 称为一致进动。 M 的三个分量 m_x , m_y 和 M_z 为

$$\left. \begin{aligned} m_x &= m e^{-t/\tau} \cos \omega_0 t \\ m_y &= m e^{-t/\tau} \sin \omega_0 t \\ M_z &= M \{1 - (m/M)^2 e^{-2t/\tau}\}^{1/2} \end{aligned} \right\} \quad (10-104)$$

式中弛豫时间 $\tau = M / H \lambda$, $\omega_0 = \gamma H$ 。当材料的阻尼因子 λ 或 α_0 大时, 弛豫时间短, 转动完成得快。

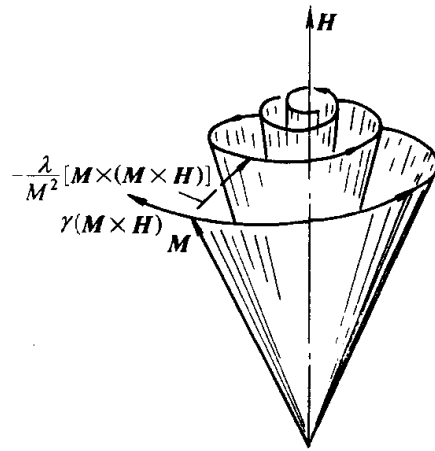


图 10-30 磁化矢量的阻尼进动

10.4.2 张量磁化率和铁磁共振

(1) 张量磁化率 若一个小样品处于沿 z 方向的均匀直流场 H_0 和任意方向的均匀交变场 h 下, $H = H_0 + h e^{j\omega t}$, 且 H_0 远大于 h_0 , 则 M 绕 H_0 作强迫进动。其直流分量 M_z 与交变分量 m 之和为 $M = M_z + m e^{j\omega t}$, m 与 h 的比值为张量磁化率 χ

$$\chi = \begin{vmatrix} \chi & j\chi_a & 0 \\ -j\chi_a & \chi & 0 \\ 0 & 0 & \chi_{//} \end{vmatrix} \quad (10-105)$$

$$m = \chi h \quad (10-106)$$

式中 $\chi = \chi_0 \frac{\omega_{\text{res}}^2 + j\omega\omega_r}{\omega_{\text{res}}^2 - \omega^2 + 2j\omega\omega_r}$, $\chi_a = \chi_0 \frac{\omega\omega_H}{\omega_{\text{res}}^2 - \omega^2 + 2j\omega\omega_r}$

$$\chi_{//} = \chi_0 \frac{\omega}{j\omega + \omega_r}, \quad \omega_{\text{res}} = (\omega_H^2 + \omega_r^2)^{1/2}, \quad \omega_H = \gamma H_0$$

$$\omega_r = \lambda / \chi_0, \quad \chi_0 = M_0 / H_0$$

交变的 m_x 和 m_y , 既依赖于 h_x 又依赖于 h_y 。张量磁化率源于 M 的右旋固有进动。故当外加交变场含有右旋的分量且 $\omega = \omega_{\text{res}}$ 时, 发生铁磁共振。进动振幅大, 大量吸收能量。张量磁化率的性能又称旋磁性。根据此性能发展了多种不可逆器件, 用于微波频段直至光频段。

(2) 铁磁共振, 一致进动 一致进动的铁磁共振的 ω 与共振场 H_{res} 的关系可通过对进动方程和 H_0 作用下的平衡方程联立求解来确定。若略去阻尼项, 得到如下两式:

$$(\omega / \gamma)^2 = \frac{1}{(\mu_0 M \sin \theta)^2} (E_{\theta\theta} E_{\phi\phi} - E_{\theta\phi}^2) \quad (10-107)$$

$$E(\theta, \phi) = \min \quad (10-108)$$

E 为总自由能。例如对于具有垂直各向异性的无织构多晶薄膜,

$$E = -\mu_0 H_0 M_s [\cos \theta \cos \theta_H + \sin \theta \sin \theta_H \cos(\phi - \phi_H)] + \left(\frac{1}{2} \mu_0 M_s^2 - K_u \right) \cos^2 \theta \quad (10-109)$$

从以上关系可得到

$$(\omega / \gamma)^2 = [H_{\text{res}} \cos(\theta - \theta_H) - M_{\text{eff}} \cos 2\theta] [H_{\text{res}} \cos(\theta - \theta_H) - M_{\text{eff}} \cos^2 \theta] \quad (10-110)$$

将上式与不同方向 θ_H 和 ϕ_H 的 H_{res} 的实验值拟合, 可以鉴别铁磁共振, 并确定 γ 值和有效磁化强度 $M_{\text{eff}} = M_s - 2K_u / \mu_0 M_s$, 当 M_s 已知时, 可确定 K_u 。式 (10-109) 及式 (10-110) 的形式随各向异性而变。

(3) 铁磁共振线宽 从式 (10-102) 可得到铁磁共振信号的线宽为

$$\Delta H_{\text{in}} = \frac{2\lambda\omega}{\gamma^2 M_s} \quad (10-111)$$

线宽的定义为铁磁共振吸收峰或复量磁化率虚数的半极大值的两个磁场 H 值之差。 ΔH_{in} 代表来自阻尼的本征线宽。它们的机制有多种, 且随材料而异。多晶材料的铁磁共振线宽有附加的等效线宽 ΔH_{eff} 。

$$\Delta H_{\text{poly}} = \Delta H_{\text{in}} + \Delta H_{\text{eff}} \quad (10-112)$$

ΔH_{eff} 是由于晶粒的取向不一引起磁各向异性场的不一致和种种缺陷引起的磁化和内场不均匀, 使多晶中共振场不一致导致的线宽表观增宽, 而且引起附加的阻尼。因此不能简单地认为相同成分与结构的单晶和多晶材料具有相同的 ΔH_{in} 及阻尼, 钇铁石榴石单晶可有低达 0.1 Oe 的线宽。文献中金属铁磁薄膜的最小线宽约为 30 Oe。

10.4.3 自旋波共振

自旋波可以被热运动能激发, 表现为强磁体的 M_s 随 T 上升按 $T^{3/2}$ 下

降。当强磁体中有合适的磁不均匀时，自旋波可以被同频率的一致进动激发将能量传给晶格，为铁磁共振线宽的一个机制。在强磁薄膜中，满足一定边界条件——表面自旋钉扎时，均匀微波场可以激发某种自旋波驻波，这是一种非一致进动的模式，薄膜厚度各点进动的振幅不等。这种自旋波共振为基特尔首先提出，成为实验上直接检测自旋波和确定交换作用常数的方法之一。例如，对于具有垂直磁各向异性的强磁薄膜，外场沿薄膜法线时的自旋波共振条件为

$$\omega_{\text{res}} = \gamma \left(H_{\text{res}} - M_{\text{eff}} + \frac{2Ak^2}{\mu_0 M_s} \right) \quad (10-113)$$

式中， $k = \frac{n\pi}{L}$ ， L 为膜厚度， n 为沿厚度方向膜中支持的自旋波半波长数，图 10-31 为最早在玻莫合金膜中观察到的共振谱，图中给出了各共振峰相应的 n 。第一峰相应于 $n = k = 0$ ，即为一致进动铁磁共振峰。这时式(10-113)与 $\theta = \theta_H = 0$ 的式(10-110)相同。式中 A 为交换常数。从图 10-31 中自旋波共振场的间距可求出该合金的 $A = 1.22 \times 10^{-15} \text{ J/m}$ 。

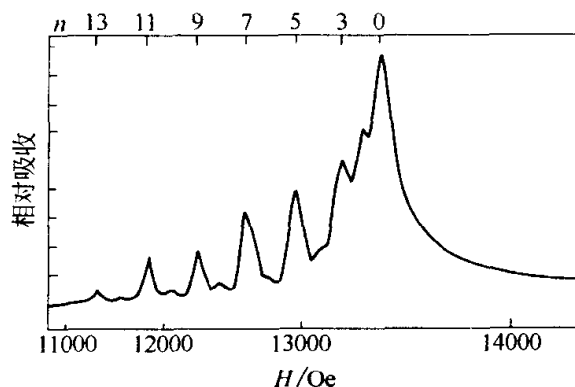


图 10-31 玻莫合金膜中的自旋波共振谱

10.4.4 磁化与反磁化的动态过程

(1) 磁化转动过程与磁化率

的频散 当只有微小的交变场 $h_x = h_0 e^{j\omega t}$ 作用于强磁体，而没有直流场时，得交变的复量磁化 m 及复量磁化率 $\tilde{\chi}$ 。磁畴中的 M_s 绕一个等效的各向异性场 H_K 作强迫进动。令 $H = ih_x e^{j\omega t} + kH_K$ ，代入朗道-利夫希茨方程，可得

$$\tilde{\chi} = \frac{m_x}{h_x} = \chi_0 \frac{\omega_0^2 + j\omega\omega_c}{\omega_0^2 - \omega^2 + 2j\omega\omega_c} \quad (10-114)$$

式中， $\omega_0 = \gamma H_K$ 称为自然共振频率， $\omega_c = \lambda / \chi_0$ 为弛豫频率， $\chi_0 = \frac{M_s}{H_K}$ 为静态转动磁化率， $\tilde{\chi} = \chi' - j\chi''$ 中虚数 χ'' 的出现表征 m_x 的相位落后于 h_x ，交流磁损耗正比于 χ'' 。当 $\omega \ll \omega_0$ 时， $\chi' \sim \omega$ 的关系较平缓，当 $\omega \approx \omega_0$ 时， χ' 及 χ'' 均出现峰值，如图 10-32 所示。

(2) 畴壁位移的动态过程和磁化率的频散 畴壁位移动态过程为畴壁中磁化矢量转动的图形沿位移方向 x 传播，这相当于畴壁中的磁化矢量绕位

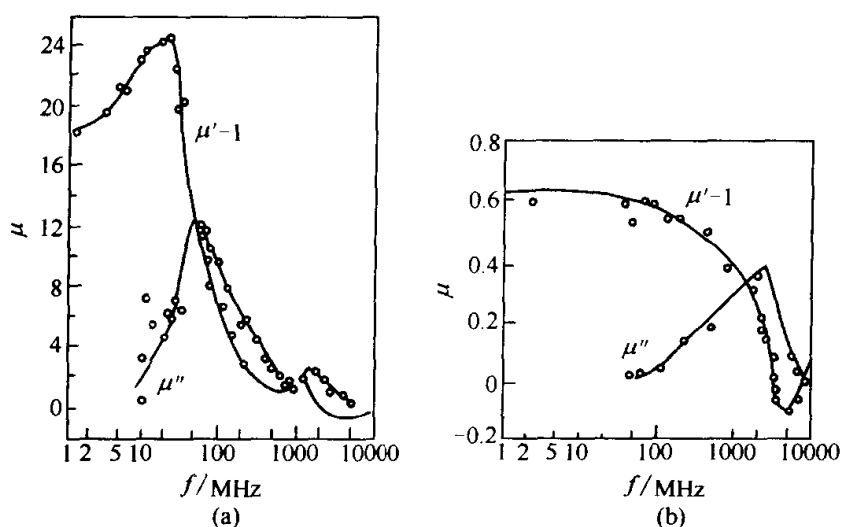


图 10-32 Mg-Fe 铁氧体的初始磁化率的频散曲线

(a) 大块样品; (b) 粉末样品

移方向 x 的进动。这是由于外磁场使畴壁发生位移时，畴壁产生动态磁荷以及沿位移方向 x 的内场。根据朗道-利夫希茨方程可导出一个 180° 畴壁位移的动态方程为

$$m_w \ddot{x} + \beta \dot{x} + \alpha x = 2\mu_0 H M_s \quad (10-115)$$

$m_w = \frac{\mu_0}{\gamma^2 \delta_0}$ 为畴壁的有效质量。这是由于运动的畴壁能量比静止时提

高，相当于一个动能， $m_w v^2/2$ 。阻尼因子 $\beta = \frac{\mu_0 \lambda}{\gamma^2 \delta_0}$ 。 δ_0 为畴壁厚度的单

位， $\delta_0 = \sqrt{A/K}$ 。 α 为畴壁位移的恢复力系数， $\alpha = 4M_s/D\chi_0$ ， D 为畴宽。从上式可得

$$\tilde{\chi} = \chi_0 / \left[1 - \left(\frac{\omega}{\omega_0} \right)^2 + j \frac{\omega}{\omega_c} \right] \quad (10-116)$$

$\omega_0 = (\alpha/m_w)^{1/2}$ 为畴壁共振频率， $\omega_c = \alpha/\beta$ 为弛豫频率， $\chi_0 = \frac{4\mu_0 M_s^2}{\alpha D}$ 为静态初始磁化率。一般，畴壁的共振频率远比转动的自然共振频率低。图 10-32 为 Mg-Fe 铁氧体的初始磁化率的频散曲线。(a) 为大块样品。有两个剧烈频散区，低频相应于畴壁位移，高频相应于磁畴转动共振或自然共振。(b) 为磨碎成粉末后的样品，畴壁位移及低频剧烈频散消失。

(3) 反磁化动态过程 近来，磁存储器件要求快速反磁化，反磁化动态过程颇受重视。反磁化可通过畴壁位移或磁化转动来完成。磁化转动的速度比壁移快。磁化转动的反磁化开关时间依赖于材料的阻尼系数，亦依赖于反磁化场的强度和脉冲长度。

10.5 强磁材料及应用^[5]

10.5.1 强磁材料按组成与结构的分类^[2]

(1) 单质 在元素周期表中, 只有金属呈现强磁性。3 个过渡族铁磁金属铁、钴、镍, 具有高于室温的居里温度, 长期以来获得广泛应用的铁磁材料多数为铁、钴、镍及其合金。此外, 有 6 个稀土金属有强磁转变, 其中钆的 T_c 略高于室温, 其他具有很低的居里点温度。

(2) 合金

① 以铁磁元素为基的合金, 构成了一大类有应用价值的强磁材料。为了改变各种性能, 已形成多种合金。Fe-Co35% 合金有高饱和磁化强度。Ni-Fe 合金系以高磁导率和低矫顽力为特点。Fe-Co-Ni 系中出现了恒磁导率合金。稀土-过渡族合金利用过渡族元素保持高居里点, 并利用稀土元素强轨道——自旋耦合获得高磁晶格各向异性性能, 为高性能永磁及磁光存储介质提供若干种材料。在铁磁合金中加入非铁磁元素可带来性能的改善, 如 Ni-Fe 合金中加入铜、铬、锰等以提高电阻率和磁导率。Fe-Si, Fe-Al 和 Fe-Al-Si 的磁导率比 Fe 有所提高, Fe-Nd 合金中加入 B, N 或 C 使 T_c 及 M_s 提高及改善各向异性性能。

② 非铁磁元素构成的铁磁合金, 有特殊的应用和科学意义。如反铁磁元素锰和铬与正常金属组成的铁磁合金。著名的郝斯勒合金系 Mn-Cu-M, (M = Sn, Al, Ge, Zn, As, In, Sb, Bi 和 Ga 等) 和相似的 MnAg₅Al, Mn-Al-C 等以及二元系 MnBi, MnSb 和 MnAs 等。含 Cr 的如 CrTe, CrAs。此外含过渡元素的, 如 MnSbPt, Cr-Pt 等。又如顺磁过渡族金属和正常金属构成的弱铁磁合金 ZrZn₂, Se₃In 和 Au₄V 等。

(3) 非金属化合物

① 铁氧体^[4,11]。铁氧体为含铁和其他过渡元素的氧化物。磁铁矿 Fe₃O₄ 为古代最早发现的尖晶石结构的铁氧体。20 世纪 40 年代发展了一系列尖晶石型软磁铁氧体, 使非金属强磁材料开始进入工业生产及应用。20 世纪 50 年代出现了六角晶系磁铅石型铁氧体永磁及超高频铁氧体, 并发展了石榴石型铁氧体。它们都是亚铁磁性, 故饱和磁化强度比金属铁磁材料低, 但电阻率高, 在 $10^3 \sim 10^{10} \Omega \cdot \text{cm}$ 。故适用于频率较高的交变场, 直至微波及光频。大大弥补了长期以来金属磁性材料在交变场特别是高频下的不足, 改变了强磁材料在技术应用上的面貌。铁氧体的另一晶系为钙钛矿型, 又称正铁氧体, RFeO₃。在正铁氧体单晶中首先发现泡畴, 磁泡存储器以石榴石单晶薄膜为基本材料并获得应用。

② 其他强磁化合物^[5,9]。还有多种具有强磁转变的化合物, 如 EuO, CrO₂。在钙钛矿型的 RMnO₃ 的系列化合物中, 在低温下有铁磁相变, 且随温度降低在 T_c 附近发生绝缘向金属性导电的转变, 并出现了庞磁电阻效应。此

外, 一些过渡族金属与硫系、磷系及高价元素形成许多强磁性化合物。

(4) 非晶强磁合金^[5] 非晶合金又称金属玻璃。它不具有晶格的长程序, 有似于玻璃的结构。非晶合金系的本征性能 M_s 及 T_c 随成分变化的规律与相近的晶态合金系相似, 但稍低, 这是由于非晶合金中常含有约 20% 的玻璃态形成元素如硼, 磷、硅等之故。非晶合金由于没有长程晶格序, 因而常具有较高的电阻 ($100 \sim 200 \mu\Omega$), 有利于降低涡流损耗; 没有宏观磁晶各向异性只有应力不均匀性导致的不高的磁弹性各向异性, 约 $10 \sim 100 \text{ J/m}^3$; 没有晶界等不均匀性或晶体缺陷, 因而 3d 族元素为基的非晶合金带成为优异的软磁。与非晶相似的纳米晶粒合金, 由于晶粒间交换耦消除了各晶粒中的磁各向异性, 亦获得优异的软磁性能。但含有稀土的非晶合金膜可有高的垂直各向异性及低的噪声, 成为优良的磁光存储介质。

10.5.2 强磁材料的应用^[5]

(1) 软磁 软磁是指高磁导率及低矫顽力的材料, 因而容易磁化亦易于退磁, 交变场下磁滞回线面积小而磁损耗低, 是电工和电子技术的基础材料。用于电机、变压器、继电器、电感、互感以及电磁铁的磁芯等。良好的软磁性能要求尽可能低的磁各向异性和磁致伸缩, 单相和低的内应力, 高的电阻率以降低交变场下的涡流。

① 纯铁具有高 M_s , 相当高的 μ 及低的 H_c , 但电阻率低。常用于仪表及电磁铁的磁轭及要求不高的大功率, 低频软磁。其 μ 及 H_c 值依赖于所含杂质, 真空或氢气热处理可以改善其软磁性能, 但有磁性老化现象。

② 硅钢及铁硅铝合金。Si 的加入可增加纯铁的电阻率, 降低铁心损耗, 降低磁晶各向异性, 表 10-7 为纯铁与 3% Si-Fe 的主要物理性能。含硅量从 0.5% ~ 4.25% 的硅钢片及纯铁为主要的功率软磁, 用于电工技术。通过冷轧再结晶工艺以获得 (110)、[001] 织构的单取向和 (001)、[100] 织构的双取向硅钢片 (3% Si), 沿易磁化方向有优良的软磁性能。加铝也可改善纯铁的软磁性能, 但脆性增加。

表 10-7 纯铁与 3% Si-Fe 的主要物理性能

性 能	bccFe	3% Si-Fe	性 能	bccFe	3% Si-Fe
$\mu_0 M_s / \text{T}$	2.2	2.0	$\lambda_{111} / \times 10^{-6}$	-20	-4
$\rho / \mu\Omega \cdot \text{cm}$	10.1	48	μ_i	150	1500
$T_c / ^\circ\text{C}$	771	745	μ_m	5000	40000 取向
$K_1 / 10^4 \text{ J/m}^3$	5	3.6	$H_c, \text{ A/m (Oe)}$	80(1.0)	8(0.1) 取向
$\lambda_{100} / \times 10^{-6}$	25	23			

在 Fe-Si-Al 三元合金系中, 10% Si-5% Al-Fe 的 K_1 和 λ_s 均近于零, 因而获得高达 20000 的初始磁导率 μ_i , 但脆性高, 只能制成粉末及薄膜使用。

由于 M_s 较低, B_s 约为 1.2 T , 不宜作为大功率软磁。高硬度及高 μ_i 使其可用作磁记录磁头材料。

③ 镍-铁合金系。常称为坡莫合金。富镍的面心立方镍-铁合金为著名的高磁导率软磁合金系。合金的 K_1 既依赖于成分又依赖于无序转度及热处理。图 10-33 为初始磁导率随合金含 Ni 百分比的关系, μ_i 最大值相应于各向异性及磁致伸缩都很小的成分及热处理。

此合金系可分为 3 个主要成分区, (a) 约含 78% Ni 的合金。在这个成分附近, 经快冷, K_1 及 λ_s 均很小。5Mo-79Ni-16Fe 经最佳热处理, μ_i 可达 10^5 。(b) 65% Ni 的合金表现出强烈的磁场热处理效应。(c) 50% Ni 为此合金系 B_s 最高的成分, 其 B_s 约为 1.6 T , 比 78% Ni 合金几乎高 1 倍, 且也有强烈的磁场热处理感生各向异性效应。沿热处理磁场方向磁滞回线为矩形, 有高的最大磁导率 μ_m 。垂直方向磁滞回线倾斜, 磁导率几乎不随磁场变化。Ni-Fe 合金有良好的延展性, 故可轧制成直至 $2.5\mu\text{m}$ 的薄带, 以适用于较高频率的交变场。

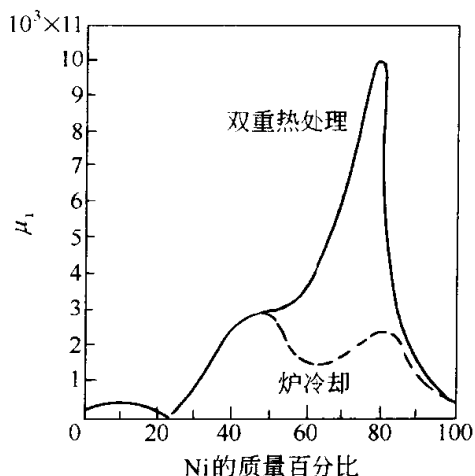


图 10-33 Ni-Fe 合金的 μ_i
对含 Ni 量的关系

④ 铁钴合金。铁-钴合金的 B_s 比纯铁可增加约 10%, 50CoFe 合金的 $B_s = 2.45\text{ T}$, 其 K_1 近于 0 但 λ_s 较高, 故 H_c 约为 160 A/m (2 Oe), μ_m 约为 5000, 加 2% 的 V 可使软磁性改善。

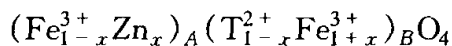
⑤ 非晶软磁合金。非晶合金由熔体快冷获得。用专门的设备可制成厚度仅为几十微米的非晶薄带。在 100 kHz 频率下其性能优于铁氧体及坡莫合金薄带。与晶态合金相应, 非晶合金有 Co-Fe 基, Fe 基, Ni-Fe 基及 Co 基非晶, 它们的性能各有优缺点。代表性能见表 10-8。

表 10-8 一些非晶合金的磁性

磁性能指标	$\text{Fe}_{67}\text{Co}_{18}\text{B}_{14}\text{Si}_1$	$\text{Fe}_{81}\text{B}_{13.5}\text{Si}_{3.5}\text{C}_2$	$\text{Fe}_{40}\text{Ni}_{38}\text{Mo}_4\text{B}_{18}$	$\text{Co}_{72}(\text{FeMoBSi})$
B_s/T	1.80	1.60	0.88	1.00
$H_c/\text{A}\cdot\text{m}^{-1}(\text{Oe})$	4(0.05)	3.2(0.04)	1.2(0.015)	1.4(0.018)
$\mu_m(\times 10^3)$	200	260	500~750	12
$\rho/\mu\Omega\cdot\text{cm}$	130	130	160	

⑥ 软磁铁氧体。由于涡流效应, 用于交变场下的金属性磁性材料必须制成相互绝缘的片材, 厚度从约 0.1 mm 直至微米量级, 或制成粉末。软磁

铁氧体为非金属, 电阻率达 $10^3 \sim 10^{10} \Omega \cdot \text{cm}$, 因此当交变场频率达 GHz 时, 仍可使用块状铁氧体作为磁芯。软磁铁氧体多数为与 Fe_3O_4 相似的尖晶石结构, 其代表性结构分子式为



A 为 O^{2-} 构成的四面体, B 为八面体, 在 T_c 以下 A 晶坐的离子磁矩与 B 坐的反平行, 为亚铁磁性。T 为 Mn 或 Ni 等, MnZn 及 NiZn 铁氧体分别用于音频及射频段, 少量掺杂如 Fe^{2+} , Co^{2+} 的加入可调整各向异性, 磁致伸缩磁导率及其温度系数及频率依赖性。此外具有易平面的六角晶系铁氧体的自然共振频率高, 适用于超高频。

(2) 永磁材料^[5,6] 永磁材料具有高矫顽力, 因而不易退磁。在磁路不闭合时仍可保持较高的剩磁, 提供应用所需的磁场或磁矩。永磁材料要求高 H_c 及 $B_R = \mu_0 M_R$, 软磁的 H_c 常小于 1000 A/m (约 10 Oe), 而硬磁的 H_c 常大于 10^4 A/m , 表 10-9 给出一些代表性永磁材料的磁性。

表 10-9 一些代表性永磁材料的磁性

名 称	B_R /T(10^4G)	H_c		$(BH)_m$		附 注
		/kA·m ⁻¹	/Oe	/kJ·m ⁻³	/MGOe	
36% 钴铜	1.04	13	163	7.8	1.0	可加工
铝镍钴合金	1.28	51	641	44	5.5	铸造
单畴铁钴微粉 ESD	0.68	76	955	24	3.0	电解 Fe-Co 粉 + Pb
铁铬钴合金	1.25	52	653	40	5.0	可加工
六角铁氧体	0.38	235/240	2953/3016	28	3.5	烧结
锰铝碳合金	0.56	180	2262	44	5.5	可加工
铜镍铁合金	0.55	42	528	11	1.4	可加工
钐钴合金	0.9	675/1200	8482/15079	160	20	烧结
钕铁硼合金	1.23	880/960	11000/12000	290	36.4	烧结

表中 H_c 有两个数值时, 第一个相应于 B - H 回线的 BH_c , 第二个相应于 M - H 回线的内禀矫顽力 MH_c , 一般 $MH_c \geq BH_c$, 只有当 MH_c 高时, 两者才有明显的差别。 $(BH)_m$ 称为磁能积, 它表征永磁材料提供磁能的大小。永磁回路中空气隙体积 V_g 中的磁能 $\mu_0 H_g^2 V_g = \mu_0 HB V_m$, V_m 为永磁体体积。 H 和 B 相应于永磁体的工作点。理想的永磁材料具有矩形 M - H 回线及高的 M_R 和 MH_c , 因此寻求优质永磁材料的方向是高 M_s , 高的单轴各向异性常数及合适的制备工艺, 使材料具有使易向平行的结晶织构和高的矫顽力。

① 单畴型永磁。铝镍钴系合金最早出现于 20 世纪 30 年代, 大量研究发展出十余个品种, 成为 20 世纪 70 年代前性能最好的永磁材料, 并证明它们属于单畴型永磁。铁基并含有 Al, Ni, Co 及其他少量元素的多元合金从

约 1250℃ 单相区冷却至 750~850℃ 时, 发生失稳 (spinodal) 分解, 成为富 Fe, Co 的强磁 α' 相与富 Ni, Al 的弱磁相。 α' 相沿 $\langle 100 \rangle$ 轴析出成为直径约 10nm, 长 100nm 的针状体, 近于强磁单轴微粒。在磁场下脱溶分解可使析出物沿最接近磁场的 $\langle 100 \rangle$ 方向择优取向。若材料本身为 $\langle 100 \rangle$ 柱晶或单晶体, 则磁场热处理后的析出物的长轴均沿磁场方向, 成为沿易向具有矩形回线的各向异性永磁。单轴的各向异性主要来源于析出物的形状各向异性, 但 H_c 仅约 45~170kA/m (约 500~2000Oe), 说明其反磁化为不均匀转动且析出物并非完全的单畴。Fe-Cr-Co 为 20 世纪 70 年代初按铝镍钴合金的原理发展出的另一合金。性能相似但有延展性, 具有可加工的特点, 50~60 年代对铝镍钴单畴机制的研究引发了用电解法制备细长状单畴超细微粒永磁的研究。简称 ESD, 并获得成功, 但由于价格高未成商品。

② 铁氧体永磁。六角对称的磁铅矿铁氧体 $\text{BaO}_6\text{Fe}_2\text{O}_3$ (BaM) 或 PbM 及 SrM, 具有单易轴, K_u 约为 $3 \times 10^5 \text{ J/m}^3$ 。用陶瓷工艺可制成 H_c 比铝镍钴高的永磁, 在磁场中用压机将预烧粉末成型, 烧结后得到各向异性永磁, 沿成型磁场方向六角轴择优取向, 具有高 H_c 的矩形 $M-H$ 回线。铁氧体为亚铁磁性, M_s 不高, 故 $(BH)_m$ 不够高。铁氧体永磁价低而性能中等, 取代了许多铝镍钴的市场份额。

③ 稀土永磁。20 世纪 60 年代末以来发展出几种高 H_c 、高 $(BH)_m$ 的优质稀土过渡族合金永磁, 是当前磁性最佳的永磁品种。稀土金属 4f 电子具有大的轨道矩及轨道自旋耦合, 因而磁晶各向异性强, 且多为单轴, 有利于高矫顽力。但居里点低, 经与过渡金属形成合金后 T_c 得到提高。 RCO_5 最早问世, R 为 Sm 或 Pr 或混合稀土 (Ce, La, Nd, Pr)。单相的 RCO_5 的烧结体的高磁晶各向异性导致高成核场及钉扎场及高矫顽力。磁场成型可得到各向异性永磁。 $\text{Sm}(\text{Co}, \text{TM})_{7.5}$ 是另一类脱溶硬化的复相永磁。TM 为 Cu 及其他过渡金属。其性能优于 RCO_5 。当前性能更好且价格低于前两种的是钕铁硼永磁合金。其特殊之点在于非磁元素硼的加入提高了铁钕合金的居里点, 改善其永磁性能, 新的稀土永磁合金仍被广泛研究。

永磁体可提供稳定的磁场而不需电流, 因而有多种应用, 如马达、电机, 扬声器及多种机电产品及生活用品。永磁市场主要为稀土永磁及铁氧体永磁占领, 其他永磁材料各有特点也占据少量市场份额。将永磁铁氧体或稀土合金的粉末与树脂或橡胶等介质混合制成的粘结永磁具有加工精度好, 尺寸精度高, 甚至柔软可变形等优点, 但剩磁及磁能积降低。此外还有一种 H_c 较低的半永磁, 亦有其特殊的用途。应该说明, 大量应用的磁记录介质实际上也是一种永磁。

(3) 磁记录^[5,12] 磁记录是指将信息转化为记录介质的磁化, 并可将

记录的磁化再转为信息的技术。根据需要有模拟式及数字式,广泛用于录音、录像及计算机和多媒体的录码和各种磁卡。磁记录也应包括磁泡及磁随机存储器。其中发展最为迅速的是硬磁盘,20世纪90年代以来信息存储的面密度每年以60%的增幅发展,2000年的最高水平的报道达 100Gb/in^2 ^①。每位的尺寸达100nm以下,已超过了预期的超顺磁极限。

① 磁记录介质。记录介质是一种永磁薄膜,其矫顽力的典型数值为500~4000Oe。记录密度愈高,愈要求高的 H_c 。在数字记录中,每位信息相当于介质中被磁化的微小区域,剩磁在相反的两个方向以产生0和1。因此要求有高的 M_s 及剩磁,产生尽可能高的表面磁场以被磁头读出。实际上 M_s 约为500~1000G ($\mu_0 M_s$ 约为0.63~1.26T),已实用化的记录介质主要有两大类。

a. 颗粒型不连续介质。由磁性单畴微粒与复合高分子介质混合而成,实用化的磁性颗粒主要有以下几种。亚铁磁的针状 $\gamma\text{-Fe}_2\text{O}_3$,包Co或掺Co的 $\gamma\text{-Fe}_2\text{O}_3$,其性能优于前者,铁磁性 CrO_2 ,具有高 M_s 的针状金属铁颗粒,片状钡铁氧体颗粒。

b. 薄膜型连续介质。高密度磁记录介质多采用薄膜型连续介质。采用电解、真空蒸发或溅射等工艺制备。所用的磁性材料包括早期的Co, Co-P, Co-Ni, Co-Ni-Cr, Co-Ni-Pt及近期的Co-Cr合金,包括非晶合金,如CoCrPtTaB,后者已用于高密度硬盘。

② 写入头。常用的感应式写入头是一个带有缝隙的微小环形电磁铁。将电流信号转换为缝隙中的磁场信息,使磁介质磁化而实现写入。为了能在高 H_c 的介质中写入信息且尽量减少噪声,对写入头材料有以下要求:高饱和磁化强度,高磁导率及高使用频率,低矫顽力,高居里温度,小的磁致伸缩系数等。使用的材料有铁氧体多晶或单晶,玻莫合金,Fe-Co-Ni, Fe-Al合金等,近几年铁基合金中加N以提高 M_s 的磁头材料正在研究与开发。

③ 读出头。通常,无电源的感应式写入头可用作读出头。运动的磁介质中的存储信息在读出头中产生感生电动势。但高密度、小型化磁盘的发展使位尺寸不断减小,磁盘线速度降低。感应式读出头的灵敏度不足以检测到储存的信息曾经成为高密度硬盘存储的主要困难。近10年各向异性磁电阻和巨磁电阻读出头的出现成为硬盘密度高速发展的关键技术^[8]。当前的硬磁盘使用的是感应式写入头与巨磁电阻读出头的复式磁头。更加灵敏的磁性隧道结磁电阻读出头正在开发。磁电阻读出头还在向软盘发展以提高软盘的存储密度。

① $1\text{in}^2 = 6.4516 \times 10^{-4}\text{m}^2$ 。

④ 垂直磁记录。目前,实用化的磁记录均通过在介质平面内磁化来实现。当存储密度增大时,平面内存储的位长度减小。因而其退磁因子增大使稳定性降低。这时虽可通过减小介质厚度来改善,但厚度减小会使剩磁磁场减小,且有一定限度。垂直磁记录采用在沿介质法线方向磁化,尺寸愈短,退磁因子愈小,愈稳定,有利于高密度的发展。但由于工艺上的原因,垂直磁记录迄今未能商品化,但恰用于磁光存储。

⑤ 磁光存储^[7]。磁光存储为一种垂直磁记录。通过激光辐照提高介质局部区域的温度,降低其 H_c ,用垂直磁场写入,用磁光 Kerr 效应读出。记录介质为稀土-过渡族亚铁磁非晶合金膜。成分接近抵偿点,因而 M_s 很低,有利于垂直易向的出现。由于稀土金属而有强的垂直各向异性及高 H_c ,且居里温度低, $T_c < 200^\circ\text{C}$ 。用聚焦的功率约 10mW 的激光可使光斑区的温度升高至 T_c 附近,进行磁化或反磁化。磁光 Kerr 效应是指线偏振光在强磁体表面反射后偏振面旋转的效应。Kerr 转角的方向随 M 的方向而有左右之分,因此可根据 Kerr 转角的方向来鉴别介质位中的磁化是向上还是向下,以读出存储的数字信息。极 Kerr 效应的转角的大小,依赖于亚铁磁稀土-过渡族合金,如 Tb-Fe-Co,各次晶格沿法线的磁化强度的方向和大小。由于成分接近补偿点,总磁化强度低,但 Kerr 角并未抵偿。实际上在室温下 Kerr 角主要由过渡族元素 Fe, Co 的磁化提供。Kerr 角的大小决定读出信号的信噪比。磁光存储取垂直磁化的原因之一是由于用极 Kerr 效应读出可得到较大的转角,磁光盘属高密度。20 世纪 80 年代曾认为它是硬盘的竞争对手。10 多年的发展表明,磁光盘在存储密度方面不及硬盘磁记录,主要受限于光的干涉效应对读出的分辨率的影响。目前 640MByte 的 3.5 英寸^① 磁光盘使用波长为 685nm 的激光,其存储面密度为 1.2Gb/in^2 ,远小于硬盘的最高水平。新近推出的以交换耦合的三层膜为介质的 3.5 英寸超分辨磁光盘将存储密度提高到 1.3GByte 及 2.4GB。磁光盘的优点是介质-头间距约 1~2mm,光盘可取出,作为备份便携式存储器,远比软盘容量高,但价格较高。

⑥ 磁随机存储器 MRAM^[8]。20 世纪 50~60 年代,计算机内存为由直径小于 1mm 具有矩形磁滞回线的铁氧体环为存储单元的阵列组成。以后为 Si 随机存储器取代替。随着集成电路的发展,主内存 DRAM 的密度还在不断提高。20 世纪 90 年代以来,以具有巨磁电阻的自旋阀或磁性隧道结为存储单元的研究有新的发展。其存储密度、存取时间和价格均可与半导体存储器竞争。但兼有独特的优点:断电不丢失信息,抗辐射及抗恶劣环境。可能在几年内上市。其数字信息 0 和 1 由一个亚微米尺寸的铁磁金属薄膜的两个

① 1 英寸 (in) = 25.4mm。

不同方向的剩磁态表征, 当剩磁在不同方向时, 该薄膜所属的器件 (自旋阀或磁性隧道结) 的电阻不同 (GMR), 因而可以方便地读出。磁性隧道结 MRAM 存储单元示意图 10-34。

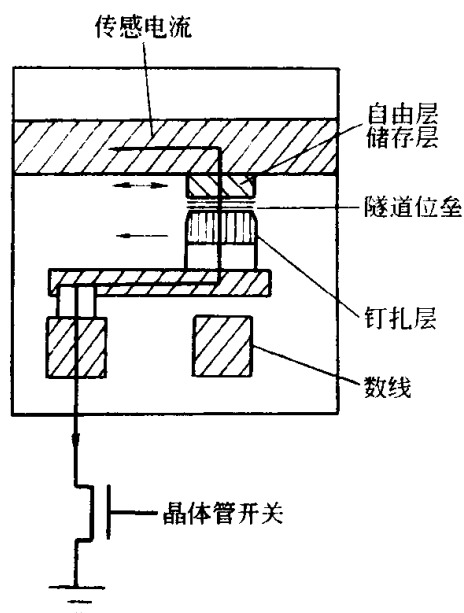


图 10-34 磁性隧道结 MRAM 的存储单元示意图 [S. Tehrani, IEEE MAG. 2000, 36: 2752]

⑦ 磁泡存储器^[2,6]。在具有垂直磁各向异性的石榴石外延薄膜中, 磁畴常为 180° 的条状畴。在外磁场作用下发生畴壁位移, 最后成为圆柱畴, 用 Kerr 效应从上面观察, 貌似水泡, 称为磁泡, 通过磁场可以使磁泡产生, 消灭和移动, 因而可以存储信息并可进行逻辑运算, 磁泡存储器价格高, 但不需要机械部件。目前有一个小的市场。

(4) 旋磁性的应用^[4,11] 高电阻的铁氧体的出现使磁性材料进入到微波领域, 出现了一系列特殊功能的器件, 一定程度上改变了微波技术的面貌。多数铁氧体微波器件均基于铁氧体的旋磁性或张量磁化

率, 构成不同类型的不可逆器件。微波沿传输线正向及反向传输呈现出不同的性能或阻抗。如回旋器: 使平面偏振波的偏振面旋转; 隔离器: 使正向传输衰减很小, 反向波衰减很大; 环行器: 使正、反向的波转向不同通道; 相移器: 成为相控阵天线的关键器体以及滤波器等。此外还出现了非线性器件, 如倍频器、混频器等。

旋磁材料主要为铁氧体。其主要要求如下:

- ① 适当高的磁化强度, 因旋磁性正比于 M_s ;
- ② 低损耗, 因而要求小的有效线宽和高电阻率及小的微波下虚数电介常数;
- ③ 适当高的自旋波线宽 ΔH_k , 它与对高微波功率的承受能力相关。

对旋磁材料的具体要求随微波波长及器件不同而异。可通过改变铁氧体成分及制备工艺来达到, 旋磁材料包括尖晶石铁氧体、石榴石铁氧体和六角铁氧体 3 大晶系。如镁系铁氧体, 镍系铁氧体, 锂系铁氧体, 钇及钇钒石榴石, 钙钒石榴石, M 和 W 型六角铁氧体。后者主要用于毫米波段。此外, 旋磁器件还用于光波, 例如光隔离器, 成为激光通讯等技术中的重要器件。其材料取对红外光透明的石榴石单晶。具有延迟功能的自旋波器件则采用外延石榴石单晶薄膜。

(5) 磁弹性效应、磁卡效应、热膨胀反常及磁性液体^[11]。磁弹性效应之一为磁致伸缩, 利用大磁致伸缩的软磁制作大功率超声换能器已有多年历史。磁致伸缩材料有金属镍、镍锌铁氧体等。磁致伸缩还可用来实现微小位移, 用于精密控制上。磁致伸缩的逆效应为应力影响磁化曲线。据此原理已制成不同称量的秤, 从克到吨。

磁卡效应的应用由来已久。磁卡效应指物质被磁化和绝热去磁时引起磁熵的变化及温度的变化。低温下利用顺磁盐的绝热去磁以获得超低温已有几十年的历史。近两年利用铁磁体居里点附近较大的磁卡效应以实现致冷, 发展无氟电冰箱的技术吸引了人们的注意。在稀土 Gd 及其化合物、钙钛矿锰氧化物中均发现了巨大磁卡效应的材料。

居里点附近的自发磁致伸缩使强磁体在居里点附近的热膨胀反常, 因而有些合金有很小的或不热的热膨胀系数, 称为英钢 (Invar)。如 Fe-Ni, Fe-Pt, Fe-Ni-Cr 合金, 在玻璃与金属焊接等方面有重要的应用。

磁性液体为由 Fe_3O_4 等超顺磁性的超细微粒高度弥散于有一定粘滞性的液体中形成的悬浮液或胶体, 它既是液体又具有强磁性, 具有特殊用途, 如动态密封, 阻尼, 磁性检疵, 磁畴观察, 磁性墨水及印刷。它们的特点之一是受磁场控制。

10.6 巨磁电阻效应和磁电子学^[8,12]

10.6.1 磁电阻和巨磁电阻

磁电阻是指磁场使电阻发生变化的现象。1857 年凯尔文首先发现了铁的磁电阻。强磁体的磁电阻称为各向异性磁电阻 AMR。来源于磁畴中电阻率的各向异性。通常, 沿磁场方向的磁电阻比 $\left(\frac{\Delta R}{R}\right)_{\parallel} > 0$, 而垂直于磁场方向的 $\left(\frac{\Delta R}{R}\right)_{\perp} < 0$, 可在低磁场下饱和。饱和值约为 1% ~ 5%。正常磁电阻 OMR 普遍存在于所有的金属和半导体中。它来源于磁场对电子的洛伦兹力, $\frac{\Delta R}{R} > 0$, 磁场低时 $\Delta R/R$ 很小, 但无饱和。20 世纪 70 年代, OMR 及 AMR 均用于传感器, 90 年代初 AMR 开始用于硬磁盘读出头。

1988 年在纳米金属 Fe/Cr 多层膜中发现了巨磁电阻 GMR。其 $\Delta R/R$ 为负值, 绝对值比 AMR 可高 1~2 个数量级。图 10-35 为 Fe/Cr 多层膜的磁滞回线和磁电阻比曲线。

巨磁电阻引起了科学与技术界的广泛重视。研究及应用开发迅速发展, 以后在铁磁颗粒与非铁磁金属组成的不均匀合金中也观察到 GMR。

尔后, 两种新的巨大磁电阻效应得到发展: 铁磁/绝缘体/铁磁结构的隧

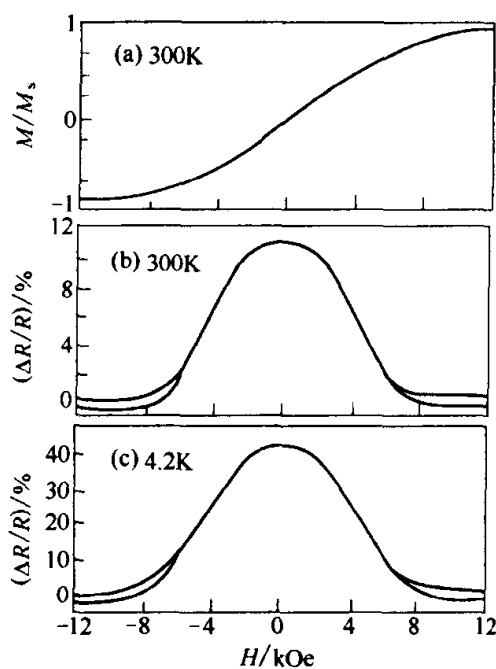


图 10-35 Fe/Cr 多层膜的磁滞回线和磁电阻比曲线

道结巨磁电阻 TMR 和钙钛矿型锰氧化物系统中的庞磁电阻 CMR。前者提供了一种高内阻的巨磁电阻结构, 在应用上有吸引力, 后者出现了高达 $10^6\%$ 以上的磁电阻比, 并有丰富的物理内容。

10.6.2 自旋相关导电

如何解释 AMR, GMR, TMR 和 CMR 呢? 人们提出了自旋相关导电的理论。传统的导电理论基于电子(或离子)电荷转移及电子经受的散射, 而与电子自旋无关。不同自旋电子的导电性能没有区别。自旋相关导电是指, 导电依赖于载流子自旋的方向。不同材料又有不同的自旋相关导电的机制。

① 自旋相关散射。适用于铁磁金属的 AMR, 金属多层结构和金属颗粒膜的 GMR。

② 自旋相关隧穿几率。适用于磁隧道结 MTJ 的 TMR。

③ 自旋极化电子导电。正在 CMR 材料和新的自旋电子器件研究中发展。

(1) 双电流模型 根据半经典理论, 金属电阻率可由下式表示

$$\rho = m^* / ne^2 \tau \quad (10-117)$$

式中, n 为费米面电子的浓度, 只有 Fermi 面上的电子才对导电有贡献; m^* 为传导电子有效质量; τ 为其散射的弛豫时间, 与平均自由程 λ 呈正比,

$$\tau \sim \lambda \sim 1/|V|^2 N(E_F) \quad (10-118)$$

式中, V 为散射矩阵元, $N(E_F)$ 为费米面能态密度。正常金属和半导体中上两式中各参量与电子自旋无关, 因而电阻率与自旋无关。3d 族铁磁金属则不同, 在 T_c 以下, 由于交换作用发生自发磁化, s 和 d 电子不同自旋的次能带发生劈裂, 如图 10-36 所示。

上两式改写如下

$$\rho_\sigma = m_\sigma^* / n_\sigma e^2 \tau_\sigma \quad (10-119)$$

$$\tau_\sigma \sim \lambda_\sigma \sim 1/|V_\sigma|^2 N_\sigma(E_F) \quad (10-120)$$

式中, σ 代表自旋向上及向下, $\sigma = \uparrow$ 或 \downarrow 。因而电阻成为自旋相

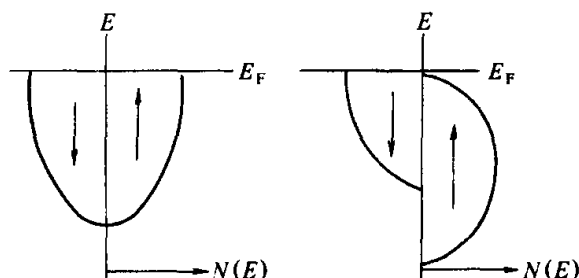


图 10-36 正常金属和铁磁金属正负自旋次能带的态密度示意图

关。 $\sigma = \uparrow, \downarrow$ 相应于与 M_s 方向相同及相反。在大块铁磁金属中, 传导电子因碰撞而不断改变自旋方向。自旋保持不变的平均距离, 自旋扩散长度约为微米量级, 平均自由程为几十纳米, 因此只有在纳米尺度的铁磁体中, 可以认为传导电子在输运过程中自旋保持不变, 可将其导电分解为自旋向上 $\sigma = \uparrow$ 和向下 $\sigma = \downarrow$ 两个相互独立的电子导电通道, 相互并联。各自的电阻分别为 ρ_{\uparrow} 及 ρ_{\downarrow} 如式 (10-119), 总电阻为

$$\rho = \rho_{\uparrow} \rho_{\downarrow} / (\rho_{\uparrow} + \rho_{\downarrow}) \quad (10-121)$$

其值小于 ρ_{\uparrow} 或 ρ_{\downarrow} 中的任一个。此模型称为双电流模型, 最早为莫特提出^[2,8]。

(2) 自旋相关散射与 GMR。一个 FM/NM/FM 三层膜示意于图 10-37。电流与膜面平行 (CIP), 但电子的运动是沿曲折的路径, 在界面上发生自旋相关散射。

若铁磁体 FM 的能带如图 10-36 所示。 $N_{\downarrow}(E_F) > N_{\uparrow}(E_F)$, $\sigma = \downarrow$ 的传导电子受到散射大, 而 $\sigma = \uparrow$ 的电子的散射小, 故 $\lambda_{\downarrow} < \lambda_{\uparrow}$, $\rho_{\downarrow} > \rho_{\uparrow}$, 是为自旋相关散射。当相邻铁磁层的 M_s 反平行时, $\sigma = \uparrow$ 及 $\sigma = \downarrow$ 的传导电子均在 M_s 与其自旋平行的铁磁层界面受到小的散射而在 M_s 与其自旋反平行的 FM/NM 界面受到较大的散射, 总电阻较高。如 10-37 上图所示, 而当两铁磁层的 M_s 平行时, 自旋与 M_s 平行的电阻低, 自旋与 M_s 反平行的电阻高, 前者使总电阻短路, 呈低阻态, 如 10-37 下图所示。因此通常得到负的各向同性的巨磁电阻, 获得 GMR 的条件之一是 $\rho_{\uparrow}/\rho_{\downarrow} \neq 1$ 。

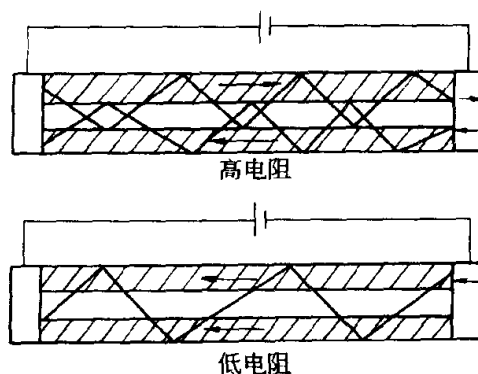


图 10-37 一个 CIP 三层膜的 GMR 示意图

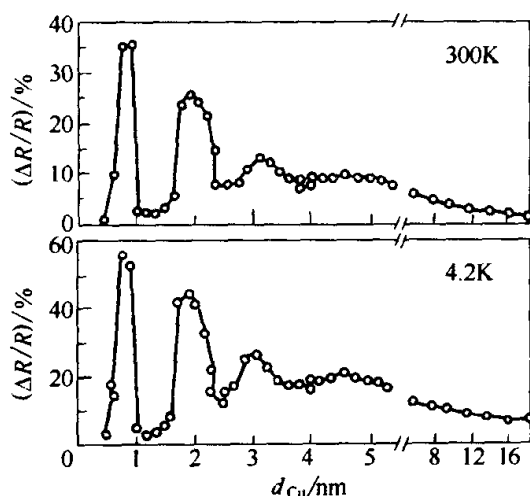


图 10-38 Co/Cu 多层膜的磁电阻比与 Cu 层厚度的关系

阻高, 前者使总电阻短路, 呈低阻态, 如 10-37 下图所示。因此通常得到负的各向同性的巨磁电阻, 获得 GMR 的条件之一是 $\rho_{\uparrow}/\rho_{\downarrow} \neq 1$ 。

(3) 多层膜与自旋阀。GMR 被发现以来, 在多种多层膜中观察到 GMR。并发现 FM/NM 多层膜中铁磁层的磁化通过非铁磁金属层发生层间交换耦合, 并随非磁金属的厚度变化而在铁磁与反铁磁耦合间振荡^[2,5,8]。当反铁磁耦合时可观察到 GMR, $\Delta R/R$ 大。当为铁磁耦合

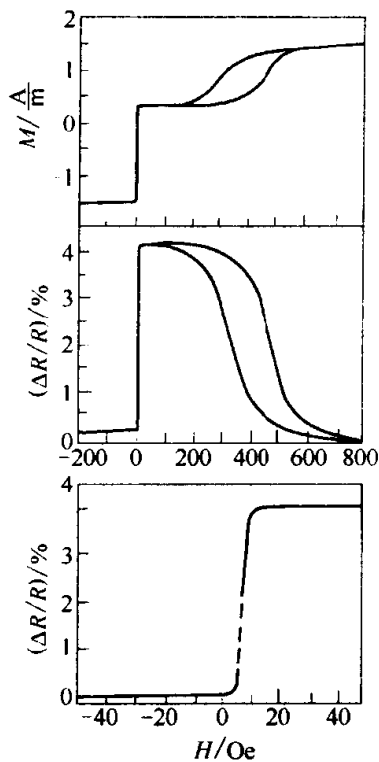


图 10-39 一个自旋阀的磁滞回线和磁电阻曲线

时, 没有 GMR。图 10-38 给出 $\Delta R/R$ 的振荡, 多层膜虽然可有高的 $\Delta R/R$ 值, 但这时处于反铁磁耦合, 故饱和磁场 H_s 高, 磁场灵敏度 $S = \Delta R/R/H_s$ 并不高。

为降低 H_s , 常采用的方法是适当增厚非磁层, 使层间耦合近于零。典型的非耦合的基本结构为 $FM_1/NM/FM_2/AFM$, AFM 为一反铁磁层, 通过其界面的交换耦合, 使相邻的 FM_2 铁磁层的 M_s 钉扎在某个方向, 非磁层 NM 的厚度约 2.3nm, 交换耦合近于零。铁磁层 FM_1 称为自由层, 在很低的磁场下可使其与 FM_2 平行或反平行, 从而得到足够大的 $\Delta R/R$ 和灵敏度, $\Delta R/R/H_s$ 。其基本性能见图 10-39, 这种结构称为自旋阀, 为实用化多层膜的主要结构。

(4) 磁性隧道结 MTJ 和其 GMR 效应 以绝缘层 I 代替自旋阀的非磁金属层即为 MTJ, 常用的基本结构为 $FM_1/I/FM_2/AFM$ 。但工作电流垂直于膜面 (CPP)。由于绝缘层 I 的厚度仅 1~2nm, 可有隧道电流跨过 I 层, 隧道电流

与费米面的能态密度 $N(E_F)$ 相关, 且与自旋相关。从图 10-40 可以看出, 当 FM_1 及 FM_2 的 M_s 平行时为低电阻态, 当两者相反时为高电阻态, 通常其 GMR 效应为负值, 各向同性。磁性隧道结的磁电阻比值可高于自旋阀, 最近数据^[12], 室温下 $\Delta R/R$ 达到 50%, 且为高内阻, 故在应用上有吸引力。此外隧道结可用以研究铁磁金属中隧道电子自旋的极化度。

(5) 自旋极化电流及 CMR 材料 从图 10-36 可以理解铁磁体中的传导电流是自旋极化的。利用超导/绝缘/铁磁金属的隧道结曾给出 Fe, Co, Ni 的隧道电流的自旋极化度 P 的数值为 +40%、+35% 和 +23%。新近的数据更高些, 可达 50%, 自旋极化度的定义为 $P = \frac{n_{\uparrow} - n_{\downarrow}}{n_{\uparrow} + n_{\downarrow}}$ 。 n_{\uparrow} 和 n_{\downarrow} 为电子流中自旋向上及向下的浓度。 P 为正, 意味着电流的自

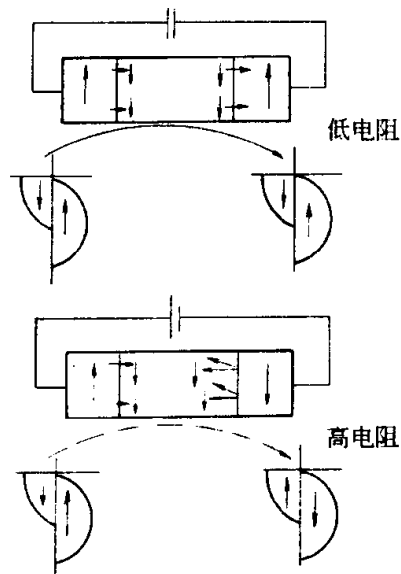


图 10-40 一个 CPP 结确立 GMR 的示意图

旋极化方向与 M_s 的方向相同, 这显然与图 10-36 的结论相反。过去曾认为正的 P 意味着隧道电流主要来源于 s 电子, 而 s 带被 d 的自发磁化在相反方向交换劈裂。但新近发现, 隧道电流 P 的符号及大小与绝缘材料的选择有关, 而不完全决定于铁磁层的材料^[12]。

传导电子自旋极化度 $P=1$ 的材料是很有意义的, 若用这种材料制成 CPP 的隧道结或金属三层膜的两个 FM 层电极, 则当磁场使两铁磁层从平行变为反平行时, 其电阻可以从金属性导通变为绝缘性能。半金属材料 and 掺二价离子的钙钛矿型锰氧化物等, 原则上为 $P=1$ 的材料。

$RE_{1-x}M_x^{2+}MnO_3$ 属钙钛矿结构, 是一大类材料, RE 为稀土离子 La, Nd 等, M 为二价金属离子如 Ca, Sr, Ba, Pb 等。当掺杂量合适, 如 x 约为 1/3 左右时, 顺磁铁磁相变和绝缘/金属相变在相近的温度下出现, 电阻率在相变点附近出现极大值, 这种材料在磁场作用下出现了相变点的升高及超大的磁电阻, 图 10-41 为 $Nd_{0.7}Sr_{0.3}MnO_3$ 薄膜的电阻率在零场下及 8T 的磁场作用下的 $\rho \sim T$ 曲线及磁电阻比 $\Delta R/R \sim T$ 的关系。其磁电阻比超过了 10^6 , 即磁场使电阻率改变 6 个数量级, 因而被称为庞磁电阻 CMR。

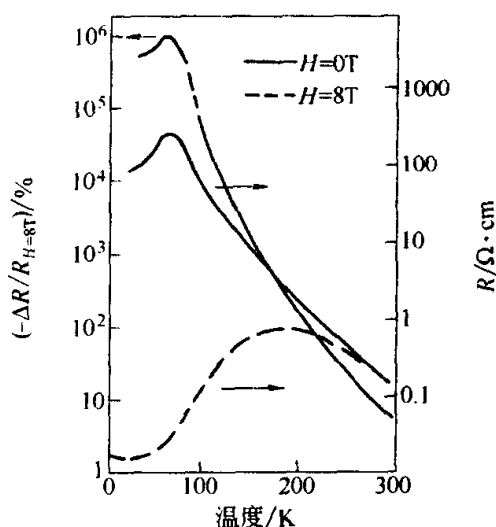


图 10-41 $Nd_{0.7}Sr_{0.3}MnO_3$ 薄膜的电阻率和磁电阻随温度的变化

这类材料的磁相变, 金属/绝缘体相变和庞磁电阻的机制为当前的研究热点。早在 20 世纪 50 年代, Zener 等人就提出双交换作用来解释这类材料的铁磁性及金属性导电共存的现象。掺二价金属后为了保持电中性该化合物中的锰离子出现变价。 Mn^{3+} 及 Mn^{4+} 共存。 Mn^{3+} 上的 4 个 d 电子中 3 个电子为局域的低能态, 一个电子为高能态 e_g , Mn^{4+} 中的 3 个 d 电子为局域态, 高能的 e_g 为空轨道。 $Mn^{3+}-O^{2-}-Mn^{4+}$ 中 O^{2-} 的一个 p 电子可能转移到 Mn^{4+} 的空 e_g 轨道, 同时 Mn^{3+} 中的 e_g 电子转移给氧离子。由于洪德法则的限制, 这个电子自旋方向必须与 Mn 离子磁矩相同, 这种双交换作用导致了铁磁性及金属性电导, 而且导电电子的自旋方向均与 M_s 的方向相同。故原则上其自旋极化度 $P=1$ 。但单纯的双交换作用给出的电阻值与磁相变温度远不能与实验一致。进一步深入研究考虑了电子-声子相互作用形成的极化子效应等仍在进行。由于这类材料的 T_c 低, 且 CMR 出现在 T_c 附近, 巨大磁电阻出现在低温及强场下。如何提高 CMR 材料的 T_c 及降低工作磁

场就成为获得应用的重要方向。另一方面,将 $P=1$ 的 CMR 材料作为隧道结的磁性层以获得高的隧道磁电阻也是有意义的课题。

10.6.3 磁电子学和巨磁电阻的应用^[2,5,8,12]

(1) 磁电子学^[8,12] 巨磁电阻效应于 1988 年出现以来,受到全世界科技界的广泛而持久的重视,其原因一方面是由于重要的应用价值,另一方面,自旋相关导电将导电与自旋联系起来,使传统的电子学和微电子学只考虑电荷的运输的面貌发生变化,增加了通过调节电子自旋的新功能。一个新的学科,磁电子学或自旋电子学正在形成。目前正在探索的有以下几个重要的新方向。

① 基于自旋极化输运特性的全金属晶体管。可能的结构类似于自旋阀但具有发射极、基极和接受极。与半导体晶体管比,它的最大特点是尺寸可以更小及有记忆能力。

② 将自旋极化电流注入到超导体内以寻求新的现象及应用。

③ 将自旋极化电流注入到半导体中,或寻求室温下的铁磁半导体,以便开发新的自旋相关,有记忆能力的微电子器件并与集成电路技术相结合。最近已发现了高于室温的透明磁性半导体 $\text{Ti}_{1-x}\text{Co}_x\text{O}_2$ ($x=0.06$)。

④ 寻求 $P=1$ 的新材料及工艺以实现制备(绝缘/金属)的开关材料。

⑤ 最近几年,出现了一个新的发现——自旋转移,用电流直接使铁磁体的磁矩反向或激发自旋波,它是自旋相关导电的逆效应。

(2) 自旋转移^[12] 自从古代发现磁铁矿以来,只有用磁场才能使磁矩转动。在物理上和技术上这都是多年以来公认的事实。因此,用电流直接使磁矩转动、反转或激发自旋波,而不通过磁场是物理上新的突破。简称为自旋转移。

纳米尺度的磁性多层膜中相邻铁磁层磁矩反平行排列时的电阻大于平行排列。通常电流中自旋方向与铁磁层磁矩方向相反的传导电子受到该层磁矩更强烈的散射,因而具有更大的电阻;而自旋方向与铁磁层磁矩方向相同的传导电子受到的散射较弱,因而具有较小的电阻。巨磁电阻的逆效应可从牛顿第三定律来思考:电子自旋有磁矩和动量矩,电子与铁磁层磁矩散射过程有动量矩的交换。不同自旋的传导电子受到铁磁层的散射既然不同,铁磁层受到来自不同自旋电子动量矩的交换也应不等。因此,多层膜中的磁矩被扰动或转动,导致自旋波的激发甚至磁矩反转。换言之,电流的自旋相关散射或自旋极化电子流通过一个小的铁磁导体时,电流将一定的自旋动量矩转移给磁矩系统,使其磁矩转动、进动,甚至反转其方向。这个过程称为自旋转移。

虽然多层膜的巨磁电阻在 1986~1988 年发现并一直是磁学和材料科学的研究热点,但巨磁电阻的逆效应则被忽略。对自旋转移的理论预言迟至 1996 年才给出。

迄今已在 Co-Cu 多层膜, 颗粒合金, Ni 纳米线和锰氧化物隧道结中证实了电流直接使磁矩发生反转或激发自旋波。

自旋转移是一个物理上新的突破, 也是对自旋相关散射和自旋极化电流更全面的认识。

(3) 实用化的 GMR 器件^[8] 目前已经或正在实用化的磁电子学器件均基于 GMR 及 TMR 效应, 主要有 3 类。

① 传感器。GMR 传感器于 1994 年进入市场, 其性能优于已在市场上的半导体和 AMR 磁电阻传感器。磁电阻传感器可以传感磁场, 特别是对微弱磁场的传感。如可用于伪钞识别器, 弱磁场测量等。更广泛地应用是各类运动传感器, 如对位置、速度、加速度、角度、转速等的传感, 在机电自动控制、汽车工业和航天工业等方面有广泛地应用。

② 磁记录读出磁头。读出磁头属于磁场传感器中最为突出的应用。图 10-42 为 GMR 读出磁头的示意图。

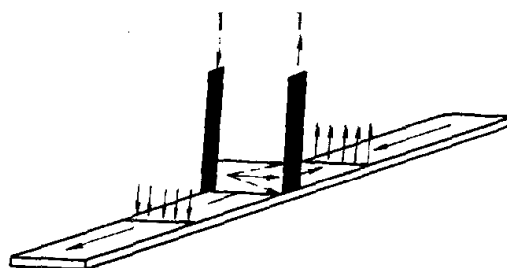


图 10-42 GMR 读出头的示意图

20 世纪 90 年代以来, 利用 AMR 读出头使磁盘记录面密度达到每平方英寸 1 千兆位 ($1\text{Gb}/\text{in}^2$) 及 $3\text{Gb}/\text{in}^2$, 用 GMR 多层膜 (自旋阀) 进一步提高到 $5\text{Gb}/\text{in}^2$ 、 $10\text{Gb}/\text{in}^2$ 及 $60\text{Gb}/\text{in}^2$ 。最新用隧道结的报道达 $100\text{Gb}/\text{in}^2$ ^[12]。

③ 磁随机存储器 MRAM。最近, 在巨磁电阻用于计算机内存的主要组成部分——随机存储器 RAM 方面获得较大进展。1995 年报道了开关速度为亚纳秒的自旋阀型 MRAM 记忆单元及由 16Mb 的 MRAM 晶片组成的 256Mb 的 MRAM 芯片的设计报告。最近新的报道, 已用自旋阀及隧道结制备了 16Kb 及 1Mb 的芯片, 预计几年可能上市, 很有竞争能力。表 10-10 为对 MRAM 性能预测及与其他 RAM 的对比。

表 10-10 对 MRAM 性能的预测及与半导体存储器的对比

存储器	DRAM	FLASH	SRAM	MRAM
容量	256Gb	256Gb	$180\text{Mb}/\text{cm}^2$	$>256\text{Gb}$
速度	150MHz	150MHz	913MHz	$>500\text{MHz}$
寻址时间	10ns	10ns	1.1ns	$<2\text{ns}$
写入时间	10ns	$10\mu\text{s}$		$<10\text{ns}$
擦除时间	$<1\text{ns}$	$10\mu\text{s}$		
保持时间	2.4s	10 年		无穷
写入擦除次数	无穷	10^5	无穷	无穷

注: 表源于 Wolf, S. A. et al. IEEE MAG. 2000, 36: 2137。

参 考 文 献

- 1 Morrish. A. H. The physical principles of magnetism. New York: John Wiley & Sons, Inc., 1965
- 2 翟宏如. 金属磁性. 见冯端. 金属物理. 第四卷. 第十四篇. 北京: 科学出版社, 1998
- 3 Van Vleck. J. H. The Theory of Electric and Magnetic Susceptibilities. London: Oxford University Press, 1932
- 4 Smit J. and Wijn H. P. J. Ferrites, The Netherlands: Philips Technical Library. 1959
- 5 O'Handley. Robert C. Modern Magnetic Materials. New York: John Wiley & Sons, Inc., 2000
- 6 戴道生, 钱昆明, 钟文定, 廖绍彬. 铁磁学. 北京: 科学出版社, 1987
- 7 Mansuripur M. Physical Principles of Magneto-optical Recordings. London: Cambridge University Press, 1993
- 8 钟山系列会议磁电子学选集. 物理学进展. 1997, 17:2
- 9 Vonsovskii S. V. Magnetism. New York: John Wiley & Sons, Inc., 1971
- 10 近角聪信. 强磁性体の物理. 东京: 裳华房, 1978
- 11 都有为. 铁氧体. 南京: 江苏科技出版社, 1996
- 12 翟宏如. 第三届全国磁性薄膜与纳米磁性学术会议论文集. 兰州. 2001

第 11 章 铁电性质

撰稿人 刘治国

铁电体 (ferroelectrics) 是电介质的一个亚类, 其基本特征是具有自发电极化并且这种电极化可以在外电场作用下改变方向。由于自身结构的原因, 铁电体同时具有压电性和热释电性, 此外一些铁电晶体还具有非线性光学效应、电光效应、声光效应、光折变效应和反常光生伏打效应。铁电体这些性质使它们可以将声、光、电、热效应互相联系起来, 成为一类重要的功能材料。本章拟对电介质、铁电体、压电材料和热释电材料作一介绍。关于铁电体的光学性质在本书第 13 章会有简单地介绍。

11.1 电介质的极化与铁电性

11.1.1 电介质的极化

材料可按其对外电场的响应方式区分为两类, 一类以电荷长程迁移即传导的方式对外电场作出响应, 这类材料称为导电材料。另一类以感应的方式对外电场作出响应, 即沿电场方向产生电偶极矩或电偶极矩的改变, 这类材料称为电介质, 这种现象称为电介质的极化。广义而言, 一些导电材料也可表现出介电性, 但这不是电介质物理的研究范畴。通常, 绝缘体都是典型的电介质。

电介质又可分为非极性电介质和极性电介质两大类。前者由非极性分子组成, 在无外加电场时分子的正负电荷重心互相重合, 不具有电偶极矩。只是在外加电场作用下正负电荷出现相对位移, 才出现电偶极矩。后者由极性分子组成, 即使是在无外场时每个分子的正负电荷重心也不互相重合, 具有固有电矩, 它与铁电性有密切关系。

电介质的极化有 3 种主要基本过程, 即材料中原子核外电子云畸变产生的电子极化; 分子中正、负离子相对位移造成的离子极化和分子固有电矩在外电场作用下转动导致的转向极化。此外, 与材料中的缺陷相关联还有一些其他极化机制, 但其重要性不如前 3 种。

描述一个中性分子电极化状态的参量是电偶极矩, 其表达式为

$$p = ql \quad (11-1)$$

其中 q 为分子中正电荷的总量 (显然其中负电荷的总量为 $-q$), l 为正负电

荷重心之间的位矢，其方向由负电荷重心指向正电荷重心。介质的电极化强度是描述电介质极化状态的基本物理量，其定义为该电介质单位体积内电偶极矩的矢量和，即：

$$\mathbf{P} = \sum \mathbf{p} / V \quad (11-2)$$

表征电介质在静电场作用下介电性质的主要参量为介电常数 ϵ 。由电学知识可知，电介质电位移矢量 \mathbf{D} 与外加电场 \mathbf{E} 的关系为

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} = \epsilon \epsilon_0 \mathbf{E} \quad (11-3)$$

其中 \mathbf{P} 为电极化强度，在电场较小的情况下有

$$\mathbf{P} = \chi \epsilon_0 \mathbf{E} \quad (11-4)$$

式中， χ 称为介电极化率， ϵ_0 为真空介电常数，其数值为 $\epsilon_0 = 8.85 \times 10^{-2}$ F/m，由上述两式可求出

$$\mathbf{D} = \epsilon_0 (1 + \chi) \mathbf{E} \quad (11-5)$$

由此可知，介电常数 ϵ 与介电极化率 χ 只差一个常数 1，它们可以用来描述介质的介电性质。在各向同性介质中 \mathbf{D} 、 \mathbf{P} 、 \mathbf{E} 方向相同， ϵ 和 χ 均为标量。在各向异性介质如单晶体中 \mathbf{D} 、 \mathbf{P} 、 \mathbf{E} 的方向并不一定相同，由于 \mathbf{D} 、 \mathbf{P} 、 \mathbf{E} 均为矢量，亦即一阶张量，将它们联系起来的物理常数 ϵ_0 和 χ 为二阶张量。

在均匀极化的电介质内部其各分子电偶极矩的正端总是和另一电偶极矩的负端相连，正负端的电荷恰好互相抵消而使介质内部呈现电中性，但在电介质的表面这种抵消不再能实现。电偶极矩正端露头处便呈现正电荷，负端露头处呈现负电荷，由于这些电荷束缚于电偶极子，不能自由移动，这种电荷称为束缚电荷。束缚电荷的面密度 σ 满足

$$\sigma = \mathbf{P} \cdot \mathbf{n}_0 \quad (11-6)$$

其中 \mathbf{n}_0 为表面外法线上的单位矢量。这些面束缚电荷在电介质内部产生一个电场，称为退极化电场，其方向与 \mathbf{P} 相反。若电介质受交变电场作用，其极化行为与在静电场作用下的行为有所不同，特别是在较高频率外电场作用下，电位移矢量 \mathbf{D} 的变化滞后与外电场 \mathbf{E} 的变化，可导致极化强度 \mathbf{P} 的减小和介电常数 ϵ 的下降。考虑位于外电场 $E(t) = E_0 \cos \omega t$ 中的电介质，其电位移矢量也应作简谐变化，但具有滞后相位 $\delta(\omega)$ ，这里 δ 为 ω 的函数，即 $D(t) = D_0 \cos[\omega t - \delta(\omega)]$ ，容易求出

$$D(t) = \epsilon'(\omega) E_0 \cos \omega t + \epsilon''(\omega) E_0 \sin \omega t \quad (11-7)$$

其中

$$\begin{aligned} \epsilon'(\omega) &= \frac{D_0}{E_0} \cos[\delta(\omega)] \\ \epsilon''(\omega) &= \frac{D_0}{E_0} \sin[\delta(\omega)] \end{aligned} \quad (11-8)$$

这意味着在交变电场作用下介质的介电特性要用两个物理量 ϵ' 和 ϵ'' 来表示，

它们都是频率的函数。此时电位移矢量可分解为两个分量，一个与外电场同位相，另一个比外电场落后 90° 位相。令 $D^*(t) = D_0 e^{i(\omega t - \delta)}$, $E^*(t) = E_0 e^{i\omega t}$ ，这种关系也可方便地用复数表示为

$$D^*(t) = \epsilon^*(\omega) \cdot E^*(t) \quad (11-9)$$

其中复介电常数 $\epsilon^*(\omega) = \epsilon'(\omega) - i\epsilon''(\omega)$ (11-10)

式中， $\epsilon'(\omega)$ 为复介电系数的实部，它是静电介电系数在复变电场下的推广； $\epsilon''(\omega)$ 为复介电系数的虚部，它正比于介质在交变电场下的能量损耗。

介质的极化特性与其晶体结构有着深刻的内在联系。按照其对称性，晶体可分为 7 大晶系，32 种点群。对这 32 种点群的具体分析表明属于所有这 32 种点群的任一种的材料都可具有介电性。这些点群中有 20 种点群不具有中心对称，具有这 20 种点群结构的晶体的电偶极矩可因弹性形变而改变，因而具有压电性并被称为压电体，在压电体中具有惟一极轴（又称自发极化轴）的 10 种点群可出现自发极化，即在无外电场存在的情况下也存在电极化。一般说来，自发极化可因温度的改变而变化，具有这 10 种点群的晶体可出现热释电性，并被称为热释电体。热释电晶体中又有一部分晶体其自发电极化可在外电场的作用下改变方向，而且电极化矢量 \mathbf{P} 与外电场 \mathbf{E} 呈类似于磁滞回线的关系，这些晶体被称为铁电体。显然，铁电体同时具有热释电性、压电性和介电性，因而这类晶体受到人们特别的关注。晶体的介电性、压电性、热释电性及铁电性相应的对称群和典型晶体见表 11-1。关于电介质物理与电介质材料更详细地介绍可参阅文献 [2]。

表 11-1 晶体结构对称性与其介电性的关系^[1]

性能特征	对 称 群	代 表 晶 体
介电性	32 种点群	高岭土 $\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})$, ($\bar{1}$) 硫酸铜 $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, ($\bar{1}$) 亚硝酸钾 KNO_2 , (m) 正长石 KAlSi_3O_8 , ($2/m$) 刚玉 Al_2O_3 , ($\bar{3}m$)
压电性	$1, 2, 4, 3, 6, \bar{6}$ $m, 2mm,$ $4mm, 3m,$ $6mm, \bar{6}m2$ $222, 422, 322,$ $622, 23$ $\bar{4}, \bar{4}2m,$ $\bar{4}3m$ 20 种点群	低温石英 SiO_2 , (322 , 或 32) 罗息盐 $\text{NaKC}_4\text{H}_4\text{O}_6 \cdot 4\text{H}_2\text{O}$, 简称 RS, (2) 锗酸铋 $\text{Bi}_{12}\text{GeO}_{20}$, (23) 氧化锌 ZnO , ($6mm$) 硫化锌 ZnS , ($\bar{4}3m$) 高温石英 SiO_2 , (622) 锗酸铅 $\text{Pb}_5\text{Ge}_3\text{O}_{11}$, ($\bar{6}$) 铌酸锂 LiNbO_3 , ($3m$)

续表

性能特征	对 称 群	代 表 晶 体
热(释)电性	1,2,4,3,6 $m, mm2,$ $4mm, 3m,$ $6mm$	铌酸锂 $\text{LiNbO}_3, (3m)$ 钛酸钡 $\text{BaTiO}_3, (4mm, mm2, 3m)$ 碘硫化铋 $\text{SbSI}, (mm2)$ 硫酸三甘酐 $(\text{NH}_2\text{CH}_2\text{COOH})_3 \cdot \text{H}_2\text{SO}_4$, 简称 TGS, (2) 硫化镉 $\text{CdS}, (6mm)$
铁电性	1,2,4,3,6 $m, mm2, 4mm,$ $3m, 6mm$	磷酸二氢钾 KH_2PO_4 , 简称 KDP, ($mm2$) β -钼酸钆 $\beta\text{-Gd}_2(\text{MoO}_4)_3, (mm2)$ 铌酸锂 $\text{LiNbO}_3, (3m)$ 钛酸钡 $\text{BaTiO}_3, (4mm, mm2, 3m)$ 钛酸铅 $\text{PbTiO}_3, (4mm)$ 六水硫酸铝胍 $\text{C}(\text{NH}_2)_3\text{Al}(\text{SO}_4)_2 \cdot 6\text{H}_2\text{O}$, 简称 GASH, ($3m$) 硬硼钙石 $\text{CaB}_3\text{O}_4(\text{OH})_3 \cdot \text{H}_2\text{O}, (2)$ 亚硒酸三氢锂 $\text{LiH}_3(\text{SeO}_3)_2, (m)$

11.1.2 自发极化与铁电性

1920 年法国人瓦拉赛克 (Valasek) 首次报道了第一个被发现的铁电体酒石酸钾钠 ($\text{NaKC}_4\text{H}_4\text{O}_6 \cdot 4\text{H}_2\text{O}$) 的奇异电学性质, 这标志着铁电体研究的开端。20 世纪 50 年代以来人们发现和合成的铁电体种类急剧增加, 现已达到 200 多种。早年铁电体是科学家实验室中的珍品, 被当作研究结构相变的典型材料。20 世纪 80 年代以来随着高科技对新材料需求的日益扩大和材料制备技术的迅速提高, 铁电体作为一类新型功能材料而崭露头角。

如前所述, 铁电体是一类具有自发极化, 而且这种自发极化可以在外电场作用下改变方向的电介质。因此, 自发极化及其变化便成了铁电体研究的核心问题。通常晶体的铁电性只存在于某一特定温度之下, 在此温度之上铁电体变成顺电体, 其自发极化消失, 铁电相与顺电相之间的转变称为铁电相变, 该温度称为居里温度 T_c 。铁电相变可按其结构转变的特征分为两类, 即位移型铁电相变和有序无序型铁电相变。铁电相的晶格参数一般与具有较高对称性的原顺电相的差别很小, 相关的结构分析实验必须有较高的精度。

按其结构和自发极化产生的机制, 铁电体可分为四类, 它们是含氧八面体的铁电体、含氢键的铁电体、含氟八面体的铁电体和含其他离子基团的铁电体, 详情可阅读参考文献 [3]。现仅以最具代表性和为数最多的一类铁电体——钙钛矿型铁电体为例进行较详细地介绍。钙钛矿结构这一名词源自人们在很早以前发现的矿物钙钛矿。其分子式为 ABO_3 , 其中 A、B 分别为不同的金属离子, O 为氧离子。20 世纪 80 年代以来由于新发现的高温超导体和超大磁电阻材料也都有 ABO_3 结构, 钙钛矿结构成了物理学家的新宠。细

察具有这种结构的晶体的奇异特性包括铁电性，人们发现它们都和其中的“氧八面体”有关。图 11-1(a)为钙钛矿结构晶体原胞的示意图。金属离子 A 位于立方体的 8 个顶角上，氧离子 O 位于六个面心位置上并构成氧八面体的 6 个顶点。金属离子 B 位于原胞的中心，同时也位于氧八面体的中心。显然，没有畸变的钙钛矿结构具有对称中心，其正负电荷重心互相重合，因而不具有自发极化，也不可能具有压电性。钙钛矿结构的一个重要特征是其结构的不稳定性。某些这类晶体如钛酸铅(PbTiO_3)在高温下具有标准的钙钛矿结构，但随着温度下降和晶格中离子振动减弱，位于氧八面体中心的 B 离子即钛离子变得不稳定而出现向氧八面体某一顶角方向移动的倾向，并终于在 490°C 发生位移型铁电相变，由顺电相变为铁电相。在此过程中离子移动的细节见图 11-1 (b)。显然该图中画出的只是钛离子移动的一种可能方式。考虑到顺电相 PbTiO_3 具有立方结构，实际上钛离子的位移可以向着氧八面体六个顶角中的任意一个发生。在钛离子移动的同时位于原胞顶角上的铅离子也向相同的方向移动，其后果是该原胞由立方结构变成了四方结构，同时原胞中正电荷的重心不再与负电荷的重心重合，从而导致了自发极化的出现。自发极化的方向与四方铁电相的 c 轴一致。

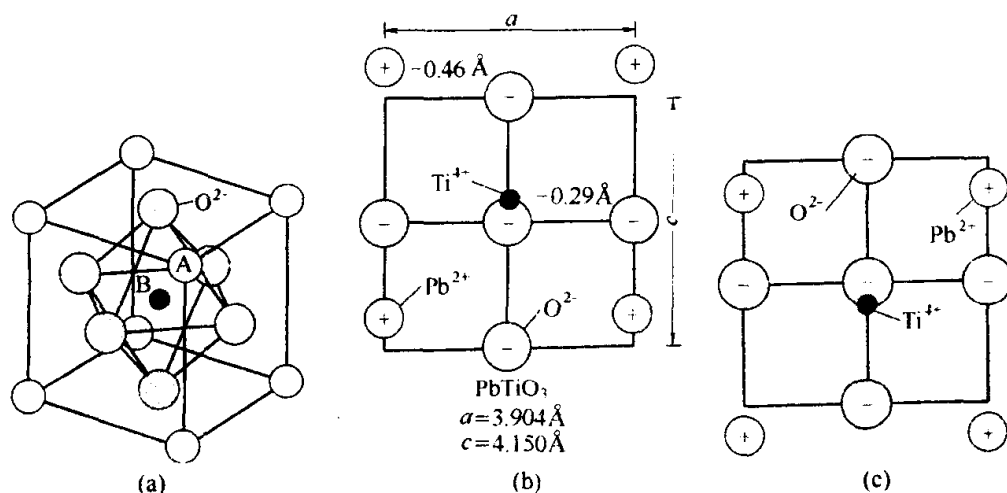


图 11-1

- (a) 具有分子式 ABO_3 的钙钛矿结构示意图；
 (b) PbTiO_3 中离子相对位移导致了自发极化产生 ($1\text{\AA}=0.1\text{nm}$)；
 (c) PbTiO_3 中离子相对位移的另一种模式，它导致了相反的自发极化

一些典型的位移型铁电体的居里点，相变前后的空间群及与自发极化相关的参量列于表 11-2。

另外一些铁电体中其某种原子或原子团在晶胞中有两个或两个以上平衡位置。在较高温度下这些原子或原子团无序地占据这些平衡位置。其总体的自发极化为零，为顺电相。在较低温度下 (T_c 以下)，这些原子或原子团优

表 11-2 一些位移型铁电体相变的参量^[3]

化合物	T_c/K	空间群	同极原子位置对称性	原点移动/ \AA	同极原子	$\Delta z/\text{\AA}$	$P_s/10^{-2}\text{C}\cdot\text{m}^{-2}$
NaNbO_3	73 ± 10	三角 $\rightarrow P6ma$			Nb	[0.060]	11.7 ± 5
SbSI	296 ± 2	$Pna2_1 \rightarrow Pnam$	m	0.041	Sb	0.144 ± 0.040	25 ± 3
$\text{Ba}_{1.25}\text{Sr}_{3.75}\text{Nb}_{10}\text{O}_{30}$	348 ± 15	$P4bm \rightarrow P\bar{4}b2$	222	0.013	Nb	0.106 ± 0.022	
$\text{Pb}_{10}\text{Fe}_3\text{Nb}_5\text{O}_{30}$	388 ± 15	$R3m \rightarrow R\bar{3}m$	$\bar{3}m$	0.118	Nb	0.091 ± 0.100	
BaTiO_3	399 ± 5	$P4mm \rightarrow Pm3m$	$m3m$	0.075	Ti	0.132 ± 0.009	25 ± 1
$\text{Ba}_5\text{Ti}_3\text{Nb}_5\text{O}_{30}$	505 ± 15	$P4bm \rightarrow P\bar{4}b2$	222	0.052	Ti, Nb	0.174 ± 0.100	
KNbO_3	708 ± 5	$Bmm2 \rightarrow Pm3m$	$m3m$	0.160	Nb	0.160 ± 0.014	30 ± 3
PbTiO_3	763 ± 15	$P4mm \rightarrow Pm3m$	$m3m$	0.465	Ti	0.299 ± 0.040	
LiTaO_3	891 ± 5	$R3c \rightarrow R\bar{3}c$	$\bar{3}$	0.197	Ta	0.197 ± 0.008	50 ± 2
$\text{Bi}_4\text{Ti}_3\text{O}_{12}$	949 ± 5	$Fmm2 \rightarrow Fmmm$	mm		Ti	[0.215]	50 ± 10
LiNbO_3	1468 ± 5	$R3c \rightarrow R\bar{3}c$	$\bar{3}$	0.269	Nb	0.269 ± 0.006	71 ± 2
$\text{Ba}_{10}\text{Cu}_5\text{W}_5\text{O}_{30}$	1473 ± 15	$P4mm \rightarrow Pm3m$	$m3m$	0.431	W	0.328 ± 0.100	

先占据其中的某一个平衡位置,因而总体的自发极化不再为零,为铁电相。这时铁电相的出现是无序有序相变导致的,这类铁电体被称为有序无序型铁电体。其典型代表为 KH_2PO_3 。在 KH_2PO_3 中氢离子即质子的有序化导致了这种晶体出现自发极化和铁电性。其详细介绍可参阅文献[3]。实际上许多铁电体同时具有位移型和有序无序型特征。

铁电相变是典型的结构相变,物理学家对它一直有浓厚的兴趣。20 世纪 40 年代以来人们首先完善了铁电相变的宏观理论,它以朗道的唯象理论为基础,以自发极化强度作为标志铁电相变的一个序参量,并将序参量的出现和晶体对称性的降低联系起来,从而可以从具有较高对称性的顺电相对称群中寻找相变后产生的铁电相的对称群。关于铁电相变的宏观理论详情可参阅文献[4]。依据这些理论,铁电相变可分为一级相变和高级相变。一级相变中系统自由能的一阶导数在相变点不连续,亦即熵和体积的变化是不连续

的，因而有相变潜热。在二级相变中熵和体积的变化是连续的，而自由能的二阶导数所确定的一些热力学量，如比热容和压缩率则有不连续的跃变。三级相变或更高级的相变可由此类推。铁电体的介电系数在居里点附近常出现反常性增大。根据铁电相变的热力学理论，铁电体居里点以上顺电相的介电反常可由居里-外斯定律描述，即：

$$\epsilon = \frac{C}{T - T_c} \quad (T > T_c) \quad (11-11)$$

式中， C 为居里-外斯常数。位移型铁电体的特点之一是 C 较大，在 10^5K 量级。而有序无序型铁电体的 C 值较小，通常在 10^3K 量级。 PbTiO_3 的介电常数与温度的关系见图 11-2，它是典型的一级相变铁电体。在相变点介电系数变得非常大，由图还可以看出相变点附近介电系数与温度的关系较好地满足居里-外斯定律。

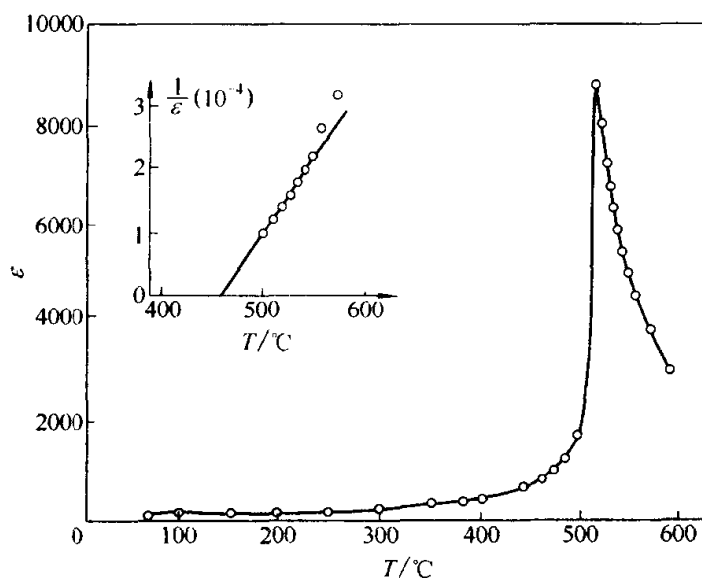


图 11-2 PbTiO_3 单晶沿 C 方向介电系数随温度的变化^[5]

从原子水平上描述铁电性起因的微观理论直到 20 世纪 60 年代才取得重大进展，这就是关于铁电相变的“软模”理论^[6,7]。这一理论揭示了铁电相变的共性并指出铁电相变是更具一般性的结构相变的特殊情况。“软模”理论以晶格动力学为基础。玻恩和黄昆早就指出，一个晶体可以稳定存在的条件是它对于其晶格中所有频率的简正振动是稳定的。这意味着各格点因振动偏离其平衡位置时总存在一个作用使它回到平衡位置，或者说所有简正振动的频率均为不等于 0 的有限大实数^[8]。然而这些精辟见解在相变研究领域并未受到很大注意，直到考科仑和安德森提出铁电相变是与布里渊区中心光学横模的软化联系在一起的。我们不妨这样来理解“模的软化”即软模：晶格的某振动模可用一系列轻弹簧连接的小球来描写，在弹簧作用下小球各自在某

平衡位置附近作振动。如果由于外界条件的变化导致弹簧逐渐变软，则小球的振动频率将逐渐下降。若弹簧完全变软即弹性系数变为 0，则小球偏离其平衡位置后没有外力使其回到原位置。这样原子偏离平衡位置的振幅就成了静态位移，这被称作该振动模被“冻结”。这意味着由于格点间交互作用的变化，格点将到达新的平衡位置。光学模代表着原胞中格点之间的相对运动，包括正负离子之间的相对移动，这类模的软化所导致的正负离子间的相对移动不仅改变了晶格的对称性，而且导致了自发极化的产生或变化。布里渊区中心的光学模，即为波矢很小、波长很大的光学模，在甚小于波长尺度的

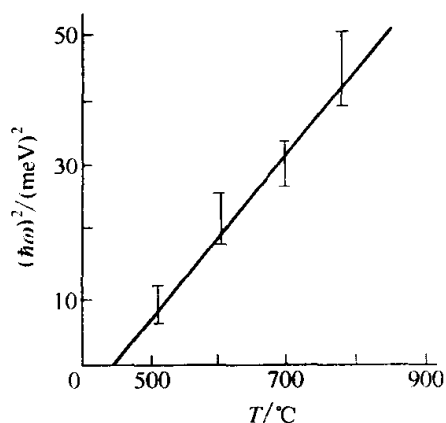


图 11-3 PbTiO_3 在居里点附近的软模

的范围内各晶胞中相应的离子同步运动。如果这种振动模被冻结，这些晶胞中的正负离子具有同样的相对位移，于是晶体呈现均匀的自发极化，或者说出现了结构相变。至于为什么只有“横模”起作用，这是由于理论计算表明光学纵模恒为正值，不会出现软化，只有光学横模有频率降为 0 而软化的可能性。图 11-3 为 PbTiO_3 在相变点的软模行为。其一个振动模频率的平方随温度线性下降至 0。研究相变中软模的实验手段主要是喇曼散射、中子散射和红外光谱。

软模的理论计算较为繁复，这里不做介绍，有兴趣者可参阅文献 [2, 3]。软模的实验和理论方法不仅适用于位移型相变，也适用于有序无序型相变，不仅适用于铁电相变，也适用于其他结构相变。

铁电晶体内自发极化一致的区域称为电畴。在理想情况下，一块铁电单晶可以是单畴的，这种单畴晶体的对称性即为铁电相的对称性，实际铁电单晶往往是多畴的。铌酸锂是单轴铁电体，其自发极化可以沿 c 或 $-c$ 方向，两种极化方向的夹角为 180° ，极化方向相反的区域称为 180° 畴，其界面称为 180° 畴界。对另一些晶体如 PbTiO_3 ，它们居里点以上的顺电相具有立方钙钛矿结构。当温度下降至居里点时，自发极化可能沿六个 $\langle 100 \rangle$ 方向发生，因而可能存在六种铁电畴，相邻两畴自发极化间的夹角可以为 180° 或 90° ，分别称为 180° 畴和 90° 畴，其晶胞和自发极化的示意图见图 11-4。由于铁电相有四方结构，其晶格参数 a 与 c 有些差异， 90° 畴壁有较大的弹性畸变。

稳定的电畴结构应使系统的自由能最低，自由能主要由静电能、退极化能、弹性能构成，畴的形态取决于三者的竞争和平衡。此外，畴结构还与晶体的历史有关。一般说来为了降低静电能，畴壁两侧电矩往往头尾相接，以

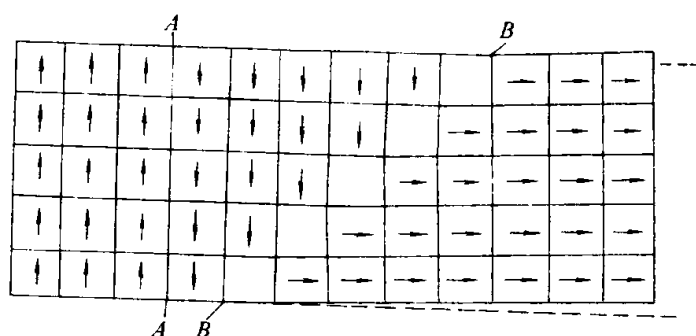


图 11-4 90°畴与 180°畴的结构示意^[2]

其中 A—A 为 180°畴界；B—B 为 90°畴界

便两侧电矩在畴壁法线方向的投影之和为 0，亦即使畴壁不带电，有时也有例外的情况，这时自发极化造成的畴壁电荷往往被自由电荷所中和。

铁电畴可以用偏光显微镜、透射电镜、扫描电镜、X 射线形貌法和化学腐蚀法进行观察，具体方法的选取由观察条件及材料自身的特征决定。图 11-5 为经腐蚀的铌酸锂晶体中电畴的光学显微照片。

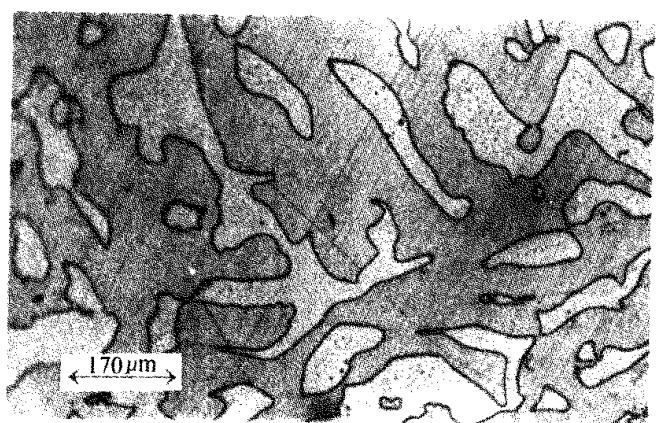


图 11-5 铌酸锂晶体中的电畴

(由朱永元提供)

铁电体的一个基本特征是其自发极化可以在外电场作用下转向。宏观极化方向的改变是通过极化方向与外场平行或含有与外电场平行的分量的新畴的成核和逐渐长大而取代旧畴实现的。本书第 18 章介绍的与相变有关的成核生长理论原则上也适用于畴成核和生长的过程。通常使晶体中电畴反转所需的外加电场是很大的，达到 kV/cm 量级。另外一种控制电畴方向的方法是先将材料加热至其居里温度以上，然后让它在在外电场中逐渐冷却至居里点以下，这样材料中产生的电畴将尽可能平行于外加电场。而这时所需的外场强度要大大低于使畴反转所需的电场强度。此外，通过在铁电薄膜生长过程中施加一个低电场可有效控制铁电薄膜极化的方向。这种方法已成功地用于

控制铁电薄膜和多层结构的极化方向^[9]。近年来人工调制畴结构的铁电晶体受到重视，利用周期或准周期排列的 LiNbO_3 片状 180° 畴结构可以实现准位相匹配，这在激光倍频和超高频电声转换方面有重要应用。关于这方面工作的介绍参见本书第 13 章。

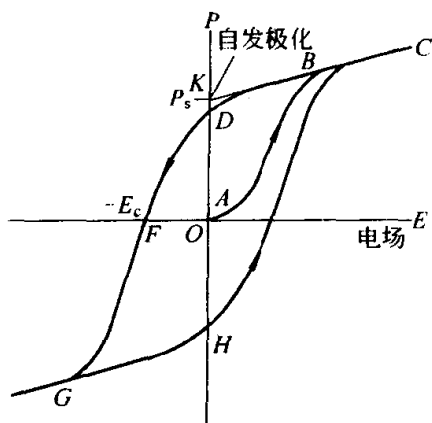


图 11-6 典型的铁电体电滞回线

铁电体表现的宏观电极化强度 P 与外加电场 E 密切相关。实际上 P 是铁电体中电畴的自发极化与感应极化的总和。由于电畴在电场作用下转向的过程总是滞后于电场， P 与 E 呈现出看上去类似于磁滞回线的复杂的“电滞回线”关系。

考虑一多畴铁电单晶体，一般说来，外加电场前它不表现出宏观电极化。若沿其某一极化轴方向施加一外电场，其极化矢量 P 与外加电场 E 的关系如图 11-6 所示。在外加电场之初，外电场很小时， P 随 E 线性增加，如 OA 段所示。此时 P 的增加主要来自感应极化和“弹性”的可逆畴壁移动。当电场进一步增加，新畴成核与不可逆畴壁移动造成 P 随 E 的增加而迅速增加，如 AB 段所示。当电场增至 B 点所在处的数值，晶体已变成单畴，自发极化对 P 的贡献趋于饱和。在此之后 E 继续增加时 P 线性增加，这种增量完全来自感应极化的贡献，如 BC 段所示。此后，若所加电场强度减小，极化强度 P 将循 CBD 曲线减小。在电场降至 0 时，仍保留有相当大的电极化强度 OD ，称为剩余极化 P_r 。将 CB 外推与 P 轴相交于 K 点， OK 表征自发极化 P_s 。此后若电场反向强度增加，极化将逐渐减小。至 F 点，极化下降为 0。 OF 表征的电场强度称为矫顽场，它表示将一个单畴铁电晶体变成宏观极化为 0 的多畴晶体所需的外加电场强度。反向电场继续增大，极化反向并沿 FG 曲线变化直至饱和。电场在大于正负饱和值之间循环一周过程中，电极化强度与电场强度沿封闭曲线 $CBDFGHC$ 变化。这一曲线称为电滞回线，它是铁电体的标志。

除了铁电晶体外，铁电陶瓷是经常被使用的材料，特别是作为压电材料。铁电陶瓷是由铁电体料粉体经过成型、烧结、极化和后加工而制成的多晶聚集体。它们具有较高的强度和硬度及较小的各向异性。另一个重要优点是其制造成本比单晶体要低得多。经烧制而尚未进行极化的铁电陶瓷中各晶粒的取向基本上是无序的。其晶粒可以是单畴或多畴的，而各畴自发极化的方向除了受铁电体自身极轴方向的限制外并无择优取向，这时的铁电陶瓷并不表现出宏观电极化。所谓极化处理是将陶瓷置于场强大于该材料矫顽场的

电场中,使铁电陶瓷中的各电畴的自发极化方向转动并尽可能与所加电场的方向一致。经过极化处理的铁电陶瓷便有了宏观电极化,这类陶瓷作为铁电、压电、热释电材料受到广泛应用,例如 PZT[$\text{Pb}(\text{Zr}_{1-x}\text{Ti})\text{O}_3$]陶瓷、PLZT[$(\text{Pb}_{1-x}\text{La}_{2x/3})(\text{Zr}_{1-y}\text{Ti}_y)\text{O}_3$]陶瓷、PMN[$\text{Pb}(\text{Mg}_{2/3}\text{Nb}_{1/3})\text{O}_3$]陶瓷及钛酸钡(BaTiO_3)陶瓷等。

如前所述,铁电体的极化方向可以用外加电场来改变,并且改变了的极化方向可以在外电场去除之后长期保存,这是一种可以开关的记忆特性。从20世纪50年代起人们就一直努力利用它的这一特性来进行数字存储,因为这种存储器具有如下的优点:在电源切除后记忆不会消失即“不挥发性”;能耗比磁芯存储器低得多;抗辐照性和快速存取功能。显然,铁电存储是否具有生命力在很大程度上取决于它能否与现存的半导体微电子技术相兼容,即足够小的体积,足够低的工作电压和制备工艺与半导体工艺相兼容。但是,早期人们制备的铁电陶瓷薄片厚度极限为 $10\mu\text{m}$,在这样的厚度自发极化反转所需电压为40V,这大大高于半导体芯片的开关电压 $2.5\sim 5\text{V}$,而且陶瓷工艺与半导体工艺难于兼容,因而为半导体科技界所难于接受。这种状况一直持续了30年。但从1986年开始情况有了很大的变化,这主要受益于氧化物薄膜制备技术在这一时期所取得的长足进展,各种现代的薄膜沉积技术使人们可以制备高质量的厚度薄达25nm的铁电薄膜,其运行所需的电压降至1V以下。而溶胶-凝胶(sol-gel)技术和金属有机物沉积(MOD)制膜技术的发展使低成本大面积铁电薄膜的制造成为可能。

作为数值计算基础的布尔代数中的“1”和“0”两种状态可以用一个铁电薄膜电容器中自发极化“向上”和“向下”来表示,一组这样的电容器就可构成一个铁电不挥发存储器,例如一个1Mb的存储器要由一千行和一千列这种铁电电容器存储单元构成。为了克服各存储单元之间的交叉干扰,每个铁电电容器被一个门晶体管(passgate transistor)和其余的电容器分隔开。由于每个铁电存储单元中含有一个晶体管和一个铁电电容,它称为“1T+1C”结构,该结构的原理和结构剖面图见图11-7与图11-8。

存储器所存数据的读出是这样进行的,若欲读出某一存储器的状态,则按前面所述使该存储器极化相类似的方法向该存储器加一个激发电脉冲,若其电场与被读存储器中铁电电容器的极化方向相同,则极化方向不变。若相反,则电容器中铁电膜的极化方向反转,随之而来的束缚电荷的变化 dp/dt 产生一个附加的输出电讯号,通过对这一输出电信号的比较和鉴别,便可知道该存储单元中极化“向上”或是“向下”,亦即读出它是“0”或是“1”。这种存储器的读出是破坏性的(destructive),读完之后为了保持原来存储的信息必须进行复位(reset),由于这种铁电存储在“写入”和“读出”时要

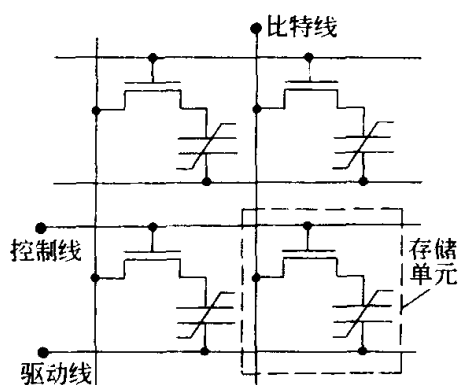


图 11-7 铁电存储器的
1T+1C 线路结构图

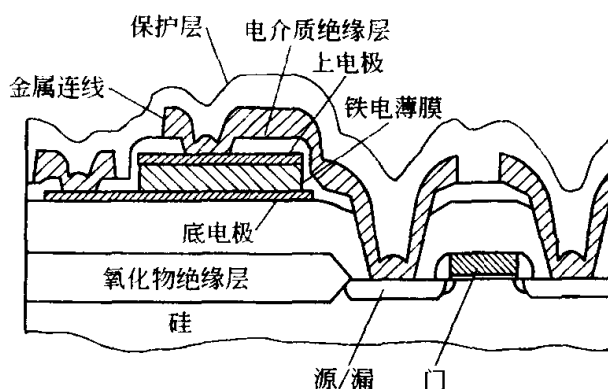


图 11-8 一个铁电存储器单元的剖面图，
显示集成在一道的一个铁电薄膜
电容器和一个 n-通道 MOS 晶体管

进行正、反向的交替极化，材料长期工作会出现疲劳和失效，因而材料的疲劳寿命是一个重要指标，一般要求达到 10^{12} 次。

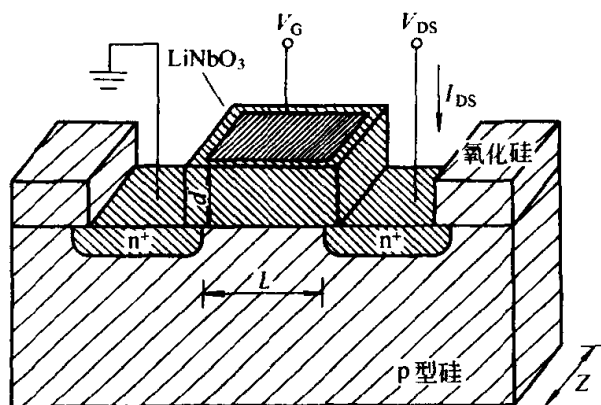


图 11-9 铁电场效应管的结构示意^[10]

为了从器件设计上克服作为疲劳原因的“破坏性读出”这一问题，近年来人们正在大力研究一种新型的非破坏性读出的铁电存储器件——铁电场效应管（FeFET），其示意图见图 11-9，它以可改变极化方向的铁电薄膜作为门电极，铁电薄膜“向上”或“向下”的极化场使 FET 开通或关断，这分别相应于存储单元的“0”或

“1”，测量 FET 的源和漏之间的电流便可知道存储单元的状态，这种存储器要在“写入”时改变铁电薄膜的极化方向，而读出时不必这样做，因而可在很大程度上解决“疲劳”问题，但 FeFET 中铁电薄膜要和半导体直接接触，界面反应和界面陷阱问题延缓了这类器件的发展，人们正在努力克服这一障碍，一种方法是采用氧化物半导体材料，第二种方法是在铁电薄膜门电极与半导体之间加一很薄的二氧化硅隔离层。

前面介绍的 1T+1C 不挥发铁电存储器用的铁电薄膜材料主要为 PZT 或 SBT。对 PZT 的研究已进行了多年，人们对其结构与性能已有较多的了解，它较大的剩余极化（ $30 \sim 40 \mu\text{C}/\text{cm}^2$ ），较好的电滞回线形状及通过调节成分改变其性能的便利使其成为一种广泛应用的材料，其主要问题是：第

一, PZT 与 Pt 电极的界面缺陷导致易于发生疲劳, 这点已通过采用导电氧化物如 LSCO 等电极在一定程度上得到克服。第二, PZT 中含有铅, 在制备和使用中可能会产生污染。SBT ($\text{SrBi}_2\text{Ta}_2\text{O}_9$) 的铁电性最早为前苏联约飞研究所的 Smolensky 所报道, 20 世纪 80 年代末 J.F.Scott 建议将 SBT 薄膜用于不挥发铁电存储器, 它具有以下 3 个明显的优点: 第一, 使用金属电极也几乎不发生疲劳, 据信这是由于其层状结构中的富氧 B_2O_3 层的作用; 第二, 其漏电流很小, 比 PZT 低两个数量级; 第三, 它可以被制成极薄的膜而不失去其大块材料的铁电性。它的缺点是剩余极化较小 ($6 \mu\text{C}/\text{cm}^2$), 而且自发极化是沿着薄膜不易发生择优取向的 a 轴方向, 但多晶膜的剩余极化对铁电存储来说倒也够了, 这些特点使 SBT 成为近年铁电薄膜研究的热点^[10]。

和其他存储器相比, 铁电不挥发存储器具有以下显著特点: 第一, 不挥发性, 其记忆基于自发极化的方向, 它不因外加电源的切断而消失; 第二, 快速存取, 达 10 ns 量级, 因为存储的速度取决于极化反转的速度, 这相当于弹性波通过薄膜厚度尺寸所需的时间; 第三, 低能耗, 因为其存储是电压驱动而不是电流驱动; 第四, 其工艺基本上可以和现存的半导体工艺相兼容, 尽管仍有要改进的地方。它还有抗辐射性, 但目前这个特点尚未派上用场。它的这些特点从近期看来使它可以用于需耗电低的智能卡、BP 机、手机等电子产品, 从远期来看, 它可能取代计算机特别是便携式计算机的磁盘从而实现低能耗、快速存取和大大减轻重量。

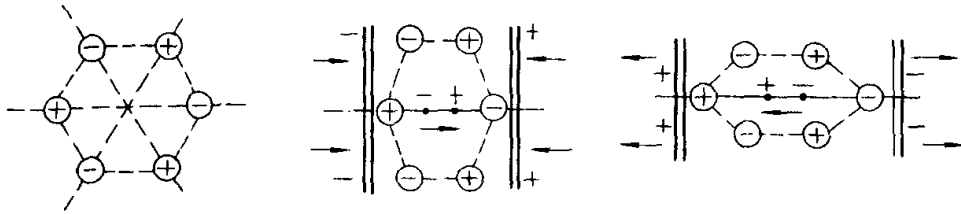
11.2 压电性与热释电性

11.2.1 压电性与压电材料

某些介质在机械力作用下发生电极化或电极化的变化, 这样的性质称为压电性。电极化的改变导致介质与极化方向垂直的两端面出现等量反号的束缚电荷变化, 看起来这是由于压力造成了电荷的变化, 这就是压电性一词的由来。具有压电性的介质称为压电体 (piezoelectrics)。上述过程的逆过程为: 在压电体的适当方向上施加外电场会导致压电体发生应变, 这种现象称为逆压电效应。

介质具有压电性的条件是其结构不具有对称中心, 在 32 类点群中有 20 类点群不具有对称中心, 属于这 20 类点群的电介质才可能是压电体。不具有对称中心的晶胞在应力作用下出现极化和表面束缚电荷的示意图 11-10。铁电体具有自发极化, 它们不仅结构上没有对称中心, 而且在无外力作用时晶胞的正、负电荷重心就不互相重合。在应力作用下, 其自发极化一般都要发生变化, 因而常见的铁电体都具有压电性。

为了对压电性进行理论分析, 我们从电介质的特征函数出发, 在系统的

图 11-10 压电晶胞受压、张应力的结构示意图^[1]

状态对其初始态偏离较小的前提下，将特征函数对独立变量展开只取到二次项，将有关热力学量对独立变量展开并只取到一次项，由此导出各种物性参数及其相互关系^[3]。

由于下面的计算中要涉及零阶至四阶张量，各种热力学量及系数的脚标较复杂，我们先对此作一些预备性的介绍。下面的计算中要涉及的热力学量按其阶数可分为 3 类：熵 S 和温度 T 为标量，即零阶张量；应力 X 和应变 x 为二阶张量；电位移 D 和电场 E 为矢量，即一阶张量。二阶张量的每个分量应有两个脚标，如应力分量为 X_{mn} ($m, n = 1, 2, 3$)，共有 9 个分量。但由弹性力学知识可知应力 X 与应变 x 均为对称张量，它们各自只有 6 个独立分量。为简单计，我们可以将双下标 mn 用一个下标 i 或 j 来表示，按约定俗成双下标与单下标有如下对应关系：

双下标	11	22	33	23, 32	31, 13	12, 21
单下标	1	2	3	4	5	6

这样， X_{mn} ($m, n = 1, 2, 3$) 便可记为 X_i ($i = 1, 2, 3, 4, 5, 6$)，利用这种表示方法，我们便可以用矩阵来表示三阶或四阶张量。注意，我们约定 m, n 的取值为 1, 2, 3。 i, j 的取值为 1, 2, 3, 4, 5, 6。

以温度 T 、应力张量 X 和电场强度矢量 E 为独立变量，相应热力学特征函数为吉布斯自由能 G ，按定义

$$dG = -SdT - x_i dX_i - D_m dE_m \quad (11-12)$$

考虑初始状态电场、应力均为 0，在偏离初始状态很小的情况下 $dE = E$ ， $dX = X$ 。 G 的展开式可记为

$$\begin{aligned}
 G = G_0 &+ \frac{\partial G}{\partial T} dT + \frac{\partial G}{\partial X_i} X_i + \frac{\partial G}{\partial E_m} E_m + \frac{1}{2} \frac{\partial^2 G}{\partial T^2} (dT)^2 + \frac{1}{2} \frac{\partial^2 G}{\partial X_i \partial X_j} X_i X_j \\
 &+ \frac{1}{2} \frac{\partial^2 G}{\partial E_m \partial E_n} E_m E_n + \frac{\partial^2 G}{\partial T \partial X_i} X_i dT + \frac{\partial^2 G}{\partial T \partial E_m} E_m dT + \frac{\partial^2 G}{\partial X_i \partial E_m} X_i E_m
 \end{aligned} \quad (11-13)$$

由式(11-12)有：

$$\frac{\partial G}{\partial T} = -S, \frac{\partial G}{\partial X_i} = -x_i, \frac{\partial G}{\partial E_m} = -D_m \quad (11-14)$$

由于 dS 、 x 、 D 均为独立变量 dT 、 X 和 E 的函数。在无应力、无电场附近将它们分别展开为：

$$\begin{aligned} x_i &= \left(\frac{\partial x_i}{\partial T} \right)_{X,E} dT + \left(\frac{\partial x_i}{\partial X_j} \right)_{T,E} X_j + \left(\frac{\partial x_i}{\partial E_m} \right)_{T,X} E_m \\ D_m &= \left(\frac{\partial D_m}{\partial T} \right)_{X,E} dT + \left(\frac{\partial D_m}{\partial X_i} \right)_{T,E} X_i + \left(\frac{\partial D_m}{\partial E_n} \right)_{T,X} E_n \\ dS &= \left(\frac{\partial S}{\partial T} \right)_{X,E} dT + \left(\frac{\partial S}{\partial X_i} \right)_{T,E} X_i + \left(\frac{\partial S}{\partial E_m} \right)_{T,X} E_m \end{aligned} \quad (11-15)$$

由式 (11-14) 和式 (11-15) 两式可求出

$$\begin{aligned} x_i &= - \left(\frac{\partial^2 G}{\partial X_i \partial T} \right)_E dT - \left(\frac{\partial^2 G}{\partial X_i \partial X_j} \right)_{T,E} X_j - \left(\frac{\partial^2 G}{\partial X_i \partial E_m} \right)_T E_m \\ D_m &= - \left(\frac{\partial^2 G}{\partial E_m \partial T} \right)_X dT - \left(\frac{\partial^2 G}{\partial E_m \partial X_i} \right)_T X_i - \left(\frac{\partial^2 G}{\partial E_m \partial E_n} \right)_{T,X} E_n \\ dS &= - \left(\frac{\partial^2 G}{\partial T^2} \right)_{X,E} dT - \left(\frac{\partial^2 G}{\partial T \partial X_i} \right)_E X_i - \left(\frac{\partial^2 G}{\partial T \partial E_m} \right)_X E_m \end{aligned} \quad (11-16)$$

利用式 (11-16) 及求导顺序可以填倒的原理，引入 6 个系数：

$$\begin{aligned} - \left(\frac{\partial^2 G}{\partial X_i \partial E_m} \right)_T &= - \left(\frac{\partial^2 G}{\partial E_m \partial X_i} \right)_T = \left(\frac{\partial x_i}{\partial E_m} \right)_{T,X} = \left(\frac{\partial D_m}{\partial X_i} \right)_{T,E} = d_{mi}^T \\ - \left(\frac{\partial^2 G}{\partial E_m \partial T} \right)_X &= - \left(\frac{\partial^2 G}{\partial T \partial E_m} \right)_X = \left(\frac{\partial D_m}{\partial T} \right)_{X,E} = \left(\frac{\partial S}{\partial E_m} \right)_{T,X} = p_m^X \\ - \left(\frac{\partial^2 G}{\partial X_i \partial T} \right)_E &= - \left(\frac{\partial^2 G}{\partial T \partial X_i} \right)_E = \left(\frac{\partial x_i}{\partial T} \right)_{X,E} = \left(\frac{\partial S}{\partial X_i} \right)_{T,E} = \alpha_i^E \\ - \left(\frac{\partial^2 G}{\partial X_i \partial X_j} \right)_{T,E} &= \left(\frac{\partial x_i}{\partial X_j} \right)_{T,E} = s_{ij}^{E,T} \\ - \left(\frac{\partial^2 G}{\partial E_m \partial E_n} \right)_{T,X} &= \left(\frac{\partial D_m}{\partial E_n} \right)_{T,X} = \epsilon_{mn}^{T,X} \\ - \left(\frac{\partial^2 G}{\partial T^2} \right)_{X,E} &= \left(\frac{\partial S}{\partial T} \right)_{X,E} = \frac{\rho^{E,X}}{T} \end{aligned} \quad (11-17)$$

于是式 (11-16) 可改写为

$$\begin{aligned} x_i &= \alpha_i^E dT + s_{ij}^{E,T} X_j + d_{mi}^T E_m \\ D_m &= p_m^X dT + d_{mi}^T X_i + \epsilon_{mn}^{T,X} E_n \\ dS &= \frac{\rho^{E,X}}{T} dT + \alpha_i^E X_i + p_m^X E_m \end{aligned} \quad (11-18)$$

这就是以温度、应力、外电场为自变量的处于电场中的弹性电介质的线性状

态方程。这 6 个系数是电介质的物性参量，它们代表了 6 种线性效应。这 6 个系数及它们所联系的物理量原则上都可用张量表示。将 l 阶张量和 m 阶张量联系起来的系数应为 $n = l + m$ 阶张量。应力 X 与应变 x 均为二阶张量，描述它们之间弹性效应的参数——弹性顺度 s 为四阶张量。电位移 D 与电场 E 均为一阶张量，描述它们之间关系的参数——介电系数 ϵ 为二阶张量。温度 T 为零阶张量，应变 x 为二阶张量，描述它们之间关系的参数——热膨胀系数为二阶张量。电位移 D 为一阶张量，应力 X 为二阶张量，描述它们之间关系的参数——压电系数 d 为三阶张量。温度为零阶张量，电位移 D 为一阶张量，描述它们之间关系的参数——热释电系数 p 为一阶张量。温度 T 与熵 S 均为零阶张量，描述它们之间关系的参数——比热容 c 为零阶张量。

利用式(11-17)前 3 个方程式中特征函数 G 的二次偏微商与求导次序无关的原理，可以得到如下关系：

$$\begin{aligned}\left(\frac{\partial x_i}{\partial E_m}\right)_{T,X} &= \left(\frac{\partial D_m}{\partial X_i}\right)_{T,E} \\ \left(\frac{\partial D_m}{\partial T}\right)_{X,E} &= \left(\frac{\partial S}{\partial E_m}\right)_{T,X} \\ \left(\frac{\partial x_i}{\partial T}\right)_{X,E} &= \left(\frac{\partial S}{\partial X_i}\right)_{T,E}\end{aligned}\quad (11-19)$$

它们表示描述正效应与逆效应的系数相等，例如式(11-19)中的第一个方程式表示压电系数与逆压电系数相等。第二个方程式表示热释电系数与电释热系数相等。压电系数 d_{mi} 为三阶张量，其 m 为 1, 2, 3，其 i 为 1, 2, 3, 4, 5, 6，它共有 18 个分量，其单位为 C/N 或 m/V。晶体的对称性对这些分量施加了更多的限制，这使实际独立分量数减少。由于自变量选取的不同和边界条件的差异，前面讨论的各个系数可以有不同的形式，有兴趣者可自行计算或查询专著^[2]。

压电材料包括压电单晶、压电陶瓷、压电聚合物和压电复合材料。压电单晶有很强的各向异性，测量晶体沿各晶轴的压电系数是一项复杂的工作。陶瓷是大量晶粒的聚集体，尽管单个晶粒表现出压电性，但只要它们在空间的分布是无序的，由于各个晶粒的效应互相抵消，总体上表现不出压电性。因此实际应用的压电陶瓷都是经过极化处理的铁电陶瓷。极化处理使铁电陶瓷保留沿极化电场方向的剩余极化，表现出单轴各向异性。应力造成剩余极化的改变就表现出压电性。

表征压电材料的物理量主要为压电系数、弹性系数、介电系数、热膨胀系数、硬度和密度。其主要功能参数为机电耦合系数 K ，其定义为

$$K^2 = \frac{\text{由逆压电效应转换的机械能}}{\text{储入的电能量}} \quad (11-20)$$

或

$$K^2 = \frac{\text{由压电效应转换的电能}}{\text{储入的机械能总量}}$$

此系数越大，材料的压电耦合效应越强。一些重要压电晶体材料和压电陶瓷材料及其主要性能参数列于表 11-3 和表 11-4。

表 11-3 几种重要压电材料的主要性能参数^[11]

材 料	点群	耦合系数 K (不同切割方式)/%	压电系数 d $/10^{-12} \text{C} \cdot \text{N}^{-1}$	相对介电常数 ϵ/ϵ_0	弹性系数 $s/10^{-12} \text{m}^2 \cdot \text{N}^{-1}$ $c/10^{11} \text{N} \cdot \text{m}^{-2}$	声速 v $/\text{m} \cdot \text{s}^{-1}$	密度 $\rho/10^3$ $\text{kg} \cdot \text{m}^{-3}$
水晶	32	10(X) 14(Y) 8.8(AT)	2.3 -4.6 -3.4	4.6 4.6 4.6	12.8(s_{11}^E) 9.6(s_{33}^E)	5700 3850 3320	2.65
罗息盐(30℃)	2	65(X-45°) 32(Y-45°)	275 30	350($\epsilon_{11}^T/\epsilon_0$) 9.4($\epsilon_{22}^T/\epsilon_0$)	52.0(s_{11}^E) 36.8(s_{22}^E)	3100 2340	1.77
(NH ₄)H ₂ PO ₄	$\bar{4}2m$	28(Z-45°)	24	15.3($\epsilon_{33}^T/\epsilon_0$)	18.1(s_{11}^E) 43.5(s_{33}^E)	3250	1.80
LiNbO ₃	3m	17(Z, 伸缩) 68(X, 切变)	6(d_{33}) 68(d_{15})	30($\epsilon_{33}^T/\epsilon_0$) 84($\epsilon_{11}^T/\epsilon_0$)	5.78(s_{11}^E) 5.02(s_{33}^E)	4160	4.7
Bi ₁₂ GeO ₂₀	23	15.5(111片, 伸缩) 23.5(110片, 切变)	$e_{14} =$ 1.14C·m ⁻²	38	1.28(c_{11}) 0.305(c_{12})	3340(纵波) 1680(表面波)	9.2
CdS	6mm	26(k_{33}) 19(k_{15})	10.3(d_{33}) -5.2(d_{31})	10.3($\epsilon_{33}^T/\epsilon_0$) 9.35($\epsilon_{11}^T/\epsilon_0$)	9.4(c_{33}^E) 9.07(c_{11}^E)	4500(伸缩) 1800(切变)	4.82
ZnO	6mm	41(k_{33}) 31(k_{15})	10.6(d_{33}) -5.0(d_{31})	11.0($\epsilon_{33}^T/\epsilon_0$) 9.26($\epsilon_{11}^T/\epsilon_0$)	21.1(c_{33}^E) 21.0(c_{11}^E)	6400(伸缩) 2945(切变)	5.68
AlN	6mm	30(k_{33})	5(d_{33})	9.3($\epsilon_{33}^T/\epsilon_0$)	34(c_{33}^E)	10400(伸缩)	3.26

表 11-4 几种重要压电陶瓷的性能参数^[11]

材 料	居里温度 $T_c/^\circ\text{C}$	密度 ρ $/10^3 \text{kg} \cdot \text{m}^{-3}$	耦合系数/%		相对 介电常数 $\epsilon_{33}^T/\epsilon_0$	介质损耗 $\tan \delta$ /%
			k_p	k_{31}		
BaTiO ₃	约 120	5.7	36	21	1700	1.0
PbTiO ₃	460~520	7.7	7~9.6	4.2~6.0	约 150	0.8~1.1
PbTiO ₃ -PbZrO ₃	180~350	7.5~7.6	25~65	15~39	460~3400	1.4~2.0
Pb(Mg _{1/3} Nb _{2/3})O ₃ -PbTiO ₃ -PbZrO ₃	170~350	7.6~7.8	30~76	22~43	550~9000	0.2~2.5
Pb(Co _{1/3} Nb _{2/3})O ₃ -PbTiO ₃ -PbZrO ₃	220~320	7.5~7.7	24~64	14~39	350~3900	
Na _{0.5} K _{0.5} NbO ₃ (热压的)	420	4.46	46	27	496	1.4
Pb _{0.6} Ba _{0.4} Nb ₂ O ₆	260	5.9	38	22	1500	1.0

续表

材 料	压电系数 $/10^{-12}\text{C}\cdot\text{N}^{-1}$		压电系数 $/10^{-3}\text{V}\cdot\text{mN}^{-1}$		弹性柔顺系数 $/10^{-12}\text{m}^2\cdot\text{N}^{-1}$	
	d_{33}	d_{31}	g_{33}	g_{31}	s_{33}^E	s_{11}^E
BaTiO ₃	190	-78	12.6	-5.2	9.5	9.1
PbTiO ₃	45~56	-4.2~-6.8	33	-3.2~-4.2	9.6	7.8
PbTiO ₃ -PbZrO ₃	71~590	-27~-274	17~40	-5.2~-16	9~20	9~16.5
Pb(Mg _{1/3} Nb _{2/3})O ₃ -PbTiO ₃ -PbZrO ₃	280~460	-79~-250	约 30	-6.5~-12.6		6.3~15.9
Pb(Co _{1/3} Nb _{2/3})O ₃ -PbTiO ₃ -PbZrO ₃						
Na _{0.5} K _{0.5} NbO ₃ (热压的)	127	-51	29.0	-11.6	10.1	8.2
Pb _{0.6} Ba _{0.4} Nb ₂ O ₆	220	-90	16.6	-6.8		11.5

压电材料是一种重要的功能材料,压电体在被发现后不太长的时间内就得到重要应用。两个公认的早期重要工作是:在1916年朗之万利用石英晶体制造了水声发射和接受装置,用于探测水中物体。这类装置即“声纳”,至今仍在海军中有重要应用。在1918年凯迪利用罗息盐晶体在机械谐振频率附近特异的电性能发明了谐振器,这为压电材料在通讯和频率控制方面的应用打下了基础。近年来压电材料在高科技中获得越来越广泛的应用,其主要应用领域参见表11-5。

表 11-5 压电材料的应用领域^[1]

应 用 领 域		举 例
电 源	压电变压器	雷达、电视显像管、阴极射线管、盖革计数器、激光管和电子复印机等高压电源和压电点火装置
信号源	标准信号源	振荡器、压电音叉、压电音片等用作精密仪器中的时间和频率标准信号源
信号转换	电声换声器	拾声器、送话器、受话器、扬声器、蜂鸣器等声频范围的电声器件
	超声换能器	超声切割、焊接、清洗、搅拌、乳化及超声显示等频率高于20 kHz的超声器件
发射与接收	超声换能器	探测地质构造、油井固实程度、无损探伤和测厚、催化反应、超声衍射、疾病诊断等各种工业用的超声器件
	水声换能器	水下导航定位、通信和探测的声纳、超声测深、鱼群探测和传声器等
信号处理	滤波器	通信广播中所用各种分立滤波器和复合滤波器,如彩电中频滤波器;雷达、自控和计算机系统所用带通滤波器脉冲滤波器等
	放大器	声表面波信号放大器以及振荡器、混频器、衰减器、隔离器等
	表面波导	声表面波传输线
传感与测量	加速度计压力计	工业和航空技术上测定振动体或飞行器工作状态的加速度计、自动控制开关、污染检测用振动计以及流速计、流量计和液面计等
	角速度计	测量物体角速度及控制飞行器航向的压电陀螺
	位移发生器	激光稳频补偿元件、显微加工设备及光角度、光程长的控制器
其他	非线性元件	压电继电器等

在压电材料的设计与应用方面特别值得一提的是我国学者发展的“声学超晶格”理论与材料^[12]。利用晶体生长技术或薄膜技术可以获得如图 11-11 所示, 具有周期性排列畴结构的晶体。其中各畴自发极化与界面垂直而且呈“头对头、尾对尾”排列。



图 11-11 压电超晶格示意,
箭头表示极化方向



图 11-12 畴界处压电系数
不连续, 可看作 δ 声源

声学超晶格可以看作一系列 δ 声源按周期或准周期排列而成。这些 δ 声源产生于压电应力不连续处, 而压电应力为压电系数与电场强度的乘积。这里考虑的为一铁电晶体, 即铌酸锂聚片多畴晶体。如上所述, 所有奇数阶张量在畴界处变号, 而压电系数为三阶张量, 因而畴界是压电系数的不连续处。在交变外电场作用下, 畴界处压电应力的不连续性将产生相应的应变 $S(u_m)$ 这里 u_m 是畴界的位置。该应变将以弹性波的方式传播, 可表示为

$$S(u) = S(u_m) \cos(\omega t - ku) \quad (11-21)$$

因而畴界可以看作是 δ 声源。假设片畴沿 z 方向排列 (参看图 11-11 和图 11-12), 若 a 和 b 分别为正、负片畴的厚度, $a + b$ 为超晶格的周期, N 为超晶格中的周期数, $N(a + b)$ 为晶片的总厚度, 如果 $\{0001\}$ 面为电极面, 平行于超晶格中的畴界, 那么在交变电场下, 它将激发出沿 z 轴传播的平面纵波。这种情况下可用下列波动方程描述:

$$\frac{\partial^2 u_3}{\partial z^2} - \frac{1}{V^2} \frac{\partial^2 u_3}{\partial t^2} = \frac{2h_{33}D_{33}}{c_{33}} \sum_m (-1)^m \delta(Z - Z_m) \quad (11-22)$$

这里 u_3 代表粒子沿 z 方向的位移, V 为声速, h_{33} 和 c_{33} 分别为压电和弹性系数, D_{33} 为电位移矢沿 z 方向的分量。

利用格林函数方法求解周期声学超晶格的弹性波波动方程式 (11-22), 可得谐振器电抗, 从而得到谐振频率:

$$f_n = \frac{n \cdot v}{a + b} \quad n = 1, 2, 3, \dots \quad (11-23)$$

其中 v 是沿 z 方向传播的纵波波速。值得注意的是, 谐振频率只决定于声学超晶格的周期 $a + b$, 而与晶片的总厚度无关。如我们所知, 如果用通常材料如单畴 LiNbO_3 晶体来制备工作频率为几百兆周的谐振器, 其晶片的总厚度太薄, 难以加工。然而生长周期为数微米的声学超晶格晶体是比较容易

的。因而利用声学超晶格可以制备工作频率在数百 MHz 到数千 MHz 的声学器件。

与逆压电效应有些类似的另一现象是电致伸缩效应。电致伸缩也是由电场引起应变的效应。但与逆压电效应不同的是：逆压电效应是线性效应，只发生在没有对称中心的点群中，而电致伸缩是非线性效应，应变与外场的关系为：

$$x = QE^2 \quad (11-24)$$

其中 Q 为电致伸缩系数，它是一个四阶张量。而且电致伸缩可以发生在各种电介质中，无论它们的点群有无对称中心，也无论它们是固体、液体或气体。电致伸缩效应一般很弱，长期未有重要应用。但近年来人们发现某些高介电常数材料和稍高于居里点温区的铁电材料具有特别大的电致伸缩，同时具有很好的重复性，温度稳定性和很短的响应时间，例如 $\text{Pb}(\text{Mg}_{1/2}\text{Nb}_{1/2})\text{O}_3$ ，(PMN) 就是一个例子，这类材料已开始实用化。

11.2.2 热释电性与热释电材料

热释电性 (pyroelectricity) 指的是某些电介质的电极化随温度改变的性质。这些电介质称为热释电体。实际上热释电体是具有自发极化的电介质。如本章第 1 节所述，32 种点群有 10 种点群可出现自发极化，属于这些点群的电介质可具有热释电性。一般说来，铁电体都具有热释电性。考虑一个单畴的铁电晶体或经过极化处理的铁电陶瓷，与极化矢量垂直的相对的两个表面附近分别出现正、负束缚电荷。通常，这些束缚电荷被等量而符号相反的自由电荷所屏蔽，总体表现出电中性。当温度改变时，电极化强度发生变化，原先的自由电荷不再正好可屏蔽束缚电荷，于是表面上就表现出电荷的存在。如果这时在相对两表面装上电极，在连接两电极的导线上就会有电流流过。

材料的热释电性质用热释电性系数 p 来表示，若热释电晶体的温度改变了 ΔT ，则其极化强度的改变 ΔP 可由下式给出：

$$\Delta P = p \Delta T \quad (11-25)$$

由于 P 为矢量， T 为标量，热释电系数 p 应为一矢量，其分量为：

$$p_m = \frac{\partial p_m}{\partial T} \quad m = 1, 2, 3 \quad (11-26)$$

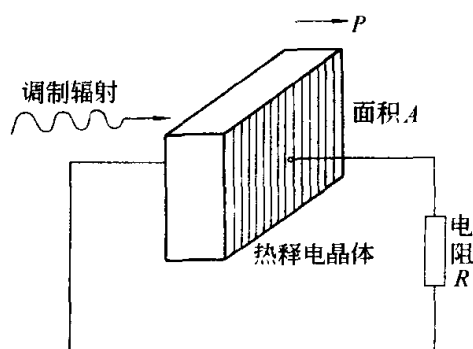
热释电系数的单位为 $\text{C}/(\text{m}^2 \cdot \text{K})$ 。

热释电材料的最重要应用是热释电传感器和红外成像焦平面。热释电传感器是把热，特别是光辐射产生的热转变为电信号的装置。其结构示意图如图 11-13 所示。为了增大效率，两相对的镀在热释电材料表面上的金属膜电极应尽可能与其自发极化 P 垂直。用调制频率为 f 的红外辐射照射热释电

体, 其温度、自发极化和表面束缚电荷都会随之变化。这样图 11-13 所示的负载电阻上便会产生一频率为 f 的电信号。如果红外辐射的强度发生变化, 那么电阻 R 上产生的电压降的幅度就发生相应的变化, 其数值 ΔV 满足:

$$\Delta V = AR \frac{dP}{dt} = AR \frac{\partial P}{\partial T} \frac{dT}{dt} = AR\rho \frac{dT}{dt}$$

(11-27) 图 11-13 热释电探测器原理示意^[1]



其中 A 为探测器面积, R 为负载电阻,

ρ 为热释电系数, T 为温度。利用此原理便可测量温度的变化。热释电传感器与其他红外传感器相比具有以下优点: 波长响应范围宽, 可在室温环境下使用, 响应快, 造价低廉。对低频小面积的探测器通常用 ρ/c_p 作为表征热释电材料的品质因子, 其中 ρ 为热释电系数, c_p 为定压比热容。显然 ρ 较大意味着较高的面束缚电荷变化率, c_p 较小的材料可在吸收热量一定的条件下产生较大的温度变化, 从而输出较大的电流和电压。几种最常用的热释电探测器材料及其室温特性列于表 11-6。

表 11-6 主要热释电探测器材料及其性质, 沿最佳方向^[13]

材 料	ρ $/(\text{nC} \cdot \text{cm}^{-2} \cdot \text{K}^{-1})$	ϵ'/ϵ_0	c_p $/(\text{J} \cdot \text{cm}^{-3} \cdot \text{K}^{-1})$	ρ/c_p $/(\text{nA} \cdot \text{cm} \cdot \text{W}^{-1})$	$\rho/(c_p \epsilon)$ $/(\text{V} \cdot \text{cm}^{-2} \cdot \text{J}^{-1})$
TGS	30	50	1.7	17.8	4000
LiTaO ₃	19	46	3.19	6.0	1470
Sr _{1/2} Ba _{1/2} Nb ₂ O ₆	60	400	2.34	25.6	720
PLZT(6/80/20)	76	1000	2.57	29.9	340
PVF ₂	3	11	2.4	1.3	1290

热释电材料器件应用的最新发展是用于红外成像系统, 即“夜视”装置。各种物体即使在黑暗的环境中也会随其温度的变化发射具有不同强度和波长的红外线。利用分立的多行多列小块热释电单元作为红外摄像机的焦平面, 每一个单元可作为一个像素。当目标发射的红外线成像于该焦平面时, 各分立单元便按其被照射强度的不同而各自产生不同强度的电信号, 这些电信号被放大和处理后可在荧光屏上还原出目标的可视图像。图 11-14(a)、(b) 分别为热释电红外成像焦平面的局部放大图和用此类器件在黑夜中拍摄的红外图像。热释电红外成像系统的最大优点是不需要低温条件, 可在室温下工作, 因而可大大降低重量和成本, 可以制成用于轻武器的夜间瞄准器等夜视器件。

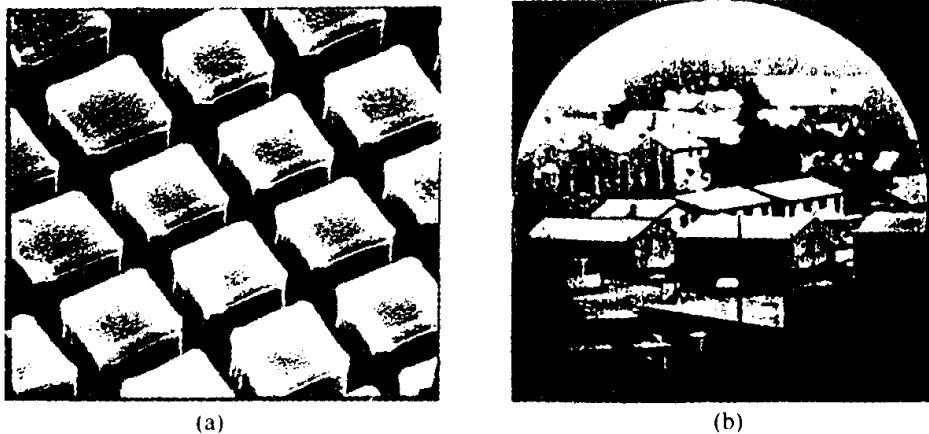


图 11-14

(a) 热释电红外成像焦平面的局部放大，每一像素的面积为 $19\mu\text{m} \times 16\mu\text{m}$ ；

(b) 用热释电红外成像系统摄取的红外光图像^[13]

参 考 文 献

- 1 师昌绪等主编，材料科学技术百科全书，北京：中国大百科全书出版社，1995，p.230. 条目作者：周志刚
- 2 方俊鑫，殷之文，电介质物理学，北京：科学出版社，1998
- 3 钟维烈，铁电体物理学，北京：科学出版社，1996
- 4 冯端等，金属物理学，第二卷，相变，北京：科学出版社，1990
- 5 Remeika J.P.Glass A M.Mat.Res.Bull.1970, 5: 37
- 6 Cochran W.Adv.Phys.1960, 9: 387
- 7 Cochran W.Adv.Phys.1961, 10: 401
- 8 Born M, Huang K.The Dynamical Theory of Crystal Lattices.Oxford: Clarendon Press, 1954
- 9 Hu W S, Liu Z.G.and Feng D.J.Appl.Phys.1996, 80: 7089
- 10 Scott J.F.Ferroelectrics Review.1998, 1: 1
- 11 许煜寰，铁电与压电材料，北京：科学出版社，1978
- 12 闵乃本，微米超晶格的概念、效应和应用，物理学进展，1993, 13: 26
- 13 Kenji Uchino, 铁电陶瓷，p.547；材料科学与技术丛书，第十一卷，陶瓷的结构与性质，W.V.Swain, 主编，郭景坤等译，北京：科学出版社，1998

第 12 章 超导电性质

撰稿人 丁世英

12.1 引言

1911 年荷兰低温物理学家 H.K. Onnes 发现有些物质从特定的温度开始会转变为完全没有电阻的状态，这就是所谓的超导（电）现象。具有这种性质的材料称为超导材料，这个与材料有关的特定的温度叫转变温度 T_c 。零电阻是超导电性的一个基本特征。

可是，在发现超导电性后的很长一段时间内，物理学家们完全不能理解物质在超导态的电阻为什么是零。1931 年这种情况有了改变，人们发现锡在转变到超导态时比热容有跳变，这是正常-超导转变，是相变的第一个证据。这样，对超导电性的理解就前进了一大步。同年，W. Meissner（迈斯纳）和 R. Ochsenfeld 发现，处在超导态的物体完全排斥磁场，即磁力线不能进入超导体内部，这一特征叫完全抗磁性，通常也叫做迈斯纳效应，是超导态的另一个基本特征。迈斯纳效应与施加磁场和冷却样品的前后次序（历史）无关，说明超导-正常转变是可逆过程，它的发现大大促进了超导研究。1935 年，超导研究工作因 H. London 和 F. London 的工作而受到进一步推动，但后来因为二次世界大战缘故而延缓，直到 1950 年 V.L. Ginzburg, L.D. Landau 理论建立以后才又迅速发展起来。

1950 年 H. Fröhlich（弗罗列希）提出了电子间通过声子相互作用作为超导电性的微观机制。1956 年，Cooper 提出超导态中的电子是两两束缚在一起的，而不是普通导体中的自由电子。1957 年，J. Bardeen, L.N. Cooper, J.R. Schrieffer 发表了现在叫 BCS 理论的超导微观理论，第一个成功地解释了超导体的性质。

1962 年 B.D. Josephson 理论上预言了超导体隧道效应，他的预言不久就被证实了。超导隧道效应现在常叫做约瑟夫森效应。在约瑟夫森效应基础上的超导量子干涉器件叫 SQUID（超导量子干涉器件），是超导体在电子学方面应用的核心部件。

自发现超导现象起到 1986 年止的几十年中，虽已发现了大量的超导体，但其转变温度均不超过 23.2K。这意味着超导现象必须用液氮冷却才会出现。液氮是一种温度最低的液体，在 0.1MPa 时沸点为 4.2K。然而氮是稀

有元素，价格贵，液化难，冷却效率低，这使超导现象的实现很昂贵，大大地限制了超导材料的广泛应用。

1986年 G. Bednorz 和 K.A. Müller 发现某些稀土元素的氧化物，也就是陶瓷材料，是超导的，其转变温度在 30K 以上，叫做高温超导体。高温超导材料的发现引起了全球性的高温超导热。随后相继发现的一系列的高温超导体，其转变温度最高达到 135K，可以用价廉的液氮（沸点在 0.1MPa 下约为 77.3K）来实现冷却。这就使超导材料的大规模应用出现了新前景，从而大大推进了高温超导的研究。当然，高温超导材料目前离广泛应用还有一定距离，比如，在制备性能优良的、价格有市场竞争力的实用超导材料方面的一些问题还有待解决。但是，转变温度的大幅度提高向超导材料的广泛应用迈出了极其重要的一步。

我国在低温超导材料的研制上有较高的水平。1986 年以来，我国的高温超导研究也一直处于国际前列。譬如我国是最早发现液氮温区超导体的国家之一，超导材料实用化的主要参数之一的临界电流密度值现在也处于世界先进水平。

12.2 超导电性的基本特征

12.2.1 零电阻效应

物质分为导体、半导体和绝缘体。常见的良导体是各种金属（如金、银、铜、铝），它们的电阻率非常低。常见的半导体是一些能隙甚小的绝缘体（如硅和锗）和某些化合物（如砷化镓，锑化钢）。验证金属导电理论的实验所需温度越低越好。金属电子论认为两方面的原因产生了电阻。一是原子的热运动（声子）对电子的散射，这种电阻随温度的下降而降低： $T \gg \Theta_D$ ，电阻随温度线性降低； $T \ll \Theta_D$ ，电阻随温度的 5 次幂下降， Θ_D 是德拜温度。二是晶体缺陷和杂质原子对电子的散射，它与温度无关。在高温时以声子贡献为主。在低温时，不纯金属以杂质贡献为主，称为剩余电阻，很纯的金属才能看到声子电阻。因此，在要验证低温下金属电阻与温度的关系要求金属愈纯愈好（参见图 12-1）。

1908 年，Onnes 将氮液化，开始进行验证经典金属电子论的实验。他能得到的最纯的金属是水银。1911 年，Onnes 发现，水银的电阻在温度降到 4K 附近时突然减小到测不出来，见图 12-2。他反复地进行测量，结果证明这个现象是真的。他称水银进入了零电阻的超导(电)态。后来，物理学家用最精确的方法也测不出超导态有任何电阻，确认了零电阻是任何超导体的基本特征之一，称为“零电阻效应”。对水银，4.2K 是其超导转变温度，用 T_c 表示， T_c 也叫临界温度。当温度高于 T_c 时，材料处于有电阻的正常态。

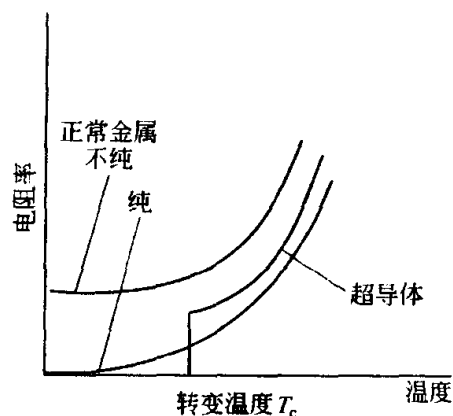


图 12-1 不同导体的电阻-温度曲线

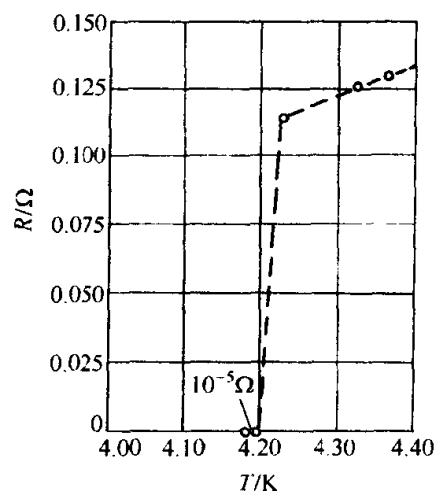


图 12-2 水银的零电阻效应

此后，科学家们对所发现的超导电性的理解进展缓慢。尽管如此，还是相继发现了大量的超导体。1931 年发现了合金超导体。现在已经证明，包括大部分金属元素，各种合金和化合物，包括氧化物和有机物都是超导体。它们的转变温度各不相同，低的小于 1 K，高的在 100 K 以上。图 12-3 是元素超导体在周期表中的位置。从表上可以看到，铁磁性金属和碱金属都不是超导体。在 1986 年以前，金属化合物 Nb_3Ge 薄膜的 T_c 最高，为 23.2K，这些材料都被称为低温超导体。

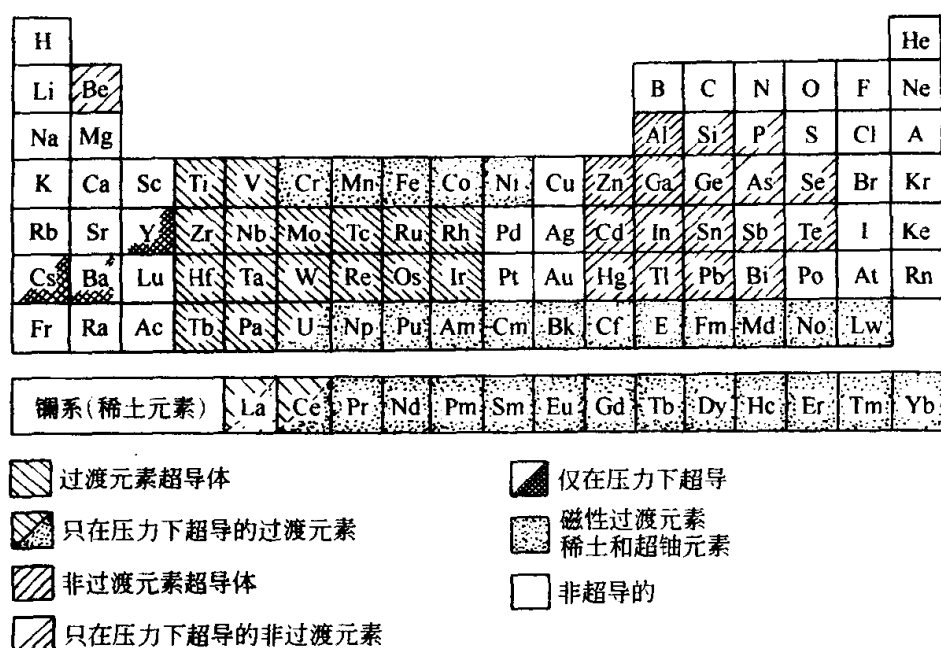


图 12-3 周期表中的超导体

12.2.2 迈斯纳效应

除了零电阻外，处于超导态的物体还完全排斥磁场，即磁力线不能进入超导体内部，这一特征叫完全抗磁性或迈斯纳效应。这是超导态的第二个基本特征[图 12-4(a)]。图 12-4(b)是迈斯纳效应的结果，也是各种超导磁悬浮的物理基础。图上半的圆盘是个悬浮的永久磁体环，下半是放在容器中的超导体。当容器中注入制冷剂以前，永久磁体放在超导体上 [说明在冷却前已经向超导体施加了磁场，参见图 12-4 (a)]，注入制冷剂以后超导体被冷却到超导态，出现抗磁性，它和永久磁体间的排斥力使永久磁体悬浮。当制冷剂蒸发完后永久磁体又落在超导体上。

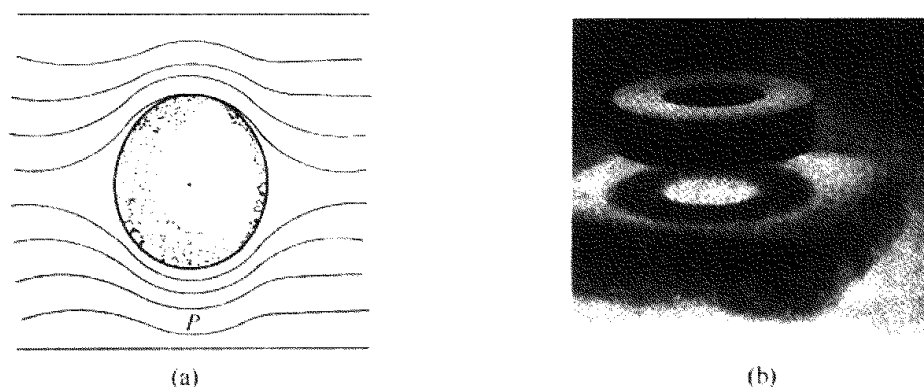


图 12-4 迈斯纳效应

(a) 超导球排斥磁通；(b) 超导体排斥力使永久磁环悬浮

非常重要的是，这一现象与怎样实现它的过程无关。也就是说，将上述先加磁场后冷却的次序改变一下：先将容器中注入制冷剂，使超导体冷却到超导态再将永磁体放在超导体上 [先冷却后加磁场，参见图 12-4 (b)]，超导体同样出现抗磁性，永磁体也悬浮起来。因而，超导态是一个热力学平衡现象，抗磁性是可逆性质，可以用平衡热力学来处理。

完全抗磁性不是理想导体的特征。假设有一个理想导体，在转变温度 T_c 以下电阻是零，把它放在磁场中冷到 T_c 以下时，它没有抗磁性。而把它先冷却到 T_c 以下后再施加磁场，它就具有抗磁性，参见图 12-5 (a)。这说明，理想导体的这种磁性质与加场过程的历史有关，不是可逆性质。理想导体的这种特性是遵从电磁感应定律的结果。

因此，零电阻效应和迈斯纳效应是超导体的两个基本特征。在人们确定一块样品是不是超导体及其临界温度时，不仅需要看有没有像图 12-2 那样的电阻-温度 ($R-T$) 曲线，还一定要看磁化率-温度 ($\chi-T$) 曲线，或等价的磁化强度-温度 ($M-T$) 曲线。图 12-6 是 2001 年发现的金属化合物二硼化镁 MgB_2 的磁化率-温度 ($\chi-T$) 曲线。其临界温度在外场为零时接近 40 K，是目前 T_c 最高的低温超导体。

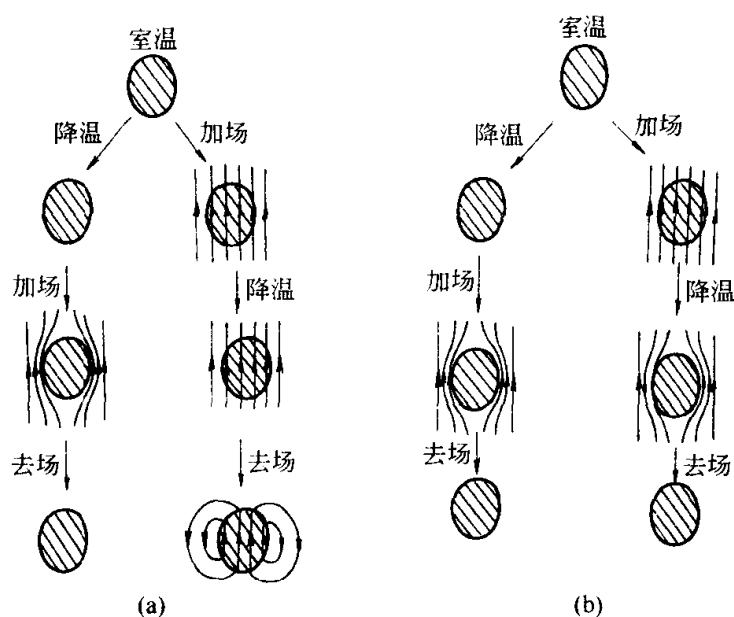


图 12-5 理想导体的不可逆 (a) 与超导体的可逆的 (b) 磁性质

(a) 理想导体是否有抗磁性取决于加磁场与冷却的次序；

(b) 超导体的抗磁性与加磁场和冷却的次序无关

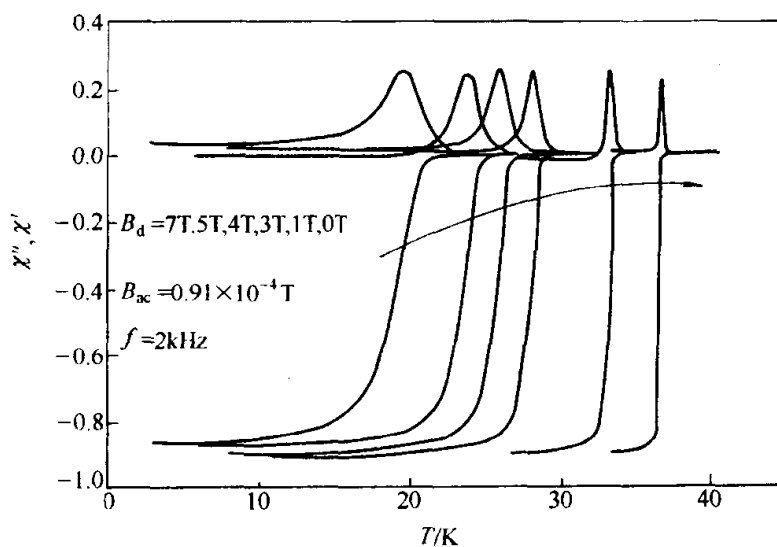


图 12-6 MgB_2 的 χ - T 曲线

χ' 是实部反映抗磁性, χ'' 是虚部, 反映交流损耗

12.2.3 超导体临界参数

进一步研究发现, 温度低于 T_c 时, 强磁场也会破坏超导态: 当磁场值超过某一个临界值 H_c 时, 材料就从超导态转变为正常态, H_c 称为临界磁场。不同超导体的 H_c 不同, 并且是温度的函数。实验发现, 临界磁场和温度的关系可以近似地表示成

$$H_c(T) = H_c(0) \left[1 - \left(\frac{T}{T_c} \right)^2 \right] \quad (12-1)$$

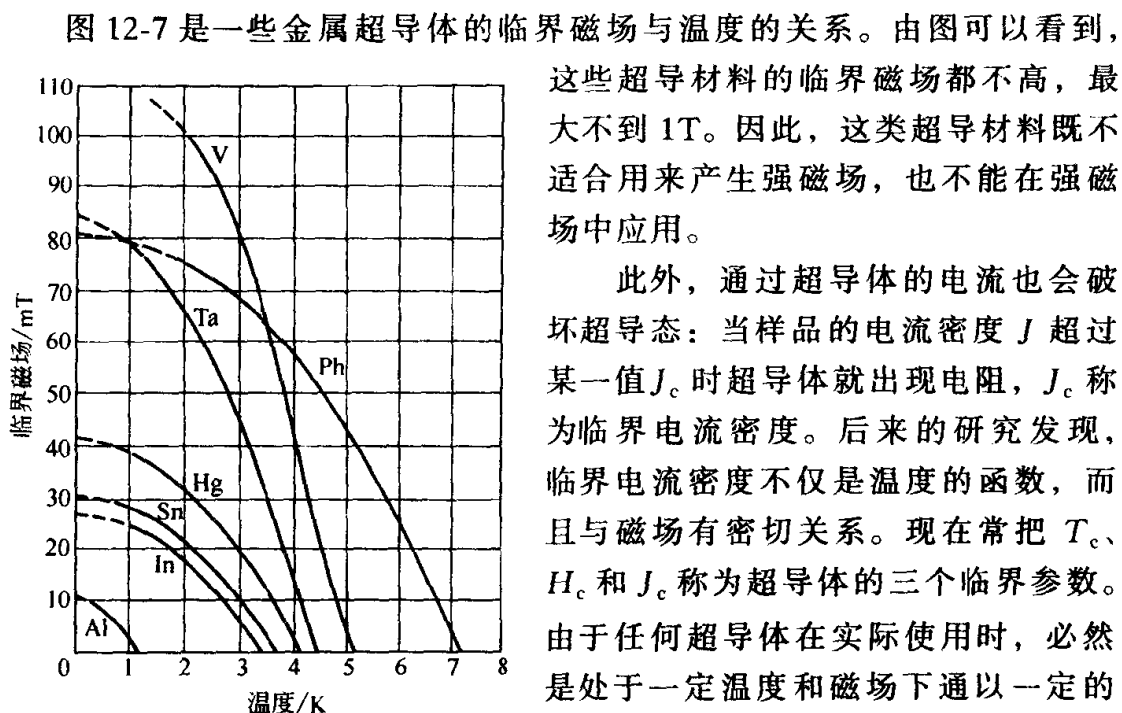


图 12-7 一些金属超导体临界磁场与温度的关系

这些超导材料的临界磁场都不高，最大不到 1T。因此，这类超导材料既不适合用来产生强磁场，也不能在强磁场中应用。

此外，通过超导体的电流也会破坏超导态：当样品的电流密度 J 超过某一值 J_c 时超导体就出现电阻， J_c 称为临界电流密度。后来的研究发现，临界电流密度不仅是温度的函数，而且与磁场有密切关系。现在常把 T_c 、 H_c 和 J_c 称为超导体的三个临界参数。由于任何超导体在实际使用时，必然是处于一定温度和磁场下通以一定的电流（样品有一定的电流密度），而温度、磁场和电流密度一定要低于这三个临界参数。这是维持超导状态的必要

条件。由于早期超导材料的三个临界参数都很低，因此不具备广泛实际应用的条件，所以超导材料的发展要求发现高临界参数超导体。

12.3 超导体的热力学性质

12.3.1 自由能

平衡热力学告诉我们，系统在平衡态时自由能最低。当外参量是温度 T 、磁场强度 H 和压强 p 时，态函数是吉布斯自由能。单位体积吉布斯自由能密度增量 dG 为：

$$dG = -SdT + Vdp - \mu_0 M dH \quad (12-2)$$

式中， M 、 S 和 V 分别是系统的磁化强度、熵和体积。上式也可写成 $dG = -SdT + Vdp - BdH$ ，多出了一项 $\mu_0 HdH$ ，相当于计及真空的磁场能量。这两种写法没有实质性差别。这里采用式 (12-2)。

迈斯纳效应告诉我们超导体是一个平衡热力学系统，其状态方程是

$$M_s = -H \quad (12-3)$$

超导体在正常态的磁性可以忽略，也是平衡态，其状态方程是

$$M_n = 0 \quad (12-4)$$

M_s 和 M_n 分别是超导态和正常态的磁化强度。超导体的相图即图 12-7，两态的相界曲线就是式(12-1)。

由式(12-1)、式(12-3)和式(12-4)完全可以导出超导体的热力学性质，这就是所谓的磁热效应。由于超导体的磁性质在正常态和超导态的差别很大，再加上磁测量实验比较容易，精确度也高，它已成为超导体相变研究的重要手段。

我们首先看磁场从 0 升到 H 而 T 、 p 保持不变 ($dT=0$, $dp=0$) 时迈斯纳态自由能的增加。将 $M_s = -H$ 代入自由能后得到 $dG_s = -\mu_0 M_s dH$ 。于是，恒温下外场从 0 增加到任意值 H 时超导态自由能的增量是

$$G_s(T, H_c) - G_s(T, 0) = \mu_0 H^2 / 2 \quad (12-5)$$

恒温下外场从 0 增加到临界值 H_c 超导态的自由能的增量就是 (见图 12-8)

$$G_s(T, H_c) - G_s(T, 0) = \mu_0 H_c^2 / 2 \quad (12-6)$$

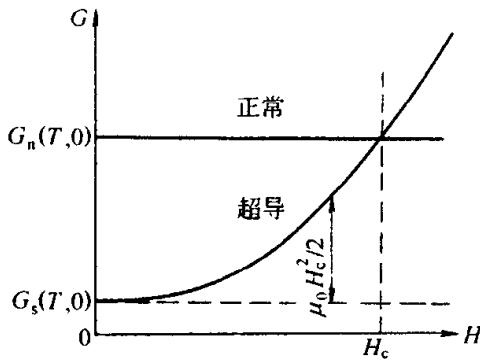


图 12-8 体的自由能

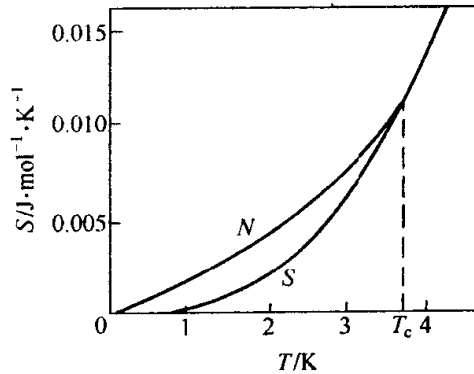


图 12-9 锡在正常态和超导态的熵

利用正常态状态方程 $M_n = 0$ 就得到 $dG_n = 0$ (见图 12-8)，即

$$G_n(T, H_c) - G_n(T, 0) = 0 \quad (12-7)$$

此等式表示这样一个事实：因为正常态没有磁性，外磁场变化时不对系统做功，超导体正常态自由能不随外磁场变化。在图 12-8 中， G_n 与磁场的关系是一条水平线。联合式(12-6)和式(12-7)就得到

$$G_s(T, 0) - G_n(T, 0) = -\mu_0 H_c^2 / 2 \quad (12-8)$$

此式的物理意义是，在 $T < T_c$ 和 $H = 0$ 时，超导体的自由能密度比正常态低 $\mu_0 H_c(T)^2 / 2$ ，所以超导态是低能态。

应当注意，虽然我们是从通过用磁场做功的方式得到式(12-8)的，但超导态的凝聚能 $\mu_0 H_c^2 / 2$ 是超导态的特性，与磁场是否存在无关。

12.3.2 熵

考虑相变过程中熵的变化。利用热力学关系式和式(12-8)可得熵差为

$$S_n - S_s = - \left[\frac{\partial (G_n - G_s)}{\partial T} \right]_{H,p} = - \mu_0 H_c \left[\frac{\partial H_c}{\partial T} \right]_{H,p} \quad (12-9)$$

由相界曲线 $H_c(T) = H_c(0)[1 - (T/T_c)^2]$ 知 $\frac{\partial H_c}{\partial T} < 0$, 故 $S_n > S_s$ 。图 12-9 是实验测定的锡在正常态和超导态的熵, 说明超导态比正常态更有序。因此超导态不仅是低能态, 而且是有序态, 由正常态转变为超导态是一种由无序的高能态向有序的低能态的“凝聚”过程。所以超导态是一种凝聚态, 通常把 $\mu_0 H_c^2/2$ 叫超导凝聚能。超导物理学是凝聚态物理学的一个重要分支。

那么, 超导态是什么方面比正常态更有序了呢? 晶体结构分析表明, 超导态的晶体结构和正常态的没有什么明显可察觉的差别。德拜温度的测量也没有发现转变前后有明显的变化, 说明晶格振动在两相基本相同。因此人们猜测, 从正常态凝聚到超导态时, 电子的有序度发生了变化。

12.3.3 潜热

按潜热 L 的定义和式(12-10)我们有

$$L = T(S_n - S_s) = - \mu_0 T H_c \left[\frac{\partial H_c}{\partial T} \right]_{H,p} \quad (12-10)$$

除临界点 ($H_c=0$, $T=T_c$) 和 $T=0K$ 外 $\frac{dH_c}{dT} < 0$, 所以由正常态向超导态的转变(凝聚)伴有潜热放出, 是一级相变。但是, 无磁场存在的凝聚是高级相变, 因为 $H_c(T_c)=0$, $S_n = S_s$, $L=0$ 。至于此时相变究竟是几级, 仅仅根据潜热是不能确定的。

在磁场中的 $N-S$ 相变有潜热是超导态存在能隙的一个证据。人们进一步猜想, 超导体由超导态转变到正常态时, 电子从能隙上面的能带凝聚到能隙下面的能带中去了。在温度升高时凝聚能密度和潜热降低, 能隙随之而逐步减小, 到临界点 ($T=T_c$, $H_c=0$) 时能隙减小到零。这样的猜想是完全合理的。

12.3.4 比热容

为判别无磁场中相变的级数, 可进一步计算两相的比热容。按比热容定

义, $c_n - c_s = T \left[\frac{\partial (S_n - S_s)}{\partial T} \right]_{H,p}$ 。将熵代入上式后得到

$$c_n - c_s = - \mu_0 T \left[\left(\frac{\partial H_c}{\partial T} \right)^2 + H_c \frac{\partial^2 H_c}{\partial T^2} \right] \quad (12-11)$$

注意 $H_c(T_c)=0$, 在零磁场下我们有

$$c_n - c_s = -\mu_0 T_c \left(\frac{\partial H_c}{\partial T} \right)^2 T_c \quad (12-12)$$

这里 $T = T_c$ 是无磁场相变条件。而 $c_n < c_s$ 表示比热容有跃变，所以是二级相变。图 12-10 是锡在两种状态的比热容与温度的实验关系，其中 $T < T_c$ 的正常态比热容可将样品放在高于 H_c 的磁场中用量热法测得，也可以利用磁测量法得到的材料方程 $H_c(T)$ 得出。

12.3.5 超导能隙

正常金属的低温比热容有两部分组成：点阵部分贡献的 $c_{nl} = A(T/\Theta_D)^3$ (A 是常数)，以及电子部分贡献的 $c_{ne} = \gamma T$ (γ 是索末菲常数)。正常态比热容为

$$c_n = A(T/\Theta_D)^3 + \gamma T = c_{nl} + c_{ne}$$

这表示在实验测量的 c_n/T 和 T^2 关系的曲线上，我们容易分别得到点阵和电子比热容。上面已提到，晶体点阵在超导态和正常态没有什么差别，点阵的振动性质也没有什么变化。这样，可以设 $c_n - c_s = c_{ne} - c_{se}$ 。精确测量表明，超导态电子比热容为

$$c_{se} = a e^{-\frac{b}{k_B T}} \quad (12-13)$$

式中， a 是与材料有关的常数， $b \approx 1.7 k_B T_c$ 。例如对于钒， $b = 1.5 k_B T_c$ 。按统计物理学，式 (12-14) 的指数律表示电子能谱的基态与激发态存在一个有限大小的能量差，或者说有一个能隙 b 。超导电子能谱存在这样大小的能隙被比热容和其他的实验（如隧道效应实验等等）所证实。

12.4 迈斯纳态的电磁性质

12.4.1 二流体模型

早期提出的一种超导态唯象理论叫“二流体理论”。这个理论把超导体中的传导电子分为超导的和正常的，超导电流（超流）没有电阻。超导态是由于正常态中的部分电子凝聚成超流电子引起的。温度为 T_c 时电子凝聚开始，温度继续降低就有更多的电子凝聚。超导电子能量低而且有序。这一模型在解释超导体的许多性质上取得了成绩，但是却不能解释在温度 T_c 时为什么会有一部分电子凝聚为超导的这类问题。

按照二流体模型，总电流和电子密度分别是

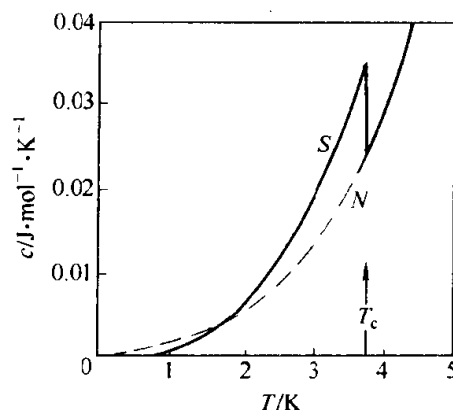


图 12-10 锡在正常态(N)和超导态(S)的比热容

$$J = J_s + J_n, n = n_s + n_n \quad (12-14)$$

式中超导电流和电子密度分别是 J_s 和 n_s , 正常电流和电子密度分别是 J_n 和 n_n 。

12.4.2 London 理论

London 理论也是一个唯象理论, 其正确性由实验确认。这个理论把超导体在超导态和正常态电磁性质的不同归结为超流满足两个 London 方程, 分别描述迈斯纳态的电学和磁学性质 (正常态的电学性质满足欧姆定律, 磁性可以忽略)。它首先是由伦敦 (F.London 和 H.London) 兄弟提出的。London 假设了两个方程

$$\mu_0 \lambda_L^2 \frac{dJ_s}{dt} = E \quad (12-15)$$

$$-\mu_0 \lambda_L^2 \nabla \times J_s = h \quad (12-16)$$

式(12-15)描述零电阻效应, 式(12-16)描述迈斯纳效应, 其中

$$\lambda_L = \sqrt{\frac{m^*}{\mu_0 e^{*2} n_s}} \quad (12-17)$$

叫 London 穿透深度, m^* 和 e^* 分别是超导电子的质量和电荷 (但是后来证明 $e^* = 2e$), h 是局域磁通密度 (是位置的函数)。

方程 (12-15) 表示稳恒情形下超导态没有电场 $E=0$, 描述了稳恒电流下的零电阻性。但是如果电流密度与时间有关, 也就是在交流电的情形下超导体是有电阻的, 这是和实验符合的。

London 方程可以描述迈斯纳效应。利用麦克斯韦方程 $\nabla \times h = \mu_0 J_s$ 后, 上式可以写成

$$\lambda_L^2 \nabla \times \nabla \times h = -h \quad (12-18)$$

计及 $\nabla \cdot h = 0$ 我们得到

$$\lambda_L^2 \nabla^2 h + h = 0 \quad (12-19)$$

在图 12-11 的半无限大情形下, 均匀外磁场 H 沿 z 轴, 电流沿 y 轴, x 轴垂直 N - S 界面, 磁通密度 h 和电流密度 J 互相垂直, 都是 x 的函数, h 仍和外场平行。在此条件下得到方程和边界条件为

$$\lambda_L^2 \frac{d^2 h}{dx^2} + h = 0, x > 0 \quad (12-20a)$$

$$h|_{x=0} = \mu_0 H \quad (12-20b)$$

解这个方程就得到

$$h(x) = \mu_0 H e^{-\frac{x}{\lambda_L}} \quad (12-20c)$$

磁场在超导体中的穿透情况参见图 12-11 (a)。

由 Maxwell 方程, $\mu_0 J_y(x) = -dh_z/dx$, 我们得到电流密度

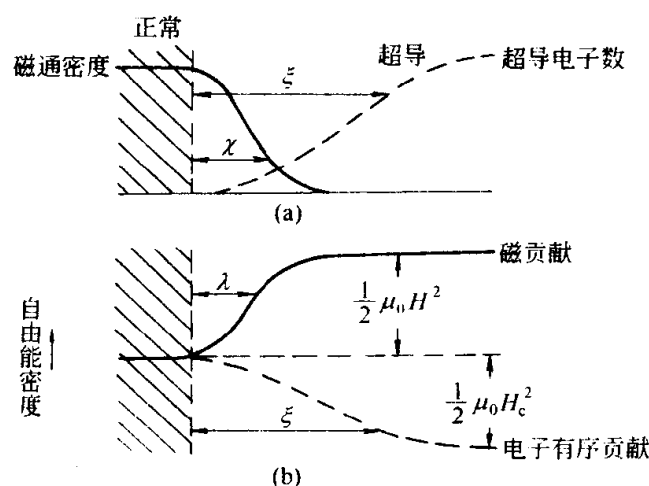


图 12-11 正常-超导界面附近的磁场电子密度和自由能的变化

(a) 超导体中磁场穿透和电子密度的变化; (b) 磁场穿透和电子密度的变化引起的自由能改变

$$J_s(x) = J_s(0)e^{-\frac{x}{\lambda_L}} \quad (12-20d)$$

式中, $J_s(0) = H/\lambda_L$ 。这就表明, 磁通密度和电流密度深入到超导体的距离 $x = \lambda_L$ 后都降低为表面处的 $1/e$, 所以 λ_L 就是磁场或电流在 x 方向的穿透深度。图12-11(a)的实线也定性表示了电流密度的穿透情况。穿透深度是磁场或电流在空间有显著变化的尺度。粗略地说, 在超导体中 $x < \lambda_L$ 的地方磁场已经穿透, 在 $x > \lambda_L$ 的地方才有迈斯纳效应。应当注意, 磁场和电流穿透的方向(x)是和电流(y)及磁场(z)的方向垂直的。磁场和电流有一定的穿透深度是 London 对“完全排斥磁通”概念的修正。穿透深度是超导体的一个特征长度。

λ_L 随超导体性质和温度而变化, 把超导体全部传导电子密度当成超导电子密度可以由式(12-17)估算出 London 穿透深度。由此得到载流子浓度高的金属超导体的 λ_L 比较小, 为 10 nm 的量级, 载流子浓度低的高温氧化物超导体 λ_L 比较大, 为 100 nm 的量级。穿透深度是可以由实验来测量的, 且实验得到的穿透深度值总是比理论的 λ_L 大。从二流体模型看, 不是全部传导电子都是超导电子, 因此根据式 (12-17), 估小了 λ_L 。因为磁场可使超导电子密度降低, 使穿透深度增加, 从而忽略磁场对超导电子的抑制也低估了 λ_L 。但是, 对于宏观尺寸的超导体 (例如样品尺寸 $L = 1\text{mm}$) 来说, $L \gg \lambda_L$, 穿透深度完全可以当成零, 迈斯纳效应是足够精确的。在另一种极端情况, $L \ll \lambda_L$ (如平行于外磁场的薄膜, 纳米超导体等) 则完全没有迈斯纳效应, 和实验相符合。

London 方程在解释超导体的宏观电磁性质方面取得了很大的成功, 但是它在解释另一些实验结果时也遇到了困难。例如, 这个模型假设超流电子

密度在 N-S 界面可以突变, n_s 与磁场无关, 特别是 London 模型无法求得 n_s 。实验结果说明, 从正常态 $n_s(x < 0) = 0$ 到超导态 n_s 的变化是逐步的, 参见图 12-11 (a)。

12.4.3 G-L 理论

比 London 模型更为深入和广泛应用的唯象理论叫 Ginzburg-Landau (G-L) 理论, 它克服了 London 理论的上述缺点。20 世纪 50 年代初, 在 Landau 的二级相变理论的基础上 G-L 理论中引进了“序参量 Ψ ”来描述超导凝聚, 发展了超导电性的唯象理论。序参量概念最初是用来处理结构相变的, 结果非常成功。后来在许多相变系统中序参量概念的应用也非常成功。在 N-S 相变系统中, 序参量是超导电子密度, 即载流子波函数 $\Psi(r) = |\Psi(r)| \exp(i\delta)$ 且 $|\Psi(r)|^2 = n_s(r)$, 其中位相 δ 一般也是位置的函数。自由能密度 F 作为序参量的函数, 并被展开为 $|\Psi|^2$ 和 $|\nabla \Psi(r)|^2$ 的级数。由自由能极小决定的稳定态中, $\Psi(r)$ 和磁场矢量势 $A(r)$ 由一对相互耦合的非线性微分方程, 即 G-L 方程来决定。G-L 理论是一个唯象理论, 其正确性由实验检验。但是 L.P. Gorkov 证明, 在一定情况下 (T 接近 T_c , 序参量和磁场在空间的变化不太剧烈等) G-L 理论可以由超导微观理论导出。

具体地说, 在 T_c 附近, G-L 唯象地将超导体的性质由自由能 F 和 G-L1 和 G-L2 两方程来决定:

$$F_s(T_0, H) = F_n(T_0, H) + \alpha |\Psi|^2 + \frac{\beta}{2} |\Psi|^4 + \frac{1}{2m^*} \left| \left(\frac{\hbar}{i} \nabla - e^* A \right) \Psi \right|^2 + \frac{H^2}{2\mu_0}$$

$$\alpha \Psi + \beta |\Psi|^2 \Psi + \frac{1}{2m^*} \left(\frac{\hbar}{i} \nabla - e^* A \right)^2 \Psi = 0 \quad (\text{G-L1})$$

$$J_s = \frac{e^* \hbar}{2m^* i} (\Psi^* \nabla \Psi - \Psi \nabla \Psi^*) - \frac{e^{*2}}{m^*} \Psi \Psi^* A \quad (\text{G-L2})$$

式中, α 和 β 是温度有关的材料参数, \hbar 是量子力学常数。G-L1 原则上能够求出 n_s , 这是 London 理论没有的。G-L2 原则上能够求出磁场和超流, 说明 G-L 理论至少能够描述弱磁场中超导体的磁性质。

G-L 方程的两个最明显的解是:

① $\Psi = 0$, 这显然描述正常态;

② $\nabla \delta = 0$, $\Psi = |\Psi| = \text{常量}$; 这是 London 方程, 描述的迈斯纳态。事实上将 $\Psi = |\Psi| \exp(i\delta)$ 代入 G-L2 后就得到它的另一种形式:

$$\nabla \delta = \frac{2\pi}{\phi_0} (\mu_0 \lambda^2 J_s + A) \quad (\text{G-L2a})$$

式中, λ 和式 (12-17) 中的 λ_L 一样, 不过常叫 G-L 穿透深度, $\phi_0 = h/e^*$,

称为磁通量子，其物理意义将在下面说明。假设 $\nabla\delta=0$ 和 $|\Psi|^2=n_s$ 就得到第二 London 方程，因此迈斯纳态是 G-L 方程的一个解。而且，从 $\nabla\delta=0$ 和 G-L 方程得出，在迈斯纳态，所有超导电子的位相相同，一个宏观材料中的所有超导电子处于同一个量子态，叫“宏观量子效应”。超导电现象是一种宏观量子现象， Ψ 是宏观波函数。量子态是微观态，宏观上观测到的物理量是大量微观量的平均，宏观现象是许多微观波函数的叠加。

在一般情况下 G-L 方程是一对耦合的非线性微分方程，只能数值求解。

12.4.4 相干长度

用 G-L 方程可以研究表面对超导电性的影响。假设 $x>0$ 是超导态， $x<0$ 是真空或绝缘体。 $x=0$ 处是 N-S 界面，没有磁场和电流。在此情形下 G-L1 和边界条件分别是

$$\frac{d^2f}{dy^2} + f - f^3 = 0 \quad (12-21a)$$

$$f|_{y \leq 0} = 0 \quad (12-21b)$$

$$f|_{y=\infty} = 1 \quad (12-21c)$$

其中 $|\Psi_0|^2 = -\alpha/\beta$ 是没有界面 ($x=\infty$) 时的序参量，只是温度的函数， $f = \Psi/\Psi_0$ 是归一化序参量， $y = x/\xi$ 是归一化变量，

$$\xi = \sqrt{-\frac{\hbar^2}{2m^* \alpha}} \quad (12-22a)$$

将式(12-21a)乘以 $2 \frac{df}{dy}$ 后积分，加上边界条件，容易得到 $f(y) = \text{th}\left(\frac{y}{\sqrt{2}}\right)$ ，即

$$\Psi(x) = \Psi_0 \text{th}\left(\frac{x}{\sqrt{2}\xi}\right) \quad (12-21d)$$

参见图 12-11 (a)，图中的虚线正比于 $\Psi^2(x) = n_s$ 。这里的结果说明，从 $x=0$ 到 ξ ， Ψ 从 0 变到 $\Psi_0 \text{th} \sqrt{1/2}$ 。定性地说， $0 < x < \xi$ ， $\Psi \approx 0$ ，没有超导电子； $x > 0$ 时 $\Psi \approx \Psi_0$ ，有超导电性。 ξ 叫 G-L 相干长度，表示在长度为 ξ 的距离内，序参量是相互关连的，不能发生明显变化。因为 α 、 β 随超导材料和温度而变化，也随超导材料和温度而变化，它也叫“温度有关的相干长度”。在低温，各种氧化物高温超导体的 ξ 约为 1nm 的量级，典型低温超导体（如铅）的相干长度为 100nm 量级，是可以测量的量。相干长度的测量是用实验检验 G-L 理论的方法之一，大量的实验结果证明了 G-L 理论的正确性。London 方程实质上把相干长度当成 0，这是它的局限性之一。相干长度和穿透深度是两个很重要的超导材料参量。

G-L 理论表明，一个有限大小的超导样品是不能全部进入超导态的：在 N-S 表面超导体一侧大约厚为 ξ 的一层是不超导的。根据实验得到的 α 与温

度的函数关系，可得到：

$$\xi = \sqrt{-\frac{\hbar^2}{2m^*a}} \propto \xi_0(1-t)^{-1}, t \approx 1 \quad (12-22b)$$

微观理论证明：

$$\xi(t) = 0.74 \xi_0 (1-t)^{-1/2} \quad (L \gg \xi_0) \quad (12-22c)$$

$$\xi(t) = 0.85 (\xi_0 L)^{1/2} (1-t)^{-1/2} \quad (L \ll \xi_0) \quad (12-22d)$$

式中， L 是电子平均自由程， $t = T/T_c$ 是归一化温度。

12.4.5 表面能

G-L 理论在早期的一个重大成就是预言 N-S 表面能（或界面能）的存在。G-L 方程表明，序参量和磁通密度在界面附近发生了变化，这种变化的程度可以分别由两个特征长度 ξ 和 λ 反映出来。从自由能 F 的定义看，序参量和磁通密度的变化必然导致自由能的变化。定义界面能为单位 N-S 界面引起超导体自由能的增加，则可用 G-L 理论计算界面能。

从 ξ 和 λ 的存在来估算界面能。在一定的温度和外场中，样品表面处正常态和超导态自由能密度之差 $G_n - G_s$ 主要由两部分决定：第一部分是一部分正常电子凝聚成超导电子，使单位体积能量降低（凝聚能） $\frac{\mu_0 H_c^2}{2}$ 。第二部分是排斥磁通引起的能量增加 $\frac{\mu_0 H_c^2}{2}$ （出现 N-S 界面时 $H = H_c$ ）。没有界面没有磁场存在时超导区的自由能密度就是 $\frac{\mu_0 H_c^2}{2}$ 。界面存在的情况参见图 12-11(b)。假设 $x > 0$ 的半无限空间是 S 态， $x < 0$ 的为 N 态，其中外磁场强度是 H_c ，平行于位于 $x = 0$ 的 $y-z$ 平面内的界面。近似地说，磁场的穿透深度是 λ ，对单位面积而言穿透体积是 λ ，则 $\frac{\mu_0 H_c^2}{2} \cdot \lambda$ 就是自由能的减少。同样， $\frac{\mu_0 H_c^2}{2} \cdot \xi$ 是自由能的增加，因为界面存在伴随着超导电子密度的完全消失。在图 12-11(b) 中分别示意出了这两种贡献，这两种贡献的总和就是界面能 σ ：

$$\sigma = (1/2) \mu_0 H_c^2 (\xi - \lambda) \quad (12-23a)$$

利用 G-L 参数 $\kappa = \lambda/\xi$ 的定义，式(12-23a)的定义可以写成

$$\sigma = (1/2) \mu_0 H_c^2 \xi (1 - \kappa) \quad (12-23b)$$

这表示， $\kappa > 1$ 时 $\sigma < 0$ ，表面能是负， $\kappa < 1$ 时 $\sigma > 0$ ，表面能是正。

界面能 σ 的定量计算必须求助于 G-L 方程，在一般情况下只能数值地求解 G-L 方程。详细的理论计算超出了篇幅，这里列出几个结果：

$$\sigma = 0, \quad \kappa = \frac{1}{\sqrt{2}} \quad (12-23c)$$

$$\sigma = 1.89\xi(\mu_0 H_c^2/2), \quad \kappa \ll 1 \quad (12-23d)$$

$$\sigma = -\lambda(\mu_0 H_c^2/2), \quad \kappa \gg 1 \quad (12-23e)$$

在一般情况下有

$$\sigma < 0, \quad \kappa > \frac{1}{\sqrt{2}} \quad (12-23f)$$

$$\sigma > 0, \quad \kappa < \frac{1}{\sqrt{2}} \quad (12-23g)$$

表面能 $\sigma > 0$ 的超导体叫第一类超导体, 此种超导体内出现 N-S 界面时在能量上不利。表面能 $\sigma < 0$ 的超导体叫第二类超导体, 此种超导体内出现 N-S 界面时在能量上是有利的。

12.4.6 磁通量子化

我们现在进一步考虑超导体内的正常区域中的磁通量。参看图 12-12, 假设块超导体在磁场中, 有“×”符号区的是正常区, N-S 界面是 C。由于磁场透入, S 区内距界面 C 约 λ 范围内也有磁通。利用 (G-L2a) 并进行环路积分得到

$$\oint (\mu_0 \lambda^2 \mathbf{J}_s + \mathbf{A}) \cdot d\mathbf{l} = \frac{\phi_0}{2\pi} \oint \nabla \delta \cdot d\mathbf{l} \quad (12-24a)$$

定义

$$\oint (\mu_0 \lambda^2 \mathbf{J}_s + \mathbf{A}) \cdot d\mathbf{l} = \Phi_L, \quad \oint \mathbf{A} \cdot d\mathbf{l} = \iint \mathbf{h} \cdot d\mathbf{S} = \Phi_A$$

式中, Φ_L 是周界 L 中的总磁通量 Φ (包括正常态中的和超导态中的) 和电流的贡献 $\oint \mu_0 \lambda^2 \mathbf{J}_s \cdot d\mathbf{l}$ 之和, Φ_A 是正常区中的磁通量。 Ψ 的单值性要求 $\oint \nabla \delta \cdot d\mathbf{l} = 2\pi N$, 所以

$$\Phi_L = N\phi_0, N = 1, 2, 3, \dots \quad (12-24b)$$

式中, $\phi_0 = \frac{h}{e^*}$ 。把普朗克常数 h 和 $e^* = 2e$ 代入后得到 $\phi_0 = 2.07 \times 10^{-15} \text{ (Wb)} = 2.07 \times 10^{-6} \text{ (mT} \cdot \text{mm}^2)$ 。伦敦首先得到这个结果, 并称 Φ_L 为全磁通, ϕ_0 为磁通量子, 是磁通量的最小值。式 (12.24b) 说明, 复连通超导体内的全磁通是量子化的。迄今已知, 除超导体外, 液氮也具有这种量子化现象。在这里, 微观现象在宏观尺度内

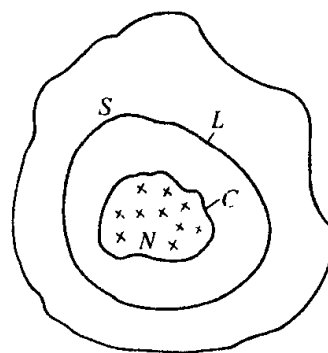


图 12-12 磁场中的复连通超导体
(有×号的是正常区 N, C 是正常-超导界面, C 以外是超导区 S, L 是积分回路)

表现出来了，所以是宏观量子效应。应当注意，超导体所围的空洞中的磁通量（即 Φ_A ）不是量子化的。如果积分回路在 N - S 界面附近，超导电流不是零。但是在 L 深入到超导体的内部时，根据迈斯纳效应有 $J_s = 0$ ，这时只有矢量势对全磁通有贡献，超导电流对全磁通量已没有贡献，全磁通就是普通磁通量了。但是 λ 和宏观尺度比是很小的，当 L 离开 C 稍微深入超导区时 $J_s \approx 0$ ，因而全磁量和 N 区内的磁通量相差甚小。一般就认为超导环（由上面可知环不一定是圆环）中的磁通是量子化的。

实验已观测到磁通量子化现象，并测得 ϕ_0 确为 $2.07 \times 10^{-15} \text{ Wb}$ ，证明 $e^* = 2e$ ，这是超导态中 Cooper 对的存在的有效证据。

12.5 第二类超导体

20 世纪 60 年代在超导材料实用化方面取得了两项突破性进展。其中一项是以发现了像 Nb_3Sn 这样的有高临界温度、高临界磁场和高临界电流密度的高临界参数超导材料。以 Nb_3Sn 为例，Kunzler 等制得的 Nb_3Sn 线材在 4.2K 和 8.8T 的临界电流密度达到 10^5 A/cm^2 。后来知道这些材料是强钉扎的第二类超导体，这类超导体有两个临界磁场：下临界场 H_{c1} 和上临界场 H_{c2} ：

$$H_{c1} = \frac{H_c}{\sqrt{2}\kappa} \ln \kappa + 0.08 \quad (12-25a)$$

$$H_{c2} = \sqrt{2}\kappa H_c \quad (12-25b)$$

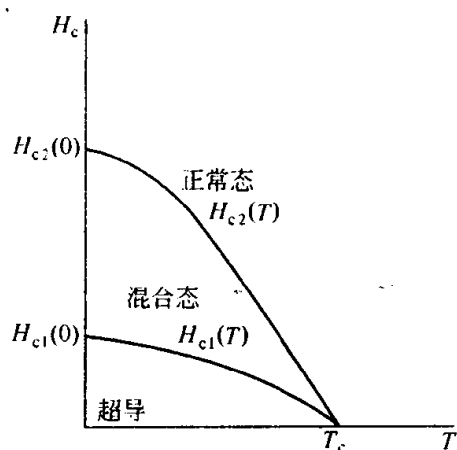


图 12-13 第二类超导体的相图

磁场小于 H_{c1} 时，超导体处于迈斯纳态；磁场大于 H_{c2} 时，完全恢复其正常态。当磁场介于 H_{c1} 及 H_{c2} 之间时，则是正常与超导的“混合态”，此时超导体一部分区域仍处于超导态，另一部分区域则处于正常态（图 12-13）。为了区别，将水银这一类只存在一个热力学临界场 H_c 的超导体称为第一类超导体。图 12-13 示意第二类超导体的相图。

由于第二类超导体的 $\kappa > \frac{1}{\sqrt{2}}$ ，所以

H_{c2} 总是比 H_{c1} 大。许多第二类超导体的一个重大特点是 $\kappa \gg 1$ ，它们的 H_{c2} 很高， H_{c1} 很低。像典型的低温超导体 NbTi 和 Nb_3Sn ， $\kappa > 50$ ，它们的 H_c 也只有几十 mT，但它们的 H_{c2} (4.2K) 分别约有 12 和 20T，于是就有可能产生很强的磁场。这些高 κ 的超导体在强磁场中的临界电流密度也可以非常高，与它们能处于混合态的性质有

关。它们在磁场中，当磁场强度超过 H_{cl} 时就有一部分磁通进入样品。进入样品中的磁通是以一个个磁通涡旋线的形式存在的，每个磁通涡旋线的磁通量是磁通量子 ϕ_0 。图 12-14 (a) 表示磁通线排成六角格子，也叫三角点阵，或叫磁通涡旋点阵，也叫磁通涡旋物质。图 12-14 (b) 表示观测磁通线的一种方法。在这一实验方法中，先蒸发铁磁性材料并使之形成纳米微粒，漂浮在超导样品上方，这些磁性微粒受磁场作用沿磁力线运动，最后沉积在磁通线的露头处，缀饰了磁通涡旋线在样品表面的图案，称叫缀饰法。理论计算和实验观察都证明，磁通涡旋物质的结构多数是三角格子，还有部分四角格子的结构。图 12-15 是缀饰法观测到的这两种结构的磁通涡旋物质。缀饰实验还证明，磁通涡旋物质也有结构缺陷，如位错，空穴，填隙等等。

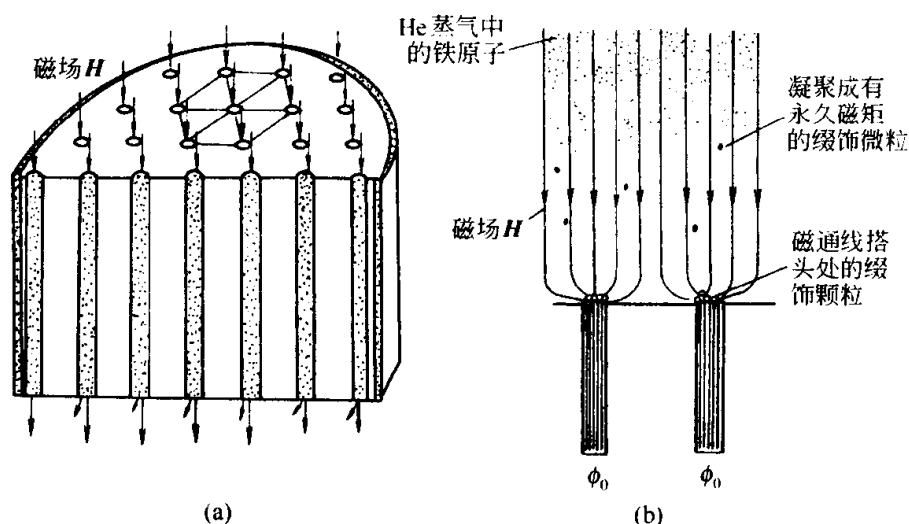


图 12-14 磁通涡旋物质的示意和观测方法

(a) 三角结构的磁通涡旋物质；(b) 涡旋物质的一种观测方法-缀饰法

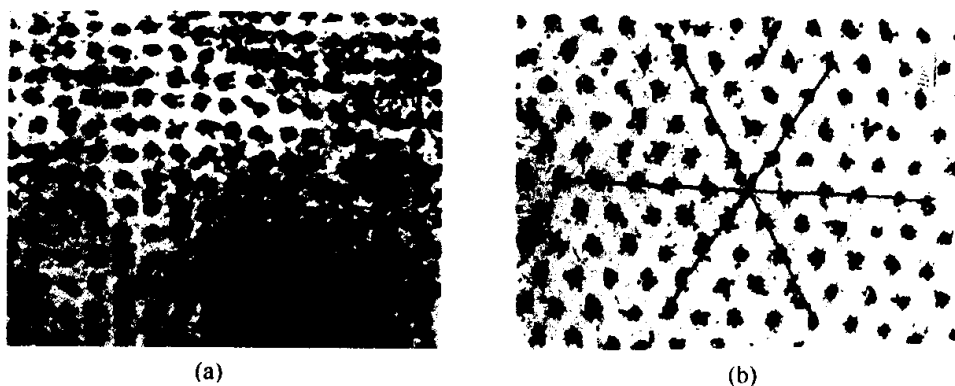


图 12-15 缀饰法观测到的两种结构的磁通涡旋物质

(a) 四角结构；(b) 三角结构

第二类超导体的重要性在于用适当的工艺引进所谓磁通涡旋线的钉扎中心后，临界电流密度很大。混合态的超导体从整体载流效果来看和完全处于超导状态一样，但它们的 H_{c2} 很大，在强磁场中完全可以在混合态下工作，表现出完全的无电阻特征。对于很强的磁通钉扎，其相应的临界电流密度 J_c 也就很大。因此，第二类超导体的很大的上临界场和临界电流密度使它们具备了作为实际强磁场材料的条件。

12.6 超导隧道效应

20 世纪 60 年代超导研究的另一项重大突破是在弱连接超导体中发现了约瑟夫森效应。所谓的弱连接超导体是两块超导体中间夹一块纳米厚度的绝缘膜，形成超导层-绝缘层-超导层 (S-I-S) 的结构，类似于夹心层很薄的三明治，如图 12-16 所示。

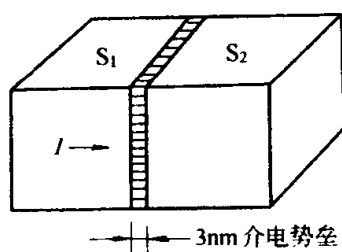


图 12-16 约瑟夫森
结示意图

由于中间的绝缘层比较薄，使它的两侧的超导体在电磁性质上弱耦合在一起。例如，当有一个很小的电流从一个超导体穿过绝缘层而流到另一个超导体时，如果电流很小，没有超过这种结构的临界电流 (I_c)，则两侧的超导体层之间没有电压，整个弱连接超导体呈现零电阻性。两超导体中间的绝缘 (真空，正常) 层也能让超导电流

通过的现象叫超导隧道效应，可用超导微观理论解释。在 20 世纪 60 年代初，这一现象首先由约瑟夫森在理论上预言，不久即被实验证实，所以也叫约瑟夫森效应，超导隧道结也叫约瑟夫森结。更具体地说这里说的是直流约瑟夫森效应。

根据以上隧道结的原理，两块超导体中间夹一层金属也可以形成约瑟夫森结 (S-N-S)。进而，超导体中间不夹东西 (真空)，而只是靠得很近也行 [扫描隧道电子显微镜 (STM) 用的结]。约瑟夫森结还可以是两块超导体的点接触，或微桥接触等结构，其关键是让两块超导体间能有弱连接而导致隧道效应。图 12-17 中示出几种约瑟夫森结。

弱连接超导体的连接弱在两个方面。一个是结的临界电流密度 (J_c) 很小，这意味着很小的电流密度就会破坏零电阻性。另一个是它对磁场极为敏感。在这里，很弱的磁场是指通过结的磁通量只要变化 $\phi_0/2$ 时就足以使通过结的电流从最大变到最小。因为 ϕ_0 是个极小的磁通量，利用这个对微弱磁通极度灵敏的特性，就制成了直流超导量子干涉器件 (DC SQUID)，它是超导量子干涉器件之一。

当约瑟夫森结两端加一个直流电压 V 时样品中就会出现频率为 f 的交

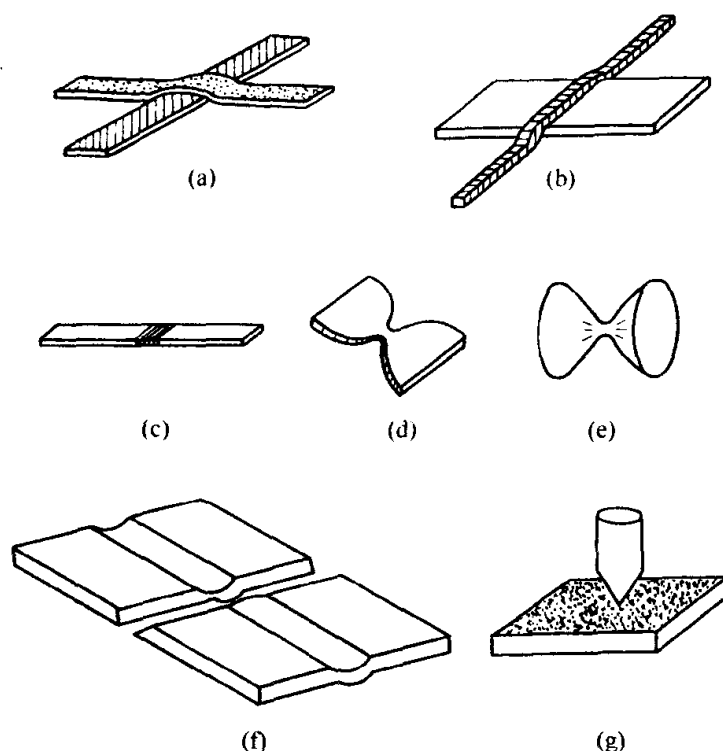


图 12-17 几种约瑟夫森结

(a) 隧道结; (b) 临近效应桥; (c) 一维微桥; (d) 二维微桥;
(e) 三维微桥; (f) 变厚微桥; (g) 点接触

流电流 (也就辐射电磁波), 叫交流约瑟夫森效应。 V 和 f 满足

$$f = \frac{eV}{\pi\hbar}$$

式中, e 是电子电荷, \hbar 是普朗克常数。 f 也叫做约瑟夫森频率。反过来, 当一个频率为 f 的电磁场照射到约瑟夫森结上后, 样品上也会出现电压 V 。按照上式, $1\mu\text{V}$ 电压的约瑟夫森频率是 4.86MHz , 是射频。利用约瑟夫森效应制成了交流 (射频) SQUID。约瑟夫森效应已成为微弱电磁信号探测和其他各种电子学应用的物理基础。

还有一种隧道结叫单电子隧道结, 或准粒子隧道结, 因为从超导凝聚体激发的单电子叫准粒子。给这种隧道结加上电压 V , 能够拆散一个 Cooper 对, 变成两个准粒子, 电压所做的功是 $eV = 2\Delta$, 2Δ 是 Cooper 对超导电子的束缚能。准粒子隧道结的 $I-V$ 特性曲线实验可以用来测量超导能隙。测量电流小于隧道结的临界电流 ($I < I_c$) 时 $V = 0$, $I > I_c$ 时隧道结上就突然出现了电压:

$$V = \frac{2\Delta}{e}$$

2Δ 就是一对超导电子被激发为两个正常电子所需的能量。这是测量超导能

隙的一种方法，相应的微分电导率-电压 (σ - V) 曲线在能隙处有尖锐的极大，图 12-18 是一个实验例子。

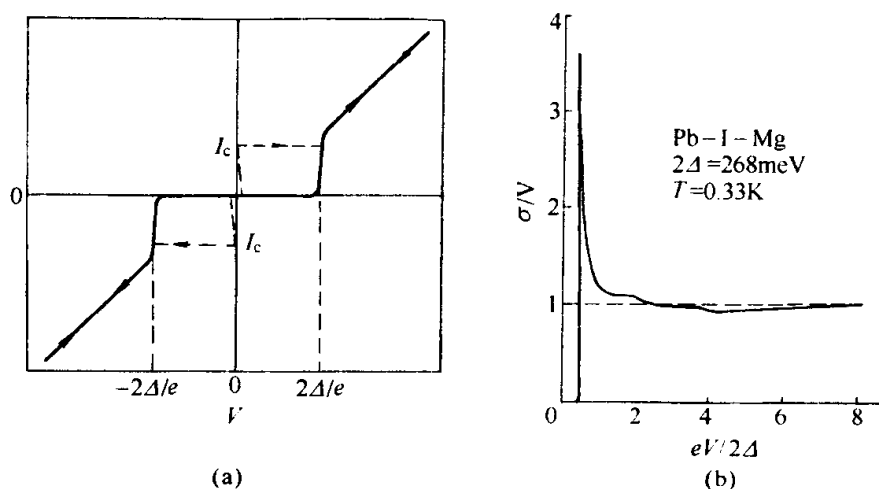


图 12-18 超导隧道结 $I-V$ 和 $\sigma-V$ 曲线
(a) S-I-S 隧道结 $I-V$ 曲线；(b) N-I-S 隧道结 $\sigma-V$ 曲线

如果不是两块超导体，而是正常金属-绝缘体-超导体 (N-I-S) 隧道结，同样方法的 $I-V$ 曲线测量将会看到，电压跳跃和相应的微分电导率的尖锐的极大发生在

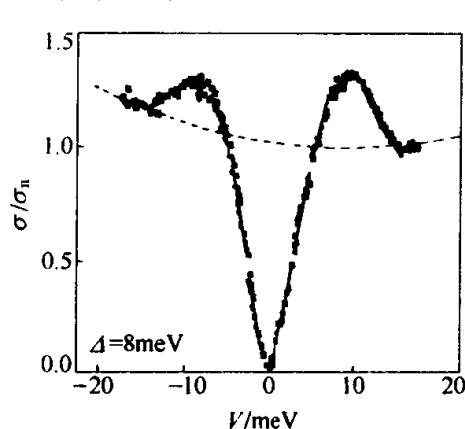


图 12-19 MgB_2 的微分电导率曲线

$$V = \frac{\Delta}{e}$$

这也是单电子隧道效应。单电子隧道效应也是测量超导能隙的一个重要方法，是超导扫描隧道电子显微镜 (STM) 的基本原理，图 12-19 就是一个实验例子。图中所示的是温度很低的情况，所以电导率的极大很尖锐，测得的能隙也准确。实验测量若比较高的温度下进行，就有热涨落，因而确定 V 跃变的位置比较模糊。图 12-19 是 2001 年发现的 MgB_2 超导体隧道结的微分电导率曲线，那里的两个微分导电

极大 (正反向电压) 不很尖锐就是这个道理，但是电导率尖峰的位置仍旧比较容易确定。图中给出了 MgB_2 的能隙 $2\Delta \approx 8meV$ ， σ_n 是正常态电导率，细抛物线是背景信号。

12.7 超导电性的微观图像

定量地表达下面的超导微观理论的图像就是超导微观理论。但是这样做

要求比较专门的物理和数学工具，超出了我们的目的。

12.7.1 实验和理论的启示

BCS 理论的出现并不是偶然的。在 BCS 理论出现之前，对超导电性的实验和理论研究已取得很大成绩。这些研究使 BCS 的成功成为可能。BCS 正是及时地吸取了这些研究成果，并在此基础上前进，取得了超导理论的突破。那么，超导理论出现前实验上已经取得了哪些重大成就，从而为超导理论的完成提供了线索呢？

实验研究已经肯定，迈斯纳态的零电阻性和完全逆磁性是可逆性质。从正常相转变到超导相是一种相变，在磁场中超导-正常的转变过程有潜热，人们用量热法和磁化法测出了潜热的大小，是一级相变；如果没有磁场，则超导-正常相变时没有潜热，但有比热容跳跃，是二级相变。所以，超导态是物质的一种新相。

相变有不同种类，它们有不同的机制。超导相变时材料的结构和晶格振动没有什么变化，所以不是结构相变。这就暗示，超导相变有可能是电子的行为引起的。实验观测到，超导体有“同位素效应”，即某超导体样品的临界温度与它的同位素原子量的平方根成反比。因为原子振动的频率与原子量的平方根成反比，同位素效应暗示着晶体离子虽不是载流子，它们的振动（声子）可能参与了超导转变过程。声子的频率和原子量的平方根成反比，所以同位素效应暗示了超导体中电子的行为和声子之间有着密切联系。

实验研究发现，N-S 相变是在极短的时间内完成的。因为导体中的自由电子数目是 $N = 6.023 \times 10^{23} / \text{mol}$ 的数量级，这就说明相变时样品中巨大数量的电子在极短时间内几乎是同步地凝聚到超导态或激发到正常态，表示超导电子系统是高度相关的，是强关联系统。

热力学研究已说明，超导态不仅比正常态的能量低，而且更有序，超导态是一种凝聚态。前面已说明，超导体相变不是原子的位置和运动的有序。那么，超导体从正常态转变为超导态就可能是电子从能量高的无序态变成能量和动量都为零的有序态了。

超导体中的一个电子要么在低能量的超导态中，要么在高能量的正常态之中，存在潜热表明在这两个态的电子能量有跃变。能谱中的能量跃变就是能隙。超导态在能隙的底部，是基态；而正常态在能隙的上部，是激发态。光吸收实验不仅能确认处于超导态的电子能量低于正常态，而且能测出超导态电子能谱中的能隙的大小。因为由于能谱中的能隙，当用光照射超导体时，如果光子的能量小于能隙，超导电子不能吸收光子，只有能量大于能隙的光子才能被超导电子吸收。于是用不同能量（频率）的光子照射超导体时就观测到超导体能隙的大小。另外，超导隧道实验也是测量能隙的好方法，

但是它是在超导微观理论出现的 20 世纪 60 年代才被发现和使用的。

比热容和光吸收等多项测量证明，超导体的能隙的量级为 $k_B T_c$ ， k_B 是玻尔兹曼常数。按 $T_c = 4\text{K}$ ，能隙约为万分之一电子伏特的量级，这个量值远小于金属的费米能 E_F （约为 1eV ）。我们知道，晶体中传导电子能量就是 E_F 的量级，因此一个电子从正常态凝聚到超导态时其能量降低到仅仅是它原有能量的万分之一，这是一个不容易理解的问题。不过，晶体的声子的能量倒是这个数量级。虽然超导体中的能隙和离子的振动（声子）作用有联系已由同位素效应暗示了，可是声子能量这样小，它们的影响是不会有静电库仑互作用势能那样大的。因此将声子与超导现象联系起来时需要合理地将其其他更强大的作用（例如库仑互作用）处理好。

早期提出的二流体理论把超导体中的电子分为超导电子和正常电子。超导电子运动形成超导电流（超流），没有电阻。超导态的出现是正常态中的一部分电子凝聚成超流电子引起的。温度降到 T_c 时这种凝聚开始，温度继续降低就有更多的电子发生凝聚。超导电子能量低而且有序，这一模型在解释超导体的热力学性质和电磁性质时取得了成绩。但是二流体理论不能解释在温度是 T_c 时为什么会有一部分电子凝聚为超导电子。一个真正的超导微观理论应当阐明，超导电子的能量为什么低，能隙有多大，有序度为什么大，然后在这个基础上解释超导态的各种性质。虽然如此，二流体理论的成功明确地告诉人们是电子凝聚导致超导态。

London 理论是超导态的另一个成功的唯象理论。这个理论把超导体在超导态和在正常态的电磁性质的不同归结为超导态满足两个 London 方程。London 理论表示，超导体的磁性质是由于超导电子波函数是完全刚性的：在磁场中，一部分电子波函数完全不受影响。London 方程给出了这个性质的一种解释。有磁场时电子的动量的平均局域值是

$$\mathbf{P}_s = m^* \mathbf{v}_s + e^* \mathbf{A}$$

此式可以重写成

$$\mathbf{P}_s = e^* (\mu_0 \lambda^2 \mathbf{J}_s + \mathbf{A})$$

和 London 方程相比后可知， $\mathbf{P}_s = 0$ 。于是 London 理论告诉人们，超导态中电子动量等于 0。和所有的唯象理论一样，London 理论也不能说明电子为什么会凝聚到超导态这样一些根本问题。

最成功的唯象理论要算是 G-L 理论了。这个理论用序参量来描述超导态、超导电子数和能隙都和序参量相联系。所以，超导态的形成是能隙的形成、超导电子的出现和增加、有序度的增加和能量的降低。比起 London 理论来，G-L 理论原则上有了决定超导电子数的方法。虽然 G-L 理论是唯象理论，但它却成功地被用来解决更多超导问题，是超导电性研究的非常有用

的理论工具。London 预言过，G-L 理论则更明确地告诉人们，超导载流子是电子对，因为 $e^* = 2e$ ，并为磁通量子化实验确认。Pippard 首先指出的，G-L 理论简明地表达出的相干长度即超导电子关联长度，电子对的尺寸。

12.7.2 电子-声子机制

微观理论的一个基础是要找到是什么原因使电子在超导态发生了凝聚。首要的是弄清楚电子之间的相互作用。一般讲，除了库仑作用外，还有几种可能使电子间产生相互作用的方式，如声子，自旋波，其他能带中的电子等。1950 年，Fröhlich 在超导微观机制研究方面迈出了重要的一步。他首先指出，电子间经声子媒介能产生新的相互作用，在一定条件下这种作用可以是吸引的，虽然这种吸引可能很微弱。普通传导电子是自由的，互相吸引着的电子处于束缚态。束缚态的形成就是能量的降低，电子的凝聚。弗罗列希的理论能够解释超导体的同位素效应。

真空中的电子之间存在库仑排斥力，电子间怎样经过声子作用而吸引呢？金属中的传导电子可以当作遵从泡利原理的自由电子。对任一个传导电子来说，它和其他电子的同性相斥力加上它和所有正离子的吸引作用合起来可以忽略，超导体中情况也基本如此。但是超导电子-离子间的吸引作用比金属良导体的要强一点。电子周围带正电荷的晶格会因为和电子的库仑作用被扰动：电子发射了声子。假设电子 1 发射了声子，晶格波动将扰动传播出去，使此扰动有可能影响其他的电子，例如电子 2。这是说，电子 2 的运动受到了电子 1 的影响，传递这种影响的是声子。两个电子 1 和 2 通过声子有一定的动量和能量交换，当声子在传播过程中被第 2 个电子吸收时第 2 个电子就获得了动量和能量，按动量和能量守恒定律，第 1 个电子损失了相同的动量和能量。这就是两个电子之间通过晶格而产生相互作用的过程，这种电子通过声子交换动量和能量就是电子间的一种新的相互作用。

按照量子理论，费米面附近的任意两个自由电子，如果其能量分别为 E_1 和 E_2 ，动量大小相等而方向相反，自旋的方向相反，且满足 $E_1 - E_2 < k_B \theta_D$ 时它们是吸引的，否则就相互排斥。金属电子论中传导电子的能量约为费米能 E_F ，位于费米面上。所以 E_1 、 E_2 和 E_F 相差不多。

12.7.3 Cooper 对

上面已经说明，电子-声子互作用机制在一定条件下可以使两个电子相互吸引。1956 年 Cooper 发表了关于在一个大量电子的系统中两个电子相互吸引后就束缚在一起而形成电子对的论文。他证明，费米面上的两个自由电子（假设是能量为 E_1 的电子 1 和能量为 E_2 的电子 2），只要它们有吸引相互作用，不管多弱和什么机制，都要形成束缚态。这样的两个处于束缚态电子的总能量，就略小于 $E_1 + E_2$ 。这样的一对能量和费米能相近，动量大小相

等方向相反，自旋相反，相互束缚在一起的电子对叫“Cooper 对”。两个电子形成了 Cooper 对后，降低了能量，这种束缚态就是稳定的。第三和第四个自由电子也会形成低能状态的 Cooper 对，如此等等。这样，正常态变成了绝对不稳定的态了。

Cooper 对是由费米面上两个电子形成的，这一点非常重要。因为如果是两个电子的系统，在物理学上叫做两体问题，即使两体间有吸引势也不一定形成束缚态，所以相互吸引的两个电子不能叫 Cooper 对。利用量子力学的不确定关系就能证明，Cooper 对电子间的距离（Cooper 对尺寸） ξ 是相当大的。 ξ 实际上就是凝聚成对的电子互作用距离，叫“BCS 相干长度”。BCS 相干长度随超导体不同而异，一般是在 $1 \sim 100\text{nm}$ 的量级，也就是说，一个 Cooper 对电子中间可以有几十到几百个原子。可见从微观尺度讲，Cooper 对是很大的，Cooper 对电子间的相互作用一定是非常的微弱。

Cooper 对中两个电子的总轨道动量和总自旋动量都是零。也就是说，尽管 Cooper 对的每一个电子的速度都和费米速度差不多，但它们的质心是不动的： $P_s = 0$ ，符合 London 方程的结论。Cooper 对中两个电子的自旋相反是泡利原理的要求。

Cooper 对电子比两个自由电子更有序。但是这种有序是方向上的有序，和位置的有序不同。它们总是面对面地移动，围绕着它们的质心旋转，所以它们也可以说是运动有序。

12.7.4 BCS 理论图像

金属的费米面附近有大量电子，更多的电子间的互作用又是怎样呢？BCS 理论认为，Cooper 的简单结果可以推广到许多相互作用着的电子系统。他们假设与超导态相关的相互作用就是上述使任意两个电子形成 Cooper 对的电声子相互作用，所有其他电子的存在对某个 Cooper 对的影响仅仅是这个 Cooper 对电子被散射后要遵守泡利原理，只能占据那些尚未被其他电子占据的态。

绝对零度下的正常自由电子气是费米能以下（费米球内）的态全被占据，球外的态全空着。如果电声子互作用使费米面上一对电子形成 Cooper 对并降低总能量，则将有更多的费米面下的电子去形成 Cooper 对，进一步降低总能量，这个过程直到平衡为止。绝对零度时，费米面附近的电子全部凝聚成 Cooper 对，大量 Cooper 对电子的出现就是超导态的形成。超导态中电子凝聚成 Cooper 对，就使它的总能量降低，并比正常态更有序。费米面附近有多少电子凝聚成 Cooper 对呢？我们可以作这样的估算。因为电子是通过吸收和发射声子而形成 Cooper 对的，声子能量的平均值大约是 $k_B \Theta_D$ ，所以电子的能量和费米能的差别也就大约为 $k_B \Theta_D$ ，这就是“费米能附近”

的含义。当所有这些电子凝聚后单位体积材料的总能量下降量就叫凝聚能密度，用实验可测量临界磁场来表示是 $\frac{1}{2}\mu_0 H_c^2(T)$ 。从这个表达式也能看出凝聚能是温度的函数，因为临界磁场是温度的函数。在 $T = T_c$ 时，凝聚能密度是零，因为 $H_c = 0$ ，没有电子凝聚到超导态。随着温度降低，凝聚能大小增加，凝聚到超导态的电子也增加。在温度降到绝对零度时凝聚能最大，为 $\frac{1}{2}\mu_0 H_c^2(0)$ 。此时费米面附近的电子全部凝聚为 Cooper 对。温度升高的过程恰好相反。

正常金属有电阻是因为载流子受到散射而损失了能量，要维持稳恒电流就需要外电场做功。但同时晶格却从散射过程中获得了能量，即焦耳热。在超导态情况下，组成 Cooper 对的电子也不断地被散射，但这种不断地散射不影响 Cooper 对质心动量，只是使 Cooper 对得以维持。所以有电流通过超导体时 Cooper 对的定向匀速运动不受阻碍，电子的能量没有损失，也就没有电阻。呈现电阻的散射应是改变 Cooper 对质心动量的散射，这种散射是一种拆散电子对的散射。拆散 Cooper 对需要的最小能量为 2Δ ，所以只有这种能量有供给来源时这种散射才会发生。在低电流密度下无法给拆对提供能量，因此能改变 Cooper 对总动量的散射事件完全被制止了，所以没有电阻。换句话说，和正常导体中电子不同的是，超导态电子受到声子的散射后又同时吸收了同样的声子，它们没有损失能量，也不需要电场力做功来补充能量和动量，所以没有电阻。

12.8 超导材料

现在已经应用的超导材料，一类是要求其承受大电流和强磁场，叫强电超导材料；另一类是利用约瑟夫森效应的，只涉及小电流和弱磁场，叫弱连接超导材料或超导电子材料。直到 1991 年，所应用的这两类超导材料都是低温超导材料。

12.8.1 强磁场超导材料

1960 年发现铌三锡 (Nb_3Sn) 的临界电流密度 J_c 在温度为 4.2K 和磁场为 8.8T 的情况下可达 10^4A/cm^2 以上，这是强电高场超导材料发展的起点。很快，人们就知道像 Nb_3Sn 这样的超导体会有这么大的 J_c 是因为高临界参数的第二类超导材料晶体缺陷对磁通线有钉扎力。现已能利用普通冶金工艺方法在高临界参数超导材料中有意识地引进大量晶体缺陷，或有意识地掺一定数量的某种杂质以增加其钉扎力，从而大大提高它们的临界电流密度。材料的临界磁场和临界电流密度很高是强电应用的基本条件。

仅仅有高的临界磁场和临界电流密度仍不能使超导材料完全实用化，因

为超导材料具有“不稳定性”这样的缺点。所谓不稳定性就是在一定的电流和磁场中，很小的扰动就可能引起磁通线沿力方向运动。磁通线一旦运动，超导体就表现出电阻，产生热量，并且其发热量和磁通扩散速度成正比。超导材料的导热能力一般都比较差(热导率小)，热扩散的速度远远低于磁通扩散速度，磁通运动产生的热量基本上来不及散发出超导体去，从而使超导体的温度迅速升高，以至于最终使超导体的温度高于它的临界温度，使超导态突然变为正常态(失超)。由于磁通线运动的速度很快，上述的失超过程实际上是在很短的时间内完成的。传导大电流的超导材料一旦失超是极其危险的，会造成设备烧坏直至更严重的后果。

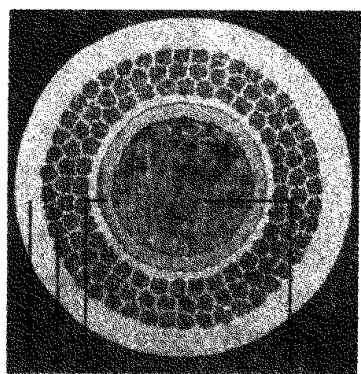


图 12-20 铜稳定化多芯扭绞超导线材 Nb_3Sn 的截面

为了克服不稳定性，已研究出两种有效方法，一是在超导线(带)材外复套一层良导金属(如银，铜或铝)，形成复合导体材料；二是将复合导体中的超导线做成细丝(直径约 $1\mu\text{m}$)，使一根复合线含有数千股细丝并扭绞，形成所谓多芯扭绞复合线材。一根 6516 芯超导纤维的、铜的稳定化多芯扭绞的商品 Nb_3Sn 线材的横截面如图 12-20 所示：1 是 CuSn 合金，2 是 Nb_3Sn 纤维，由铌纤维和 CuSn 合金扩散反应而成，3 是 Ta 层，阻止 Sn 扩散进入 4，4 是高导电率的纯铜，一旦载有大电流的超导线失

超，它能起分流，也就是稳定化作用。现在，大部分强电用超导线(带)材都是这种铜稳定化材料。由于具超导合金和金属化合物稳定化获得成功，商品超导材料的生产在 20 世纪 60 年代后期得到迅速发展，出现了各种尺寸和外形的超导材料。它们可分为两类：一类是合金，如铌钛合金等，它们的机械性能好(强度大，韧性好)，容易生产，价格便宜，性能稳定，安全可靠，是最广泛使用的超导材料。但是它们的临界磁场低，一般用作 8T 以下的磁体材料。另一类是铌三锡等金属间化合物，其优点是临界磁场高，但是机械性能差(硬而脆)，较难生产，不便于使用，价格也贵，一般用作高磁场材料。现在，利用超导体制成的小型超导磁体已经广泛用于各种科学研究中，能产生 20T 以上的场强。

12.8.2 超导电子材料

超导电子材料也叫弱电应用的超导材料，一般是做成器件后出售。多数情况是将超导体先做成薄膜，然后由膜加工成适当的元器件。例如，目前最准确的电压标准仪器，其心脏部件是 4 个铅膜-氧化铅膜-铅膜做成的 S-I-S 型约瑟夫森结。这种结很稳定，质量也好，装备在美国国家计量局作电压标

准（原理是交流约瑟夫森效应中 V - f 关系）使用几十年了。但是这类电子材料制造难度大，适用于工业规模的制备技术尚未解决。弱连接超导材料的应用目前仍限于高科技仪器方面。

12.8.3 高温超导材料

虽然目前只发现液氮温区的超导体，但人们对这一突破高度重视，被认为是 20 世纪科学上最伟大的发现之一。由于新高温超导材料研究的意义重大，受到各方面的高度重视，在众多的课题组和实验室中聚集了大批优秀科技人员，使近十几年超导材料和超导物理研究取得了飞速发展，图 12-21 表示了这种趋向（图中 * 表示高压下的 T_c ）。

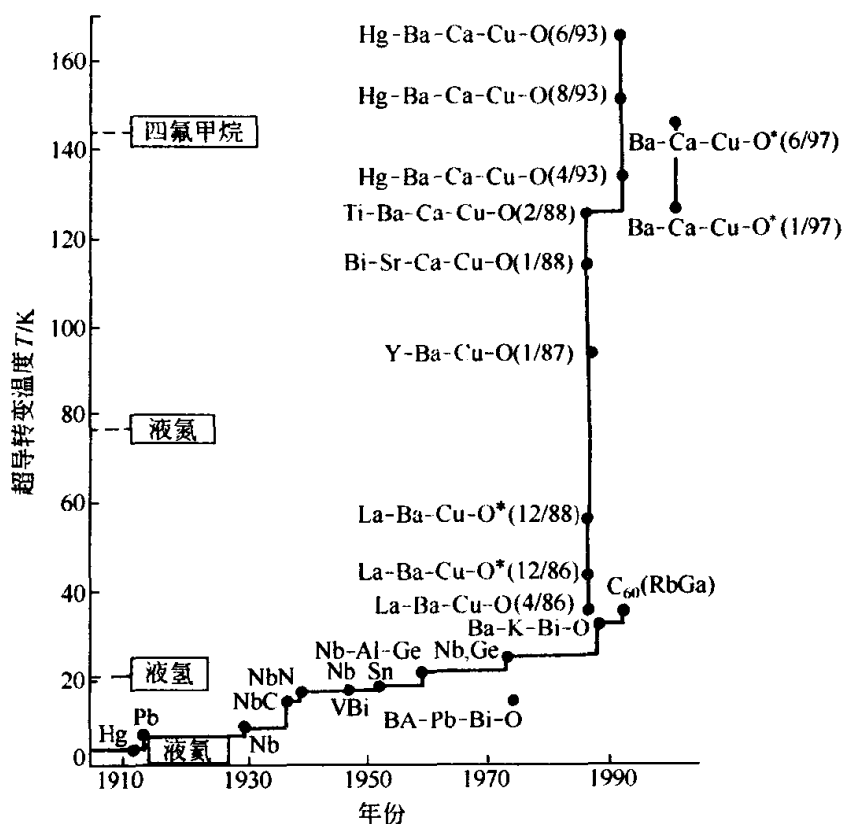


图 12-21 提高 T_c 年表

其中，(6/97) 和 (1/97) 分别指 1997 年 6 月和 1997 年 1 月

到目前为止，已发现的高温超导材料很多，其中有五种属典型复杂金属氧化物：镧钡铜氧化物体系（ $T_c = 36\text{K}$ ）；钇钡铜氧化物体系（ $T_c = 93\text{K}$ ）；铋锶钙铜氧化物体系（有多个不同的高温超导相，临界温度最高的两个相分别为 110K 和 85K ）；铊钡钙铜氧化物体系（铊系，最高的临界温度为 125K ）；汞钡钙铜氧化物体系（汞系，最高的临界温度为 133K ）。后四种体系的高温超导材料已是可以在液氮温区实现超导的材料。

这些高温超导氧化物有一些共同的特点。例如都有变形钙钛矿原胞的层状

堆砌结构, 正常相下都具有金属性质, 载流子是空穴。在采用一般陶瓷的烧结工艺制备时, 制备的材料一般属于多晶状态烧结体, 这种烧结超导材料有许多晶界, 致密度也不高, 往往造成晶粒间的弱连接, 致使材料的临界电流不高, 其力学性能也不好。后来发展了一种熔融织构 (MTG) 制备方法, 能使样品达到很高的密度, 晶界连接也大大改善, 可以获得质量较好的超导材料。

研究证明, 金属氧化物超导体不仅临界温度高, 上临界场 H_{c2} 也很高。例如钇系的 H_{c2} 在 4.2K 的温度时约相当于 Nb_3Sn 的 10 倍, 所以可作为强磁场材料, 即使在液氮中 (77K), 钇系 123 相的 H_{c2} 也远高于任何一种低温超导体。

在 77K, 高温超导材料短样品的临界电流密度已超过实用化要求的最低值 (约 $10^3 A/cm^2$)。例如钇系 123 (Y123) 相在 9T 的磁场中 $J_c = 10^5 A/cm^2$, 铋系 2223 相和铌系 2223 相带材在无磁场时 J_c 也达到了 $1.2 \times 10^5 A/cm^2$ 。对于小样品, 如高温超导薄膜和微桥, 其 J_c 则比上述数字普遍高 10 倍以上, 最高可达 $10^6 A/cm^2$ 。因此可以说, 高温超导体的各项临界参数都优于低温超导体。Y123 相超导体不仅 J_c 很高, 其他临界参数也很高, 是一种

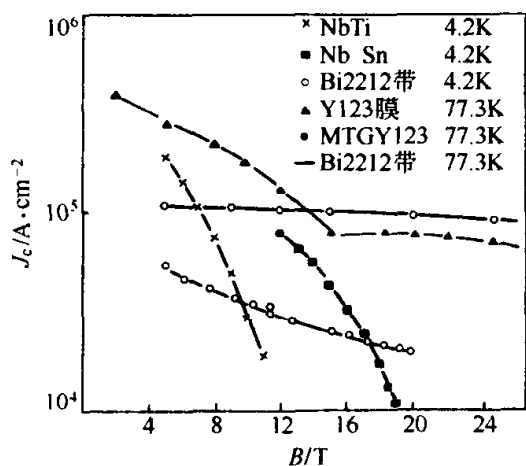


图 12-22 几种超导材料性能比较

一种非常有开发潜力的高温超导材料。如果在液氮温度下使用, 它们的临界电流密度在强磁场中, 例如在 20T 中大大超过任何一种已有的低温超导材料。银包套的 Bi2212 带材 (4.2K), Y123 薄膜 (77K) 和 MTGY123 以及几种低温超导材料的临界电流密度随磁场的变化参见图 12-22。由于它们的实用潜力很大, 高温超导材料实用化的研究开发工作在许多国家都在积极进行。

从目前情况看, 钇系容易形成单一的超导相, 制备短样相对容易。但其超导性质对氧含量特别敏感, 这一点会造成包套法加工带材时充氧的困难。已经试验研究开发的方法都设法避开这个缺点, 例如利用各种化学气相沉积 (CVD) 方法, 包括金属有机物化学气相沉积 (MOCVD) 方法来研究开发钇系 123 相带材的工作一直在一些发达国家的研究组进行着, 并且取得了不少的进展。这种方法的基本想法是以强度好的薄不锈钢带作为基带, 用 CVD 方法在基带上镀上 Y123 相薄膜, 薄膜厚度是微米数量级。让超导薄膜实现超导电性以后在最外面还镀上一层良导体金属材料, 以达到稳定化和保护的目, 这种工艺和制备 Nb_3Sn 带

材相同。利用各种熔融织构法制备大块的 Y123 材料是目前高温超导材料应用的最重要的方面之一。其不但能制备出尺寸以厘米计算的 Y123 块体,而且性能要大大超过烧结样品。因为 Y123 的临界参数很高,MTG 块材又能基本达到这些高临界参数,所以应用的潜力很大。据报道,用熔融织构块材作永磁材料在 40K 左右就能做成场强约为 1.5T 的超导磁体。Y123 相的临界磁场、临界电流等都很高,永久超导磁体的场强应当可以大大提高。特别有潜力的是 30K 温度的超导磁体已经不必要使用液氮冷却,完全可以使用小型制冷机来维持温度不变。根据迈斯纳效应,用这样的材料可以做成无摩擦的磁悬浮轴承,磁悬浮列车,磁悬浮储能器等等。我国试制的、世界上第一列高温超导磁悬浮列车已经进行过载人试验:列车导轨是钕铁硼永久磁体,车上用的 340 块 MTG 块材 Y123 是我国研制的,迈斯纳效应使列车悬浮在导轨上方。鉴于各种磁体用途的广泛性,熔融织构块材 Y123 的研究发展工作受到高度重视。

铋系存在两个高温超导相,要得到单一的 110K 超导相即 Bi2223 相和 90K 的 Bi2212 相都不困难且相的性能稳定。但其缺点是在液氮温区中 J_c 值较低,对磁场也很敏感。但是随着温度的降低,其临界电流就迅速地提高。到大约 30K 时它们的临界电流密度已经很高,对磁场的敏感程度也大大地降低,能够满足有些实际应用的要求,而且它们的临界磁场在 30K 附近比典型的低温超导体如 NbTi 线在 4K 的要高得多。

Bi2223 和 Bi2212 相的性能稳定,尤其是它们的超导电性不会受氧含量问题的困扰,这就为包套带材的加工成型带来方便,可以用管子包套粉末方法来制备长带材或线材。现在使用的都是银管中装上 Bi2223 或 Bi2212 的粉末。封好后经适当的机械加工和热处理,以达到所希望的优良的超导带材或线材。前已指出,Kunzler 等在 20 世纪 60 年代证明 Nb_3Sn 能在强磁场中传输大电流就使用了这种方法。据报道,用这种方法已经能拉出数千米长的带材。这种带材是多晶体,和烧结样品差不多,其优点不仅是容易制成长样品,而且机械性能好,比如说可以允许一定的弯曲,利用这种方法生产的铋系高温超导材料已经逐步地进入了应用阶段。人们期望用 Bi 系材料来制作磁体,电力输送电缆,电流保护器,超导储能器,超导电机等。

铊系材料性能与铋、铋两体系相比, T_c 是最高的,但也高出不多,不过铊是有毒元素,规模生产时安全问题会造成一定困难。

高温超导体实用化的重大障碍目前不是 T_c 低,而是其机械加工难度太大,用目前的技术难以制备成长线。高温超导体都是陶瓷,不能像低温超导体那样能用金属工艺成材,没有足够的韧性和强度,也不便于使用。目前,用高温超导体做成线材后,若超导性能较好(临界电流密度 J_c 等较高),则

只有厘米级的长度，做到 1m 长的材料的 J_c 已不很高了。所以高温超导应用的关键在于解决成材工艺。就目前看，铋系最有可能在近期内被开发为实用长线材料。

在弱电应用方面，已成功地开发出多种制备高温超导膜的方法，其中激光熔融法制备的钇钡铜氧膜不仅质量好，而且制备速度快。含有约瑟夫森结的回路可制成超导量子干涉器 (SQUID)，并且大角度晶界就是约瑟夫森结。用钇钡铜氧膜中长一个大角度晶界的方法已制成了 SQUID，实验结果表明，这种方法制成的 SQUID 能工作到 87K。这说明，弱电应用关键的超导材料制备技术已取得了重大进展。

镧、钇、铋、铊等系列高温超导体仅为已发现的高温超导体的一部分。现今各种高温超导体的数目虽难以统计，但可分为含铜的一类和不含铜的一类。含铜的一类又分成 p 型的（空穴导电）和 n 型的（电子导电）两类。已提及的都是 p 型的， T_c 较高，所以正在开发成有用的超导材料。n 型含铜氧化物超导体的以铋铊铜氧为代表，其 T_c 在 30K 以下，其性质和 p 型没有很大的差别。不含铜的氧化物超导体以钡钾铋氧为代表的 T_c 约为 30K，2001 年发现的 MgB_2 的 $T_c \approx 40K$ ，并且其物理性质与常规的低温超导体相近。

高温超导体除了氧化物陶瓷外，还有一类是有机超导体，越来越受到重视。有机超导体是有机物，其基本特征是含有碳氢等原子，能构成长链或环状结构的有机体。有机超导体早就受到科学家的特别重视，因为一些超导理论预言，某种具有链状结构的有机物或层状结构物质会具有高的临界温度，而且实际上氧化物高温超导体确是层状结构。1991 年 7 月发现，一种叫碳 60 (C_{60}) 的有机物，掺卤碱金属杂质 (K 等) 以后，其 T_c 已超过 30K，比已发现的其他有机超导体高得多。2000 年将 C_{60} 单晶体用场效应技术注入空穴载流子，将 T_c 提高到了 50K 以上。2001 年，已有关于 C_{60}/BrH_3 的超导临界温度达到 117K 的报道，而且没有什么迹象表明有机超导体和其他超导体的临界温度已经不能再高了。从技术上看，有机超导体的研究也是一个吸引人的领域，因为对有机物可以用“分子工程”的办法改进或创新，具有很大的竞争力。

参 考 文 献

- 1 De Gennes P. G. .Superconductivity of Metals and Alloys. W A Benjamin Inc. 1966. 刘长富等译. 北京: 科学出版社, 1980
- 2 Tinkham M., Introduction to Superconductivity. McGraw-Hill. 2nd ed. 1996. 邵惠民等译. 北京: 科学出版社, 1985
- 3 丁世英. 超导电性. 金属物理第四卷. 第 13 篇. 冯端等著. 北京: 科学出版社, 1996
- 4 张其瑞主编. 高温超导电性. 杭州: 浙江大学出版社, 1992
- 5 周午纵, 梁维耀主编. 高温超导基础研究. 上海: 上海科学技术出版社, 1999

第 13 章 光学性质

撰稿人 王慧田

关于材料的光学特性，最早可以追溯到 4000 年前古老中国的青铜器时代，当时人们已经知道，通过材料的光泽和色彩(颜色) 去确定铜合金的成分。对材料的光学性质有了初步的认识。现在，对颜色已经有了充分了解，红、绿、蓝作为三基色，任何其他颜色都可以通过这三者的适当配比可以得到。对材料的光学性质有了较深入的理解和认识。在此，将从完全不同的角度，科学地去研究材料的光学特性。通过研究诸如折射率、吸收系数和光激发以及光谱等，深刻理解其光学特性，并用于材料的表征。

20 世纪初，光与物质相互作用的研究(黑体辐射) 奠定了量子理论的基础。时至今日，光学方法对于材料的表征是极为重要和不可缺少的工具或技术，可获得采用其他技术难以获得的信息。诸如激光器、光探测器、光波导、近场光学显微镜、光子晶体等的重大技术进步，在光通信、太阳能利用、医学诊断、材料性能和结构表征以及其他光子与光电子领域发挥了极为重要的作用。

13.1 光波在线性介质中的传播

在本节将描述光作为一种电磁波在各向同性、各向异性、介质平面光波导、光纤等线性介质中的传播行为和规律。并在本节的最后，概要地介绍光子晶体。光波的传播行为以及和物质的相互作用遵从麦克斯韦方程

$$\begin{aligned}\nabla \times \mathbf{E} &= -\partial \mathbf{B} / \partial t \\ \nabla \times \mathbf{H} &= \partial \mathbf{D} / \partial t + \mathbf{j} \\ \nabla \cdot \mathbf{D} &= \rho \\ \nabla \cdot \mathbf{B} &= 0\end{aligned}\tag{13-1}$$

其中 \mathbf{E} 、 \mathbf{H} 、 \mathbf{D} 、 \mathbf{B} 分别为电场、磁场、电位移以及磁感应强度矢量， \mathbf{j} 、 ρ 则分别为电流和电荷密度。麦克斯韦方程组的前两个方程表明：变化的磁场可以激发电场，而变化的电场也可以激发磁场。后两个方程说明电场是有源的，磁场则是无源的。方程组(13-1) 中的四个方程不是完全独立的，所以问题的解不能被惟一地确定，为此还需物质方程

$$\mathbf{D} = \epsilon \epsilon_0 \mathbf{E} = \epsilon_0 \mathbf{E} + \mathbf{P}, \quad \mathbf{B} = \mu \mu_0 \mathbf{H}\tag{13-2}$$

以及电流密度方程和极化强度方程

$$\mathbf{j} = \sigma \mathbf{E}, \quad \mathbf{P} = \chi \epsilon_0 \mathbf{E} \quad (13-3)$$

式中, ϵ 、 μ 为材料的二阶相对介电常数和导磁率张量, σ 为电导率, χ 为二阶非线性极化率张量, ϵ_0 和 μ_0 为真空中介电常数和导磁率张量。从上述方程组可得到描述光波传播的波动方程

$$\nabla \times (\nabla \times \mathbf{E}) + \mu_0 \epsilon_0 \partial^2 \mathbf{E} / \partial t^2 = -\mu_0 \partial^2 \mathbf{P} / \partial t^2 - \mu_0 \partial \mathbf{j} / \partial t \quad (13-4)$$

13.1.1 光波在均匀的各向同性电介质中的传播

对均匀的各向同性电介质, 二阶张量均退化为标量, 且是无源的 ($\rho = 0$) 和非磁的 ($\mu = 1$), 还有 $\mathbf{j} = 0$ 。考虑介质的宏观极化强度 \mathbf{P} 的色散, \mathbf{P} 与电场强度 \mathbf{E} 有如下关系

$$\mathbf{P} = \sum_j \frac{N_j e^2 / m}{\omega_j^2 - i \delta_j \omega - \omega^2} \mathbf{E} = \chi \epsilon_0 \mathbf{E} \quad (13-5)$$

式中, N_j 为具有本征频率 ω_j 束缚态内的电子数密度, δ_j 为阻尼系数, e 和 m 分别为电子电荷和质量。 χ 可写为复数形式 $\chi = \chi_R + i \chi_I$ 。根据方程 (13-5) 式, χ 的实部和虚部分别为

$$\chi_R = \frac{e^2}{m \epsilon_0} \sum_j \frac{N_j (\omega_j^2 - \omega^2)}{(\omega_j^2 - \omega^2)^2 + (\delta_j \omega)^2}, \quad \chi_I = \frac{e^2}{m \epsilon_0} \sum_j \frac{N_j \delta_j \omega}{(\omega_j^2 - \omega^2)^2 + (\delta_j \omega)^2} \quad (13-6)$$

那么, 从式(13-4)~式(13-6)可以得到在均匀的各向同性电介质中, 关于 \mathbf{E} 的简单方程

$$\nabla^2 \mathbf{E} = \mu_0 \epsilon_0 (1 + \chi_R + i \chi_I) \partial^2 \mathbf{E} / \partial t^2 \quad (13-7)$$

该方程的解为单色平面波形式

$$\mathbf{E} = \mathbf{E}_0 e^{-\alpha z/2} \exp(i \omega t - i k z) \quad (13-8)$$

式中, k ($= n \omega / c = 2 \pi n / \lambda$) 为光波在均匀的各向同性介质中的传播常数, ω 光波的圆频率、 λ 真空中光波波长和 n 介质的折射率; $c = (\epsilon_0 \mu_0)^{-1/2}$ 真空中的光速。 α 则被称为吸收系数, 表示光波在介质中单位传播距离光强衰减的百分比。折射率 n 和吸收系数 α 为

$$n^2 = \frac{1}{2} \left[\sqrt{(1 + \chi_R)^2 + \chi_I^2} + (1 + \chi_R) \right] \approx (1 + \chi_R) \quad (13-9)$$

$$\alpha^2 = 2 \frac{\omega^2}{c^2} \left[\sqrt{(1 + \chi_R)^2 + \chi_I^2} - (1 + \chi_R) \right] \approx 2 \frac{\omega^2}{c^2} \chi_I \quad (13-10)$$

当光波频率 $\omega = \omega_j$ (即光频与物质系统某一束缚态的本征频率 ω_j 共振) 时, 从式(13-6)可知, $\chi_R \rightarrow 0$; 而 χ_I 达到极大, 当然吸收系数 α 也达极大。从图 13-1 可以看到一系列的 $\omega = \omega_j$ 处吸收出现峰值。在光频 ω 远离某一本征频率 ω_j 处, χ_I 趋于零, 因此 α 也达到零。从式(13-6)和式(13-9),

可以发现材料的折射率为光波场频率的函数, 此现象被称为折射率色散。在 $\omega < \omega_j$ 区域, 折射率 $n > 1$, 而且当 ω 趋于 ω_j 时, n 增大; 这一区域属于正常色散范围。相反, 在 ω_j 附近, 即近共振区域, 折射率 n 随频率的增加反而降低; 属于反常色散区域, 甚至有时 $n < 1$ 。对于折射率 $n < 1$ 的情况, 在光子局域化研究领域是引人注目的热点问题。此外, 在整个可见光谱区,

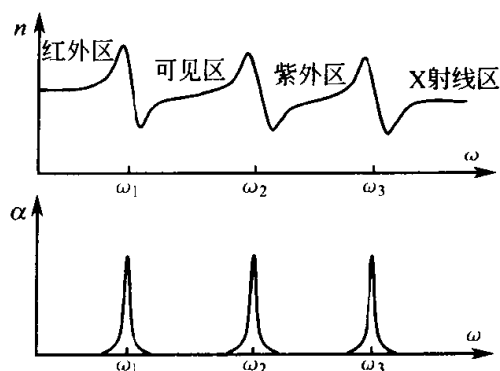


图 13-1 色散与吸收曲线

绝大多数的材料均呈现正常色散。折射率和吸收系数不仅对均匀的各向同性介质, 而且对所有光学材料均为两个极为重要的、最基本的光学参数。

13.1.2 光波在导体中的传播

光波在导体中的传播行为仍然遵从麦克斯韦方程 [式(13-1)] 或波动方程 [式(13-4)]。由于导体中电子是自由电子, 必须考虑在光波的交变电场作用下电子的真实情况。导体中的电流密度应为 $j = \sigma E / (1 - i\omega\tau)$, 而非式(13-3)所给出的简单形式, 式中电导率 $\sigma = Ne^2\tau/m$, N 为导体中的电子数密度, τ 为弛豫时间。当 $\omega = 0$ 时, 导体中的电流密度退变为式(13-3)的简单形式。将导体中的电流密度代入波动方程式(13-4), 并置 $P = 0$, 经整理后得

$$\nabla^2 E = \mu_0 \epsilon_0 \partial^2 E / \partial t^2 + \sigma \mu_0 (1 - i\omega\tau)^{-1} \partial E / \partial t \quad (13-11)$$

同样, 该方程具有与式(13-8)相同形式的解。可以容易得到折射率 n 和吸收系数 α

$$n^2 = \frac{1}{2} (\sqrt{C_1^2 + C_2^2} + C_1), \alpha^2 = 2 \frac{\omega^2}{c^2} (\sqrt{C_1^2 + C_2^2} - C_1) \quad (13-12)$$

其中 $C_1 = 1 - \sigma\tau/\epsilon_0(1 + \omega^2\tau^2)$, $C_2 = \sigma/c(1 + \omega^2\tau^2)$ 。在导体中, 通常定义吸收系数 α 的倒数为穿透深度或趋肤深度 $W = \alpha^{-1}$, 表示光波在导体中的穿透能力。在光波频段 $\lambda \sim 1\mu\text{m}$ ($\omega \sim 1.9 \times 10^{15}\text{Hz}$), 如金属铜的 $\sigma \sim 5.8 \times 10^7 (\Omega \cdot \text{m})^{-1}$, $\tau \sim 10^{-13}\text{s}$, 从上面公式得到 $\alpha \sim 740\mu\text{m}^{-1}$ 。那么, 光波在金属铜中 $W \sim 0.0014\mu\text{m}$, 同时也可估算得到 $n \rightarrow 0$ 。可见, 光波在导体中的穿透深度非常之浅, 而折射率近趋于零, 这正是导体中光波不透明的原因; 由于在光波波段金属的折射率很低, 可望在光子局域化和光子晶体中获得应用。

13.1.3 光波在各向异性介质中的传播

通常, 光波在各向异性介质中传播的情况更普遍, 因为具有光学和电光功能等性质的晶体绝大多数都是各向异性的, 所以讨论光波在各向异性介质

中的传播更有实际的意义。各向异性是指晶态物质的光学性质上的各向异性，且决定于晶体的空间对称性。在各向异性介质中，其介电常数是一个二阶张量，不再是一个标量。在主轴坐标系中，介电常数 ϵ 张量具有如下对角矩阵形式

$$\epsilon = \begin{bmatrix} \epsilon_x & 0 & 0 \\ 0 & \epsilon_y & 0 \\ 0 & 0 & \epsilon_z \end{bmatrix} = \begin{bmatrix} n_x^2 & 0 & 0 \\ 0 & n_y^2 & 0 \\ 0 & 0 & n_z^2 \end{bmatrix} \quad (13-13)$$

其中 ϵ_j 和 n_j ($j = x, y, z$) 则分别为主相对介电常数和主折射率。

将式(13-13)和式(13-2)及 $j=0$ 代入波动方程式(13-4)，在各向异性介质中则有

$$\nabla \times (\nabla \times \mathbf{E}) + \epsilon \epsilon_0 \mu_0 \partial^2 \mathbf{E} / \partial t^2 = 0 \quad (13-14)$$

该方程仍然存在平面波形式的解，即

$$\mathbf{E} = \mathbf{E}_0 \exp[i(\omega t - \mathbf{k} \cdot \mathbf{r})] \quad (13-15)$$

需要注意的是：在各向异性介质中，即使介质是均匀的， $\nabla \cdot \mathbf{E} = 0$ 也不成立。将式(13-15)代入式(13-14)可得

$$\mathbf{k} \times (\mathbf{k} \times \mathbf{E}_0) + \epsilon \epsilon_0 \mu_0 \omega^2 \mathbf{E}_0 = 0 \quad (13-16)$$

这是关于 \mathbf{E}_0 的三个未知分量 E_{0x} , E_{0y} , E_{0z} 的齐次一阶方程组。用式(13-13)计算式(13-16)中的矢积，得到如下行列式方程

$$\begin{bmatrix} \omega^2 n_x^2 / c^2 - k_y^2 - k_z^2 & k_x k_y & k_x k_z \\ k_x k_y & \omega^2 n_y^2 / c^2 - k_x^2 - k_z^2 & k_y k_z \\ k_x k_z & k_y k_z & \omega^2 n_z^2 / c^2 - k_x^2 - k_y^2 \end{bmatrix} \begin{bmatrix} E_{0x} \\ E_{0y} \\ E_{0z} \end{bmatrix} = 0 \quad (13-17)$$

在此， k_x , k_y , k_z 为波矢 \mathbf{k} 在三个主轴方向的分量。当行列式的值为零时，存在特解。从这一条件可以得到如下的圆频率 ω 与波矢分量的关系式

$$\begin{vmatrix} \omega^2 n_x^2 / c^2 - k_y^2 - k_z^2 & k_x k_y & k_x k_z \\ k_x k_y & \omega^2 n_y^2 / c^2 - k_x^2 - k_z^2 & k_y k_z \\ k_x k_z & k_y k_z & \omega^2 n_z^2 / c^2 - k_x^2 - k_y^2 \end{vmatrix} = 0 \quad (13-18)$$

对于给定的圆频率 ω 和主折射率 n_x , n_y , n_z ，该方程表示的是波矢 \mathbf{k} 空间的一个三维曲面，该曲面被称为波法线面或波矢面，由内外两层曲面构成的双重曲面，而且两个曲面均为封闭的。通常，三个主折射率 n_x , n_y , n_z 均不同，在波法线上存在四个交点。通过原点和其中关于原点对称的两个交点可以画出两条直线，它们所表示的方向称为光轴，即此时有两个光轴，这类晶体被称之为双轴晶体。实际上，有许多晶体，它们的其中两个主折射率相

等, 这样就只有一个光轴, 这类晶体称为单轴晶体。当然, 三个主折射率都相同时, 各向异性消失, 退化为各向同性。

从原点并沿光波的传播方向画一直线, 通常此直线与波法线面有两个交点, 即沿传播方向存在两个可能的 k 值, 对应于两个不同的相速度 ω/k 。 k 可以表示为 $k = n(\omega/c)s$, s 为传播方向的单位矢量, n 为要确定的折射率。将 k 的表达式代入式 (13-18), 整理可得

$$\frac{s_x^2}{n^2 - n_x^2} + \frac{s_y^2}{n^2 - n_y^2} + \frac{s_z^2}{n^2 - n_z^2} = \frac{1}{n^2} \quad (13-19)$$

其中 s_j ($j = x, y, z$) 为单位矢量 s 在主轴 j 方向的分量。这就是所谓的 Fresnel 方程。这是关于 n^2 的 2 次方程, 所以对一个传播方向 (s_x, s_y, s_z), 将有两个不同的 n^2 , 此即所谓的双折射效应。令 Fresnel 方程的解为 n_1^2, n_2^2 , 将 k 和 n^2 的两个解 n_1^2, n_2^2 分别代入式 (13-17), 可以得到相应电场矢量的两个解 E_1 和 E_2 , 那么有

$$E_{1,2} = \left[\frac{s_x}{n_{1,2}^2 - n_x^2}, \frac{s_y}{n_{1,2}^2 - n_y^2}, \frac{s_z}{n_{1,2}^2 - n_z^2} \right] \quad (13-20)$$

至此, 已得到了传播本征模的一般解。但作为特殊情况, 研究光波在主轴坐标面内的传播是有意义的。考虑在 yz 面内的传播 (即 $k_x = 0$), 此时本征值方程式 (13-18) 可以写为

$$\left[\frac{k_y^2}{n_x^2} + \frac{k_z^2}{n_x^2} - \frac{\omega^2}{c^2} \right] \left[\frac{k_y^2}{n_z^2} + \frac{k_z^2}{n_y^2} - \frac{\omega^2}{c^2} \right] = 0 \quad (13-21)$$

该式表示波矢空间的 yz 平面上的一个圆和一个椭圆。圆对应于一个本征模, 其本征折射率为 n_x , 相应的电场矢量为 $E_1 = (1, 0, 0)$, 此电场与光波的传播面 (yz 面) 正交。椭圆对应于另一个本征模, 其本征折射率 n 由 $1/n^2 = s_y^2/n_z^2 + s_z^2/n_y^2$ 给出, 相应的电场矢量为 $E_2 = (0, s_z/n_y^2, -s_y/n_z^2)$, 此电场位于含有波矢的 yz 面内。概括地讲, 在坐标面内传播的光波, 存在两种模式: 其一, 电场与传播方向正交且与另一主轴平行, 其本征折射率与方向无关等于主折射率。另一模式的电场与传播方向和另一主轴均正交。在其他两个坐标平面内的传播也同样。图 13-2 示出了坐标面与波法线面的交线

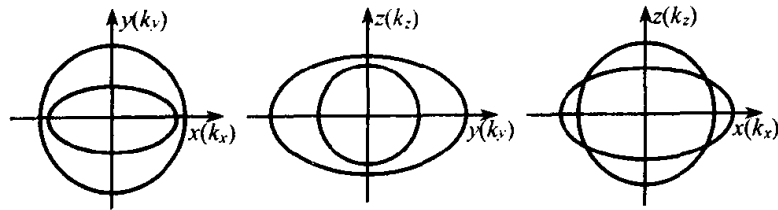


图 13-2 坐标面与波法线面的交线

线, 当 $n_x < n_y < n_z$ 时, 圆和椭圆只在 xz 面相交。

(1) 单轴晶体 各向异性介质可以分为三类。主折射率均不同的情况, 存在两个光轴, 称为双轴晶体 (如三斜晶系, 单斜晶系, 斜方晶系)。在许多晶体中, 有 2 个主折射率相同, 此类晶体只有一个光轴, 称之为单轴晶体 (如正方晶系, 六方晶系和三角晶系); 此时, 波法线面由一个球面和一个旋转椭球面构成, 这两个面只有两个交点, 惟一的光轴可令其为 z 轴。在单轴晶体中, 沿光轴 (z 轴) 方向的本征折射率 n_o , x 和 y 方向的本征折射率 n_e , 两者被称为寻常和异常折射率。尽管传播方向不同, 但有相同的本征折射率 n_o , 此类传播模称为寻常光 (o 光)。本征折射率依赖于传播方向的传播模称为异常光 (e 光)。三个主折射率都相同时, 这样的晶体在光学上是各向同性的 (如立方晶系)。

不失一般性, 可以考查 yz 面内的传播。假设波矢与光轴 (z 轴) 的夹角为 θ , o 光的本征折射率为常数 n_o , 相应的电场矢量为 $E_o = (1, 0, 0)$ 。由式 (13-19) 和式 (13-20) 可得 e 光的本征折射率和电场为

$$1/n_e^2(\theta) = \sin^2\theta/n_e^2 + \cos^2\theta/n_o^2 \quad (13-22)$$

$$E_e = (0, \cos\theta/n_o^2, -\sin\theta/n_e^2) \quad (13-23)$$

我们发现, o 光的电场矢量总是垂直于光轴和波矢构成的平面, 而 e 光的电场矢量总是位于光轴和波矢构成的平面内。

(2) 折射率椭球 折射率椭球是分析光波在晶体中传播非常有效且简明直观的几何方法。可以直接地从几何图形上, 推断光波在晶体的各方向上的传播模式以及相关各矢量之间的关系。在主轴坐标系中, 折射率椭球方程可以写成

$$\frac{x^2}{n_x^2} + \frac{y^2}{n_y^2} + \frac{z^2}{n_z^2} = 1 \quad (13-24)$$

显然, 椭球的三个半轴分别为 n_x , n_y , n_z 。如图 13-3, 从坐标原点沿光波传播方向引出单位矢量 s , 同时通过原点并垂直于 s 作一个平面, 该平面与折射率椭球相交, 截得一个椭圆, 此椭圆的长半轴和短半轴所指的方向即为允许的两个本征模式的电位移矢量 D_1 和 D_2 的偏振方向, 而长半轴和短半轴的长度则为相应模式的折射率 n_1 和 n_2 。若为单轴晶体, 即 $n_x = n_y = n_o$, $n_z = n_e$, 折射率椭球将退化为关于 z 轴的旋转椭球, 总有一个本征模 (即 o 光偏振) 的折射率为 n_o , 另一个本征模的折射率由式 (13-22) 给出。

(3) 能流 波矢 k 表示光波的波前传播方向, 而玻印亭矢量 $S = E \times H$

代表能流的传播方向。光波在均匀的各向同性介质中传播，波矢 \mathbf{k} 与玻印亭矢量 \mathbf{S} 平行，即光波的波矢与能流方向相一致。但是，在各向异性介质中情况就截然不同了。从式(13-16)可知，显然 \mathbf{k} 与 \mathbf{E} 不再垂直，但与电位移矢量 \mathbf{D} 垂直。从麦克斯韦方程可得： $\mu_0 \omega \mathbf{H} = \mathbf{k} \times \mathbf{E}$ ，即 \mathbf{H} 与 \mathbf{k} 和 \mathbf{E} 两者同时垂直，显然 \mathbf{S} 与 \mathbf{k} 不平行。说明在各向异性介质中光波的波矢与能流方向通常不一致。概括起来，在非磁的各向异性介质中的光波传播特征为：四个矢量 \mathbf{E} 、 \mathbf{D} 、 \mathbf{S} 和 \mathbf{k} 共面，且均垂直于 \mathbf{H} ；同时， \mathbf{k} 垂直于 \mathbf{D} 、 \mathbf{S} 垂直于 \mathbf{E} ， \mathbf{D} 不平行于 \mathbf{E} 、 \mathbf{S} 也不平行于 \mathbf{k} 。

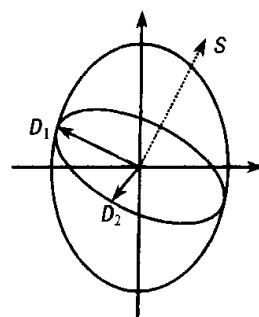


图 13-3 折射率椭球

(4) 液晶 凝聚态物质有两种极端形态：均匀的各向同性的液态和结晶固态。液态是短程有序长程无序的，具有最高的可能对称性；晶态具有长程的位置和取向有序，以及最低的可能对称性。在这两种极端形态之间，存在一些方向短程相关，而在其他方向长程有序的一种物质形态，其对称性介于液态与晶态之间。目前，液晶化合物已有几千种，依据其排列方式可分为近晶相、向列相、胆甾相和柱状相；前三种是经典的、常见的液晶相结构。

近晶相液晶的分子呈棒状、分层叠合，每层分子长轴相互平行，且与层面垂直；各层间距可以改变，单个层中分子只能在本层中活动。

向列相液晶即线型液晶，分子也呈棒状、分子长轴也相互平行，但不分层；具有僵硬棒状分子形态的化合物均呈向列相液晶形态。分子的长轴方向即为光轴，且为正单轴晶体， $n_e > n_o$ 。通常向列相液晶的双折射比较显著的， $\Delta n = n_e - n_o > 0.1$ ，因材料的温度而异。向列相液晶中的光波传播完全可以按上面描述的单轴晶体的情形处理。

胆甾相液晶中分子也是分层排列的、逐层叠合、每层中分子长轴相互平行且与层面平行；相邻两层间分子长轴逐层依次沿一定方向有一微小扭角，因此各层分子长轴的排列方向就逐渐扭转成螺旋状。该类液晶的光轴垂直层面而平行于螺旋轴，且为负单轴晶体。若将该类液晶夹在两个平行的玻璃板中间，在白光下当从不同方向观察时，会观察到不同颜色的反射光，此即选择反射性，这是由于液晶分子结构呈螺旋形排列的缘故。这种反射非常类似于晶格中的布拉格反射。进一步研究表明：透射光的颜色也因方向而异，同时反射与透射光不仅颜色互补而且两者均为圆偏光；所以选择反射性有被称为圆偏振二色性。

柱状相液晶是近年来发现的一种新的液晶结构。目前主要两种分子可以

形成柱状相：双亲分子和含有长侧链的盘状分子。柱状相均以胶束结构形式存在，通常呈现等六边形的柱状外形结构。

此外，液晶还具有旋光、电光和磁光等特性。

13.1.4 光波在平面介质薄膜波导中的传播

光波可以被约束在确定的导波介质层中传播，在此，仅讨论光波导中的线性传播问题，不涉及非线性，如调制、频率变换等。光波导的厚度约为 $1 \sim 10 \mu\text{m}$ ，且呈薄膜或带状，分别称为薄膜波导或带状波导，在此仅讨论薄膜波导。光波导是由三种不同折射率的介质构成的层状组合结构。用下标 $j=1, 2, 3$ 分别表示构成波导的三种介质——覆盖层、薄膜和衬底；相应的折射率为 n_j 。对波导结构要求三种介质的折射率满足 $n_1 \leq n_3 < n_2$ ，这样光波才可能被限制在薄膜内长距离传播，而不泄漏到周围的两种介质（衬底和覆盖层）中。研究光波在波导中的传播，一般是求解满足边界条件的麦克斯韦方程组，但是在此将采用简单、直观的射线光学和光波相干叠加的原理去处理此问题。

(1) 薄膜波导的模方程 如图 13-4，假定构成波导的介质均为无源的、非磁性的、均匀的、各向同性的。薄膜厚度为 W ，而覆盖层和衬底被假定为半无限大。考虑在薄膜内波长为 λ 的线偏振单色平面光波 A_2 ，它以入射角 θ_2 传播至薄膜-衬底界面（即 $z=0$ 面），假定所有光场均与 y 坐标无关的。若 θ_2 大于 n_2 和 n_3 间的临界角[即 $\theta_2 > \sin^{-1}(n_3/n_2)$]，光波 A_2 被全内反射进入光波 B_2 ，如图 13-4 (a)。类似地，如图 13-4 (b)，光波 B_2 也将在薄膜-覆盖层界面（即 $z=W$ 面）被全反射进入光波 A'_2 。根据电磁波的连续性条件，三种介质中的光波场中应具有相同的 x 依赖性，即 $\exp(i\kappa\beta x)$ 。其中， $\kappa=2\pi/\lambda$ 为光波在真空中的传播常数。无量纲量 $\beta=n_2\sin\theta_2$ 是以 κ 为单位的光波场波矢的 x 分量，是支配光波在波导中传播行为的重要物理量。那么，在第 j 种介质中，光波场在 z 方向的变化遵循 $\exp(\pm i\kappa b_j z)$ ，且有 $b_j=(n_j^2-\beta^2)^{1/2}$ 。确保光波在波导内传播而不泄漏，必须限定 $n_3 \leq \beta \leq n_2$ 。这样，光波在薄膜内以均匀波(体波)的形式传播（因为 b_2 是正的实数）；而在覆盖层和衬底域，因为 b_1 和 b_3 为正的虚数，光波则表现为消失波（表面波）。

显然，光波 A_2 和 A'_2 具有相同的传播因子 $\exp(-i\omega t - i\kappa b_2 z + i\kappa\beta x)$ 。类似地，光波 B_2 应具有传播因子 $\exp(-i\omega t + i\kappa b_2 z + i\kappa\beta x)$ 。现将图 13-4 (a) 和 13-4 (b) 合并成图 13-4 (c)，其中光波 A'_2 为光波 A_2 在薄膜内沿一个 z 路径传播后的光波。因为在薄膜的两个界面的反射均为全内反射，所以 A'_2 和 A_2 的振幅仅差一个位相因子 Δ 。后续的 z 路径的光波 $A''_2, A'''_2, A''''_2, \dots$ ，相对于 A_2 的位相差依次为 $2\Delta, 3\Delta, 4\Delta, \dots$ 。光波 A_2 在

位置 $x = x_c$ 和 $z = 0$ 的位相为 $-\omega t + \kappa\beta x_c$ 。光波 A'_2 在相同位置的位相是 A_2 在 $x = x_a$ 和 $z = 0$ 处的位相加上一个 z 路径的相移为 $-\omega t + \kappa\beta x_a + \kappa\beta (x_c - x_a) + 2\kappa b_2 W + 2\phi_{21} + 2\phi_{23}$ 。这样, Δ 则为上述两个位相之差, $\Delta = 2\kappa b_2 W + 2\phi_{21} + 2\phi_{23}$ 。通常, 波簇 $A_2, A'_2, A''_2, A'''_2, A''''_2, \dots$ 的叠加场为零, 只有当 $\Delta = 2m\pi$ (即相长干涉) 波导薄膜内的叠加场才不为零, 即

$$2\kappa b_2 W + 2\phi_{21} + 2\phi_{23} = 2m\pi \quad (13-25)$$

此方程称为波导的模方程。其中, ϕ_{21} 和 ϕ_{23} 分别为光波在薄膜的上下界面的全内反射相移, 由 $\phi_{jk} = -\arctan(p_j b_k / i p_k b_j)$ 确定, 其范围为 $0 < \phi_{jk} < \pi/2$; 在此, p_j 为在第 j 种介质中的偏振依赖因子, 对 TE 偏振 $p_j = 1$, 对 TM 偏振 $p_j = n_j^2$ 。

从模方程可以看出, 波导中允许传播模 (称为导模) 只是一系列具有分立的传播常数为 β_m 的波, m 为模的阶数。对于给定的膜厚和光波偏振, 传播常数 β_m 完全由模方程决定。

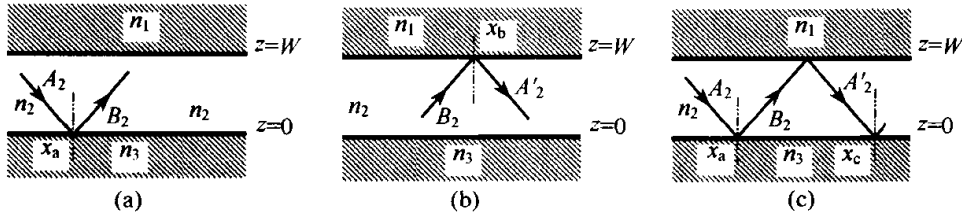


图 13-4 光在薄膜波导中的传播

(a) 薄膜下边界的全内反射; (b) 薄膜上边界的全内反射; (c) 在薄膜波导中的光波传播

(2) 薄膜波导的模 对波导的模方程式 (13-25) 可以被重新改写为

$$W = W_{23} + W_{21} + mW_2 \quad (13-26)$$

$$W_{23} = \phi_{23}/\kappa b_2, \quad W_{21} = \phi_{21}/\kappa b_2, \quad W_2 = \pi/\kappa b_2 \quad (13-27)$$

式 (13-26) 表明, 可以从两个对称波导 ($n_1 = n_3$ 和 $n'_1 = n'_3$) 构筑一个具有期望的传播特性的非对称波导 ($n_1 \neq n_3$), 这两个对称波导的特征折射率为 $n_2 = n'_2$ 。例如, 图13-5(a)和(b)示出两个对称波导, 它们分别可以承载相同传播常数 β 的 $m = 0$ 模和 $m' = 2$ 模。将图13-5(a)的下半部与图13-5(b)的上半部组合在一起, 可以获得一个非对称波导, 它具有相同的传播常数 β 且阶数为 $m'' = (m + m')/2 = 1$, 如图13-5(c)。当然, 按相同的方法, 可以得到任意的 m 与 m' 组合。图13-5中的点线表示波导结构中横向的光场分布。在薄膜域, 光波波矢的 x 和 z 分量 ($\beta\kappa$ 和 $b_2\kappa$) 均为实数, 所以薄膜内的光波为体波; 相反, 在覆盖层和衬底域, 只有 x 方向的波矢分量 ($\beta\kappa$) 为实数, 而 z 方向的波矢分量 ($b_1\kappa$ 和 $b_3\kappa$) 为虚数, 所以相应的光

波在 z 方向快速衰减, 是一种表面波(倏逝波或消失波)。

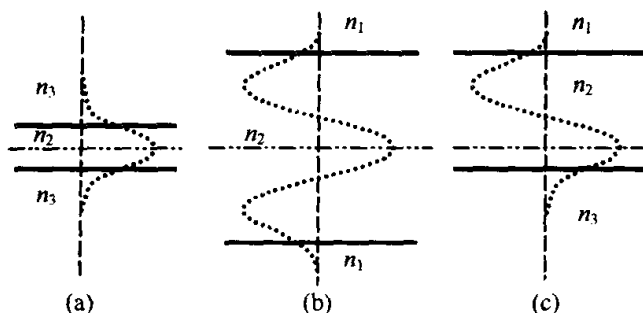


图 13-5 波导中的场分布

- (a) 具有传播常数 β 为 $m=0$ 阶模的对称波导;
- (b) 具有传播常数 β 为 $m'=2$ 阶模的对称波导;
- (c) 具有传播常数 β 为 $m''=1$ 阶模的非对称波导

13.1.5 光波在光纤中的传播

光纤即光导纤维, 是由低损耗材料制成的其结构类似于同轴电缆的圆形介质波导: 内芯是一种较高折射率的材料, 周围是另一种较低折射率的材料构成的包层, 最外层是保护层, 结构如图 13-6。

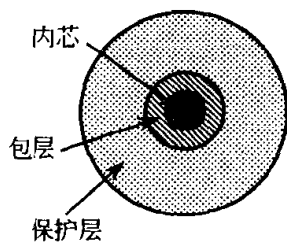


图 13-6 光纤的横断面结构

(1) 光纤的分类 光纤的种类按结构可分为阶跃型折射率光纤和梯度型折射率光纤。阶跃型折射率光纤, 内芯和包层间折射率的过渡是阶跃的。梯度型折射率光纤, 折射率分布为从中心轴线沿着径向逐渐减小。按材料又可分为: 石英光纤, 多组分玻璃光纤, 塑料光纤, 红外光纤, 紫外光纤和有源光纤等。石英光纤以 SiO_2 为主要原材料, 通过添加其他氧化物来改变其折射率, 掺杂 P_2O_5 或 GeO_2 可

提高折射率, 而 B_2O_3 则可降低折射率; 优点是损耗小、抗拉强度高、频带宽。多组分玻璃光纤由除 SiO_2 外还有 Na_2O 、 K_2O 、 CaO 、 B_2O_3 等多种材料制成, 其特点是温度低、熔点低、易加工、易获得大芯径和折射率差值大的光纤、但损耗大。塑料光纤主要以苯乙烯为芯、异丁烯树脂为包层或是以异丁烯为芯、掺氟异丁烯为包层, 其优点是柔韧性好、端面易加工、价格低廉、频谱宽, 但有耐热性差、直径均匀性差和损耗大的缺点。红外光纤一般由 TiBr - TiI 制成。紫外光纤主要用熔融石英、蓝宝石和硅树脂等材料制成。有源光纤是在光纤中掺有激光活性的 Nd 离子或 Er 离子, 使之可以自身发光的光纤。此外, 光纤还可以分为: 单模和多模光纤, 保偏和非保偏光纤, 传能、传光束和传像光纤等。

(2) 光纤的传光特性和集光本领 光纤传光依赖的物理基础仍然是全内

反射。在光纤中,若某一光线的传播路径位于同一平面内,此光线称为子午光线,包含子午光线的平面则为子午面。如图 13-7,在入射端面与光纤轴

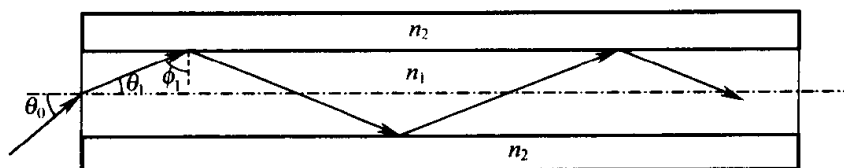


图 13-7 子午光线在光纤中的传播

线垂直的条件下,讨论最简单的平直光纤中的子午光线传播。光线从折射率为 n_0 的介质以入射角 θ_0 入射至光纤,折射角为 θ_1 ,在界面上的入射角为 ϕ_1 ,则有 $n_0 \sin \theta_0 = n_1 \sin \theta_1 = n_1 \cos \phi_1$ 。为确保发生全反射,必须满足 $\sin \theta_0 < (n_1^2 - n_2^2)^{1/2} / n_0$ 。入射角 θ_0 的最大值 θ_m ,由 $\sin \theta_m = (n_1^2 - n_2^2)^{1/2} / n_0$ 决定,不难想像它是表征光纤集光本领的一种量度; θ_m 越大说明光纤的集光本领越强。可以定义光纤的数值孔径 N.A. 为

$$N.A. = (n_1^2 - n_2^2)^{1/2} \quad (13-28)$$

可见,光纤的数值孔径仅与材料的折射率有关,而与其几何尺寸无关;可以制备数值孔径大且截面积小的光纤,使其易于弯曲,这是普通光学系统所不具有的特性。正因如此,光纤作为有效的传光媒质在光通信领域中起到了极为重要的作用。

(3) 光纤的色散与损耗 色散影响光通信系统的传输容量和误码率,而损耗影响传输距离。根据不同的机制,光纤中的色散大体上可以分为三类:材料色散,是由于光纤材料本身的折射率随频率变化,导致不同的频率分量的传播速度的不同而产生的色散,也被称为群速度色散;波导色散,依赖于光纤芯径与波长的比值,因其与光纤的结构有关也被称为结构色散;模色散,前两种色散是导致单模光纤中光脉冲展宽的主要原因,对于多模光纤情况则不同。由于在多模光纤中不同的导模(传播模)具有不同的传播路径,因而光程的不同导致不同的导模到达输出端的时间不同,光脉冲而被展宽,展宽的程度依赖于光纤的长度和导模的阶数。显然,在单模光纤中不存在模色散。在多模光纤中,模色散占支配地位,其他两种是次要的;在单模光纤中,一般材料色散是主要的色散来源,但在材料色散很小的波段,波导色散则成为主要的色散机制。

损耗就是光波在传输过程中,由于某种或某些因素所致的衰减。光纤中的损耗主要有两种:吸收损耗和散射损耗。吸收损耗又分为本征和杂质吸收,本征吸收是由于介质的原子或离子的电子跃迁所致的光吸收和分子振动所致的红外吸收,由介质的物理结构决定,是不可克服的。杂质吸收是由于

某些杂质（如金属离子、羟基 OH^- 等）的存在所致的选择吸收。散射损耗，包括瑞利散射和缺陷对光的散射所致的损耗。瑞利散射与光波长的四次方成反比，所以瑞利散射在短波段占支配地位，这也是难以克服的。缺陷（如气泡、杂质颗粒、条纹、应力等）所产生的散射源于介质折射率的局部不连续，可以通过改善制备工艺减小此类散射。

13.1.6 光波在光子晶体中的传播

光子晶体是当今非常热门的研究课题。光子晶体是介电常数被调制的人工微结构材料；其最根本的特征是具有光子禁带——光子带隙。频率位于禁带中的光波在光子晶体中不能传播；也可以理解为频率处于禁带中的光波将被光子晶体全反射。因此，光子晶体也被称为光子带隙材料，将为光信息技术提供一种全新的控制和操纵光的途径。

由于电子回路的种种固有制约，进一步提高其性能已遇到难以逾越的屏障。科学家们正在探索用光子代替电子作为信息载体的全新光子回路。光比电子具有显著的优点：光在介电材料中的传播速度远高于电子在金属导线中的传播速度，光可以携带更多的信息；介电材料的带宽远大于金属（光纤通信系统的带宽典型值为 1THz 量级，而电子系统仅为几百 kHz 量级）；特别地，光子（玻色子）间不像电子（费米子）间具有极强的相互作用，有助于降低能量损耗。尽管光子具有诸多优点，但是研制成商用的全光回路并非易事。光子晶体作为一种新材料是全光集成回路发展的关键。

光子晶体的概念源于 Yablonovitch 和 John 的早期想法。最基本的想法就是设计一种材料可以影响光子的特性，像在半导体材料中影响电子特性一样。Yablonovitch 的目的是控制材料中的辐射特性，而 John 则是通过引入随机的折射率变化去影响光子局域化。传统上，光子的操纵通常依赖于全内反射机制（如光波导和光纤），这严重地限制了光学元件的小型化程度，一则因为全反射界面相对于光波长必须是光滑的，二则为确保光波的传播对器件具有尺度要求。光子晶体具有完全不同的控制光机制，主要在于光子带隙。

为理解光子晶体中光的传播行为，最简单的方法就是与半导体中的电子和空穴的移动类比。在半导体中，电子在原子晶格中移动将经历一个周期势场，导致允许和禁戒能态的形成。对于纯而完美的半导体，在禁戒能隙（带隙）范围内，没有任何电子存在。不过，如果半导体中的本位原子缺少或被杂质占据而引起晶格的周期性破坏、或材料中在非晶格位置有额外原子，电子可以有一个位于带隙内的能态。同样在介电常数周期调制的光子晶体中，对光子而言这样的周期结构看上去恰似电子在半导体中经历的传播。此限制导致被一个禁戒区域（光子带隙）分隔开的允许能量区域的形成。当然，同样也可以通过缺陷或周期性的破坏，在光子带隙内建立一个能级（局域光子

态)。点缺陷像一个微腔，线缺陷类似于一个波导，而面缺陷则恰似一个完美的反射镜，为操纵和控制光子提供了可能。光子晶体的晶格常数必须可以与光波长比拟。例如，对光电子产业，其常用波长为 $1.5\mu\text{m}$ ，一个有用的光子晶体的晶格常数为 $0.5\mu\text{m}$ 量级。这样复杂结构的制备需要微刻蚀技术，如电子束刻蚀和 X 射线刻蚀。

电子在半导体材料中的传播由薛定谔方程支配

$$[-\nabla^2/2 + V(\mathbf{r})]\Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \quad (13-29)$$

而在光子晶体中的光子传播特性由麦克斯韦方程支配

$$[\nabla \times \epsilon^{-1}(\mathbf{r})\nabla \times]\mathbf{H}(\mathbf{r}) = \omega^2\mathbf{H}(\mathbf{r}) \quad (13-30)$$

上面两个方程实质上是线性本征值问题，其解分别由半导体中的势函数 $V(\mathbf{r})$ 或光子晶体中的介电函数 $\epsilon(\mathbf{r})$ 确定。如果存在一个带结构，那么就有可能建立一个全光子带隙的光子晶体。但是为了得到全带隙，构成光子晶体的两种材料的介电常数的比必须满足一定值。带隙就是一个禁止在光子晶体中沿任何方向传播的光子频率范围。光子晶体中的一个缺陷可以起到内表面的作用，光子可以被限制或被局域，这给出了控制光的一个新“自由度”。实际上，多层膜就是一种光子晶体（即一维光子晶体），如图13-8(a)，光子带隙只出现在一个传播方向（即多层膜的法线方向）。如果在二维甚至三维具有周期性，至少在原理上可以建立全带隙的两维和三维光子晶体，分别如图13-8(b)和(c)。

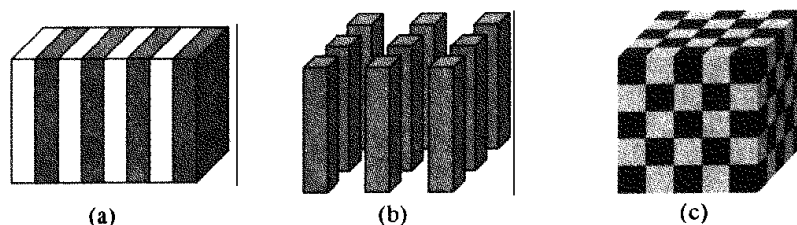


图 13-8 光子晶体

(a) 一维光子晶体；(b) 二维光子晶体；(c) 三维光子晶体

众所周知，信息技术领域的成就主要来源于有能力用复杂的方式控制半导体结构，进而控制其特性。光子晶体提供了类似的甚至更灵活的控制光的方法，因为控制光子晶体中的光比控制半导体中电子方法多且较容易。光通信系统中存在大量的分立元件，因而尽可能将它们组合到一个“集成”光学芯片上就显得特别有意义。不仅减少大量的人工，还能减少可能的故障点数。但这种集成又会带来其他问题。首先需要开发小尺寸的光学接口——即微小的能在急弯处（甚至是锐的直角）导光的波导。在微芯片上让电子拐急弯是很简单的，但是用传统的波导来让光拐急弯是不可能的。如果在规则的

光子晶体中移去一排小孔，则可能在光子晶体中构造出一个窄信道波导。当光波长位于光子晶体带隙内时，光就会被禁闭在这一线状缺陷内。在这些条件下，可以引入一些有急弯的点阵结构来使光作背向反射或转弯，已经在理论上证明损耗很小甚至为零。

由发光材料制成的光子晶体，也可用来制造与其他元件一起集成在光通信系统中的窄线宽激光器。该激光器可通过在光子晶体中引入少量的空洞来制成。这些“微腔”在光子带隙中产生了窄的“缺陷模”。当材料在很宽的光谱范围内发光时，只有与缺陷模波长匹配的光波才被放大，因为只有它才可以在该材料中自由传播，同时其他波长的光波被抑制。意味着激光器发射波长具有窄的谱宽。发光效率的增加意味着微腔二极管和光子晶体激光器比传统的设备开关速度更高，可导致更高的信息传输率和能量效率。

传统光纤有一个缺点是不同波长的光波会以不同的相速度在材料中传播。因色散发生时间延迟而展宽光脉冲，限制了可发送的信息量。一种奇异的空芯光子晶体光纤，可望解决此难题。它是这样制成的：把一系列的空心玻璃毛细管堆砌在一个实的玻璃芯周围，在加热拉伸成直径只有几微米的长光纤。这种光纤具有不寻常的特性，即便在芯径很大的情况下仍可以传输单模光。后又把中央的实玻璃芯去掉构成一个长空气腔而制成了一个更加非比寻常的光纤。此情况下，由于光子带隙限制效应，而使光实际上在低折射率的空气芯中传播；又由于光并未在玻璃材料中传播，因此它可以传输大功率的激光而不损伤自己，而且无色散不会引起光脉冲展宽。

光子晶体的提供一种全新的控制和操纵光的途径，必将在不久的将来发挥重要的作用。有学者预言：基于光子晶体的光子回路和器件，在今后 5 年内一些基础应用将会面世。其中会有高效的光子晶体激光器和极亮的发光二极管。同时，期望利用光子晶体波导将光限制在微米尺度内传播以及将光子晶体用于高分辨的光谱滤波的可行性得到验证。人们可能在 10 年内研制成商用的“芯片光谱仪”。同样在 5~10 年内，人们将证实光子晶体“二极管”和“三极管”的可能。在 10~15 年内，将首次出现光子晶体逻辑回路，同时光子晶体驱动的光学计算机样机也会在 25 年内出现。

13.2 光发射

光发射是一个极为重要的物理过程，对其深入的理解和研究，导致了一个奇异的器件——激光器的诞生。可以毫不夸张地讲，激光器的问世不仅给光学、物理学乃至整个科学研究领域带来了一场深刻的革命。其触角几乎渗透到各个领域，并充分发挥其巨大的潜力。光发射的研究一直是科学家们研

究的热门领域, 对新的光发射机理和器件的探索从未间断, 如现在广为研究的光子晶体激光器, 纳米硅基发光等。在此由于篇幅的限制不可能详细阐述所有的光发射问题, 仅就光发射的基本概念和原理, 激光产生机理和最新的光发射现象作些简明扼要的描述, 希望能起到抛砖引玉的作用。

13.2.1 光的自发辐射和受激跃迁

光波通过物质系统并与其相互作用, 电子将被激发到较高能态, 光波被吸收。另一方面, 电子一旦被激发, 最终必定返回到较低的能态, 并伴随有光子的发射或热的释放过程。这一自发的发光过程, 若是一个快过程被称为荧光, 若较慢则被称之为磷光。下面, 将阐述光的自发辐射和受激(吸收、辐射)跃迁。

光的自发辐射和受激跃迁是爱因斯坦从光量子概念出发, 重新研究黑体辐射的普朗克公式时提出的两个极为重要的概念, 由此奠定了激光的物理基础。为解释在实验中实测到的黑体辐射谱 $\rho(\nu)$ 随温度 T 和频率 ν 的分布规律, 普朗克提出了与经典概念完全相悖的辐射能量量子化假设, 成功地得到了与实验结果完全相符的黑体辐射的普朗克公式

$$\rho(\nu) = \frac{8\pi n^3 h \nu^3}{c^3} \frac{1}{\exp(h\nu/k_B T) - 1} \quad (13-31)$$

其中 h 和 k_B 分别为普朗克和玻尔兹曼常数, 而 ν 、 c 和 T 则分别为频率、光速和绝对温度。爱因斯坦从辐射与物质系统相互作用的量子观点出发提出: 辐射应包含自发辐射、受激辐射和受激吸收跃迁三个过程。

(1) 自发辐射 通常一个物质系统具有多个能级, 但是为简化且不失一般性, 只关心其中与自发辐射跃迁相关的两个能级。在 $t=0$ 时刻, 处于能量为 E_2 的能态 $|2\rangle$ 的粒子, 将按一定几率自发跃迁至能量为 E_1 的较低能态 $|1\rangle$, 同时伴随一个能量为 $h\nu = E_2 - E_1$ 的光子产生。此过程是自发的, 不需要任何外来的诱导, 因此称之为自发辐射。此种辐射几乎不受方向性的限制, 是非相干的且具有宽的谱带。考虑大量全同粒子 N_2 在 $t=0$ 时刻处于能态 $|2\rangle$, 经一定的特征时间自发地跃迁到低能态 $|1\rangle$ 的平均数为 $A_{21}N_2$, 其中 A_{21} 为自发辐射跃迁的爱因斯坦系数, 表示单位时间内从高能态 $|2\rangle$ 到低能态 $|1\rangle$ 发生自发辐射跃迁的几率。 A_{21} 只与物质自身的性质有关, 可通过计算能态 $|2\rangle$ 和 $|1\rangle$ 的本征波函数得到。可以证明 A_{21} 就是粒子在能态 $|2\rangle$ 上的平均寿命 τ_s 的倒数 $A_{21} = 1/\tau_s$ 。

(2) 受激跃迁 若黑体物质粒子与辐射场相互作用仅有自发辐射跃迁过程, 不能维持式 (13-31) 所描述的辐射场的稳定平衡。鉴于此, 爱因斯坦认为: 必然存在一种在辐射场作用下的诱导跃迁过程 (即受激跃迁)。处于低能态 $|1\rangle$ 的粒子在频率为 ν 的辐射场的作用下跃迁至能态 $|2\rangle$, 并吸收一

个能量为 $h\nu$ ($=E_2 - E_1$) 的光子, 此过程称为受激吸收跃迁。当然, 相反的过程也会发生, 粒子受到辐射场的作用, 处于能态 $|2\rangle$ 的粒子会发射一个相同能量为 $h\nu$ 的光子, 跃迁至能态 $|1\rangle$, 这一过程称为受激辐射跃迁。

受激辐射和自发辐射跃迁的区别之一是从 $|2\rangle \rightarrow |1\rangle$ 和 $|1\rangle \rightarrow |2\rangle$ 的受激跃迁几率相等, 而 $|1\rangle \rightarrow |2\rangle$ 的自发跃迁几率为零; 之二是受激跃迁几率与外辐射场成正比, 而自发跃迁与此无关。假定辐射场的能量密度的频率依赖函数为 $\rho(\nu)$, 于是 $|2\rangle \rightarrow |1\rangle$ 和 $|1\rangle \rightarrow |2\rangle$ 的受激跃迁几率应为 $W_{21} = B_{21}\rho(\nu)$ 和 $W_{12} = B_{12}\rho(\nu)$ 。

(3) 爱因斯坦系数关系 A_{21} 、 B_{21} 和 B_{12} 被称为爱因斯坦系数。 B_{21} 和 B_{12} 的大小, 与 A_{21} 一样, 仅与物质的性质有关, 而与辐射场无关, 所以可以考虑一个处于热平衡时温度为 T 的黑体辐射场的情形, 这样也不会失去其普遍性。在热平衡时处于能态 $|1\rangle$ 和 $|2\rangle$ 的平均粒子数不随时间而变。因此 $|2\rangle \rightarrow |1\rangle$ 的跃迁粒子数应等于 $|1\rangle \rightarrow |2\rangle$ 的跃迁数, 即有

$$N_2[B_{21}\rho(\nu) + A_{21}] = N_1B_{12}\rho(\nu) \quad (13-32)$$

其中 N_1 和 N_2 分别为处于能态 $|1\rangle$ 和 $|2\rangle$ 的粒子数。处于热平衡时, 在两个能态上的粒子数分布遵从玻尔兹曼分布

$$N_1/N_2 = \exp[-(E_2 - E_1)/k_B T] = \exp[-h\nu/k_B T] \quad (13-33)$$

将式 (13-31) 代入式 (13-32), 令上面两式中的 N_1/N_2 相等, 则有

$$\frac{8\pi n^3 h\nu^3 / c^3}{\exp(h\nu/k_B T) - 1} = \frac{A_{21}}{B_{12}\exp(h\nu/k_B T) - B_{21}} \quad (13-34)$$

为使此式成立, 必须同时满足下面两式

$$B_{12} = B_{21}, \quad A_{21}/B_{21} = 8\pi n^3 h\nu^3 / c^3 \quad (13-35)$$

此即爱因斯坦系数关系。

(4) 光的受激放大与布居反转 由大量粒子构成的物质系统, 在热平衡时各能态上的粒子数布居遵从玻尔兹曼分布。由于高能态 $|2\rangle$ 的能量 E_2 总是大于低能态 $|1\rangle$ 的能量 E_1 , 因此恒有 $N_2 < N_1$; 即在热平衡时高能态的布居数不可能大于低能态。从爱因斯坦系数关系可知, 在能态 $|2\rangle$ 和 $|1\rangle$ 之间的受激吸收光子数 $N_1 B_{12} \rho(\nu)$ 恒大于受激辐射光子数 $N_2 B_{21} \rho(\nu)$, 因此物质系统只能吸收光子, 而不可能有净的受激辐射光子产生。

但是, 通过外部激励 (或称为泵浦), 在非热平衡状态下, 可以实现高能态粒子数大于低能态粒子数 ($N_2 > N_1$), 即实现布居反转; 这样受激辐射将大于受激吸收, 将会有净的受激辐射光子产生而实现光放大, 不过这是以消耗外部激励能量为代价的。现在唯象地描述受激光放大的过程。频率为 ν ($h\nu = E_2 - E_1$) 的光子入射至已处于布居反转的物质系统, 并与之发生相互

作用。在此入射光子的诱导下，处于上能态 $|2\rangle$ 的粒子将回落至下能态 $|1\rangle$ ，同时伴随另一个频率为 ν 的光子发射，该光子与入射光子具有完全相同的状态（如传播方向、偏振态、频率和位相等），是完全相干的。这两个光子进一步与物质系统相互作用，诱导出更多的状态完全相同的光子，直至达到雪崩过程的建立，达到最终受激放大的目的。可以发现受激辐射与自发辐射另一个重要区别在于相干性。自发辐射场不仅在位相上是随机的、不可控的、而且出现在整个空间立体角内，是完全非相干辐射。相反，受激辐射则不然，在入射光场的诱导下，产生的受激辐射光子，具有与入射光子完全相同的状态，是完全相干的，这正是激光的最基本的物理基础。需要注意的是：在此讨论激光放大采用的是一个简化的二能级系统，实际上即使有外部激励，二能级系统也难以实现布居反转。为实现有效的布居反转，物质系统至少为三能级系统。

13.2.2 激光产生的原理

将已处于布居反转的物质系统（将可实现布居反转的物质系统称为激光增益介质）置于由两个平行反射镜构成的谐振腔内（如图 13-9），增益介质所发出的传播方向平行于两个反射镜轴线的自发辐射光在谐振腔内来回反射而多次往返穿过激光增益介质时会被不断放大，如果放大超过了所有损耗，谐振腔内储存的能量将随时间而增加，直至单程饱和增益恰好等于损耗时，达到稳态自振荡，产生激光。

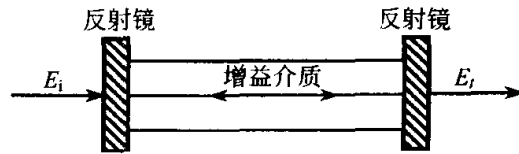


图 13-9 激光器结构原理图

光波场与增益介质的相互作用，是在近共振乃至共振区域，所以光波在其中的传播与在远离共振情况不同。不仅伴随有光子的吸收与发射，而且同时存在显著的色散效应。在激光增益介质中的传播常数由两部分（四项）组成

$$K(\nu) = (k + i\alpha/2) + (k' + i\alpha'/2) \quad (13-36)$$

$$k' = k\chi_R/2n^2, \alpha' = k\chi_I/n^2 \quad (13-37)$$

其中 $k + i\alpha/2$ 为远离激光跃迁共振频率时的介质传播常数， α 代表介质的分布损耗， $k' + i\alpha'/2$ 为激光跃迁共振所致的附加传播常数。

现在考虑一复振幅为 E_i 的平面光波场从含有激光增益介质的谐振腔左侧反射镜入射情形，假设腔的长度为 L ，且激光增益介质充满整个腔，两个反射镜对光波振幅透射率和反射率分别为 t_1, t_2 和 r_1, r_2 ，如图 13-9。由于光波场在腔内的多次反射，根据光波的相干叠加原理，在右侧总的输出光波场复振幅 E_f 为^[5]

$$E_t = E_i \frac{t_1 t_2 e^{i(k+k')L - (\alpha + \alpha')L/2}}{1 - r_1 r_2 e^{i2(k+k')L - (\alpha + \alpha')L}} \quad (13-38)$$

为了得到 k' 和 α' 的表达式, 在此不能使用第一节中 χ_R 和 χ_I 的表达式 (13-6), 因为那是在理想电子系统中得到的, 不能反映真实原子跃迁。必须使用量子力学中 χ_R 和 χ_I 的表达式

$$\chi_R(\nu) = \frac{(N_1 - N_2)(\nu_0 - \nu)\lambda^3}{8\pi^2 n t_s \Delta\nu} g(\nu), \quad \chi_I(\nu) = \frac{(N_1 - N_2)\lambda^3}{16\pi^2 n t_s} g(\nu) \quad (13-39)$$

其中 $g(\nu)$ 为洛伦兹线型函数, 而 t_s 表示自发辐射的寿命, N_1 和 N_2 分别为能态 $|1\rangle$ 和 $|2\rangle$ 的粒子数密度。将上面两式分别代入式 (13-37) 可以得到 k' 和 α' 的表达式。在经典电子振荡模型中, 那里的 χ_I 恒大于零, 所以介质仅表现为吸收。但在量子力学模型中, 此处的 χ_I 或 α' 允许小于零, 即当 $N_2 > N_1$ (即布居反转) 时, 可表现为放大 (或增益)。在此定义

$$\gamma(\nu) = -\alpha' |_{N_2 > N_1} = \frac{k(N_2 - N_1)\lambda^3}{16\pi^2 n^3 t_s} g(\nu) \quad (13-40)$$

为增益系数。同时, $\gamma(\nu)$ 随 ν 变化的曲线称之为增益曲线, 如图 13-10 (a)。所以当增益介质处于布居反转时, 从式 (13-38) 可知, 透过光波场 $|E_t|$ 完全可以大于入射光波场 $|E_i|$ 。即含有增益介质的腔具有光波的放大功能。若式 (13-38) 中的分母为零, 即

$$1 - r_1 r_2 e^{i2(k+k')L + (\gamma - \alpha)L} = 0 \quad (13-41)$$

成立时, $|E_t|^2 / |E_i|^2$ 将趋于无穷, 也就是说此时即便无任何入射光波场的存在, 仍会有有限的透射光波场存在, 此乃所谓的自振荡。实际上, 在外界激励的情况下, 激光增益介质会产生自发辐射。这种非相干的自发辐射, 作为“种子”光, 在腔内沿腔轴方向被不断放大, 因为两个平行反射镜提供正反馈而产生振荡, 实现激光输出。从式 (13-41) 可以得到实现激光振荡输出的两个充分条件

$$r_1 r_2 e^{(\gamma - \alpha)L} = 1, \quad 2(k + k')L = 2m\pi \quad (13-42)$$

前者被称为阈值条件, 必须清楚: 其先决条件是布居反转 (必要条件)。仅满足阈值条件未必能够实现振荡, 同时还必须满足上式的后者即所谓的共振 (相干相长) 条件。

激光器的纵模。从共振条件可知将有一系列共振频率 ν_m 可以满足, 这样的每一个共振频率被称为一个纵模, m (正的整数) 为相应模的阶数。为求得振荡频率 ν_m , 利用式 (13-37) 共振条件可以改为

$$\nu_m [1 + \chi_R(\nu) / 2n^2] = mc / 2nL \quad (13-43)$$

定义无源腔（即 $N_2 - N_1 = 0$ ）的第 m 阶模共振频率为 ν_{pm} （ $= mc/2nL$ ）；可见，其相邻阶模具有相等的间隔 $c/2nL$ ，如图 13-10（b）。从式（13-39）和式（13-40）可得 $\gamma(\nu) = -k\Delta\nu\chi_R(\nu)/2(\nu_0 - \nu)n^2$ ，从式（13-43）得到有源腔的第 m 阶纵模共振频率 ν_m 为

$$\nu_m \left[1 - \left(\frac{\nu_0 - \nu_m}{\Delta\nu} \right) \frac{\gamma(\nu_m)}{k} \right] = \nu_{pm} \quad (13-44)$$

其中 ν_0 为图 13-10（a）示出的增益曲线的中心频率。假定某一纵模的谐振频率 ν_m 与 ν_0 接近，可以近似地得到

$$\nu_m = \nu_{pm} - (\nu_{pm} - \nu_0) \gamma(\nu_{pm}) c / 2\pi n \Delta\nu \quad (13-45)$$

若无源腔的某一纵模共振频率 ν_{pm} 恰好等于 ν_0 。（即 $\nu_{pm} = \nu_0$ ），由上式可知 $\nu_m = \nu_0$ 。（即此时有源腔的相应纵模共振频率也恰好处于 ν_0 ）。相反，如果 $\nu_{pm} \neq \nu_0$ ，有源腔的 ν_m 相对于无源的 ν_{pm} 会向中心频率 ν_0 方向略有移动，无论 $\nu_{pm} > \nu_0$ 还是 $\nu_{pm} < \nu_0$ ，如图 13-10（c）。这一现象常称为“频率牵引”，是有源腔和无源腔的主要区别之一，源于共振跃迁所致的极化强度共振分量对色散的贡献。

单纯地从共振条件式（13-44）或式（13-45）发现，可以存在无限多个纵模（共振频率 ν_m ），但实际上，并非所有满足这一条件的纵模都能起振。如图 13-10（d），只有位于增益曲线内，且那些增益大于振荡阈值 γ_t 的纵模才能共振，并产生激光输出。通常情况下，一个激光器的输出并非单纵模的，而是一些具有分立共振频率的多纵模输出。振荡纵模的数量，取决于增益曲线的宽度和相邻纵模的间隔。腔长 L 较短时纵模间隔较大，有利于单纵模运转。同时提高振荡阈值，也是实现单模运转的有效手段。顺便提一下，有源腔（即实际的激光器）的相邻纵模间隔是不等的，依赖于相应模的阶数，与无源腔不同。

13.2.3 典型的三能级和四能级激光系统

至此，所有讨论都集中在与激光跃迁直接相关的二能级系统，但实际上的激光系统并不是简单的二能级，而至少是三能级和四能级。理想三能级系

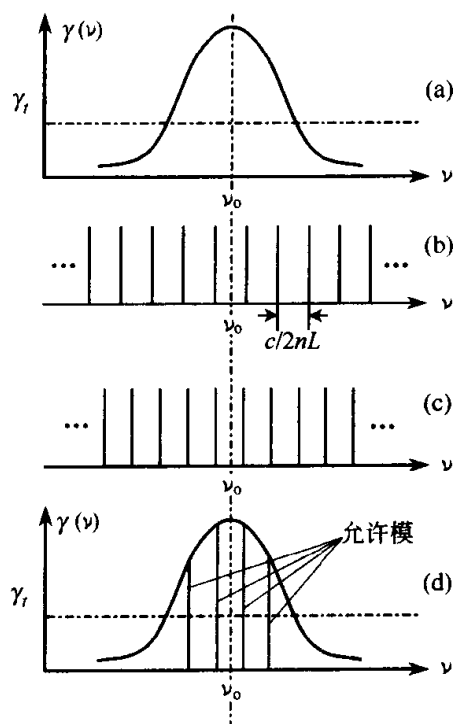


图 13-10 增益曲线与激光腔模

- (a) 增益曲线；(b) 无源腔的纵模；
 (c) 有源腔的纵模（存在频率牵引现象）；
 (d) 有源腔激光器的振荡模

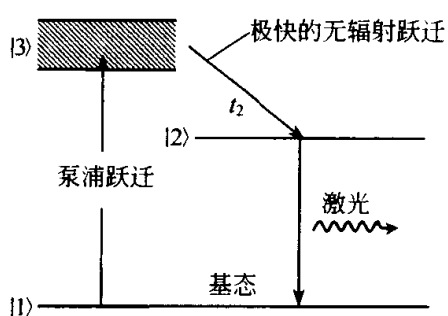


图 13-11 理想三能级系统

统的能级图如图 13-11。较低的能级要么是基态, 要么是与基态相差甚小 (其差比热能 $k_B T$ 小许多), 以至于在热平衡时, 几乎所有的粒子占据了 this 较低能级。能级 $|1\rangle \rightarrow$ 能级 $|3\rangle$ 的跃迁对应于激励 (或泵浦) 跃迁; 被激发至能级 $|3\rangle$ 的粒子, 将很快地通过无辐射跃迁至能级 $|2\rangle$, 几乎没有粒子滞留在 $|3\rangle$ 上; 能级 $|2\rangle \rightarrow$ 能级 $|1\rangle$ 的跃迁, 对应于激光跃迁, 产生激光

光输出。当外部激励足以使较高激光上能级 $|2\rangle$ 的粒子数密度 $N_2 = N_1 = N_0/2$ (N_0 为总的粒子数密度), 光学增益系数 $\gamma = 0$ 。为满足阈值条件, 激励必须进一步提高, 直至满足 $N_2 = N_0/2 + N_t/2$ 和 $N_1 = N_0/2 - N_t/2$, 其中 $N_t = N_2 - N_1$ 为反转阈值粒子数密度。在许多激光系数中 $N_0 \gg N_t$, 那么, 对于三能级系统 $N_{2\text{三能级}} \approx N_0/2$ 。

理想四能级系统如图 13-12, 基态 $|0\rangle \rightarrow$ 最高能级 $|3\rangle$ 的跃迁, 对应于激励跃迁; 被激励至能级 $|3\rangle$ 的粒子迅速地以无辐射跃迁的形式, 落回至能级 $|2\rangle$; 能级 $|2\rangle \rightarrow$ 能级 $|1\rangle$ 的跃迁, 对应于激光跃迁; 激光下能级 $|1\rangle \rightarrow$ 基态 $|0\rangle$ 的跃迁通常是以无辐射形式进行的。而且能级 $|1\rangle$ 与基态 $|0\rangle$ 的能量差远大于热能 $k_B T$, 所以在热平衡时, 激光下能级 $|1\rangle$ 上的粒子数布居可以忽略不计。于是与 N_2 相比,

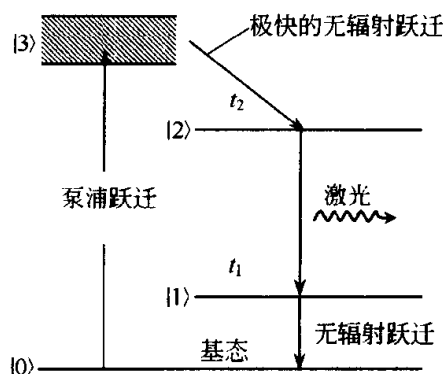


图 13-12 理想的四能级系统

N_1 可以忽略, 那么当 $N_2 \cong N_t$ 阈值条件就容易被满足。如果激励使得激光上能级的粒子数密度达到阈值 N_t , 激光器就会起振。可以比较发现, 在所有条件都相同时, 三能级系统的阈值会远远超过四能级, 它们的比值 $N_{2\text{三能级}}/N_{2\text{四能级}} \sim N_0/2N_t$ 。可见, 四能级系统更容易实现布居反转, 产生激光振荡。所以四能级系统是一个更理想的激光系统。不同类型的激光器, 可用不同的方法将粒子激发到较高的激光能级, 下面扼要地阐述两种具有代表性的 Nd: YVO₄ 和半导体激光系统。

13.2.4 Nd: YVO₄ 和半导体激光器

(1) Nd: YVO₄ 激光器 Nd: YVO₄ 晶体虽早已被发明, 但因其生长方面的问题, 直至上世纪 80 年代中后期一些基本问题解决之后, 才重新引

起注意, 成为目前非常重要的一种激光介质并发挥了非常重要的作用。

YVO_4 为四方晶体, 锆石英 (ZrSiO_4) 结构, 属正单轴晶体。 YVO_4 晶体中的 Y 离子部分被 Nd 取代 (Nd 以 +3 价的离子形式存在) 而形成 Nd: YVO_4 。Nd: YVO_4 晶体具有大的双折射 (在 1064nm 处, $n_o = 1.958$, $n_e = 2.168$), 并且其 a 轴切割时光吸收和辐射具有明显的偏振依赖性, 最强吸收和最强辐射均发生在 π 偏振方向 ($E \parallel C$ 轴), 因此非常有利于偏振输出, 与 Nd: YAG 晶体截然不同。

Nd: YVO_4 晶体的能级结构 (如图 13-13) 与其他掺 Nd 的 YAG 和 YLF 等大致相同。受激光跃迁过程如实线表示 (向上箭头代表吸收跃迁, 向下则为辐射跃迁), 无辐射跃迁过程用波浪线表示。有两个吸收带分别位于 880nm 和 808nm 附近, 分别对应于从 $^4I_{9/2} \rightarrow ^4F_{3/2}$ 和 $^4I_{9/2} \rightarrow ^4F_{5/2}$ 的跃迁。其中最重要的也是最强吸收带为 808nm 附近的一个, 峰值波长为 808.7nm, 吸收带宽约为 20nm (远宽于 Nd: YAG 的 4nm) 而且吸收截面大; 这非常有利于 LD 泵浦, 因为其宽度远大于 LD 的谱宽, 即使因温度的漂移导致的 LD 发射波长的改变也不会显著地影响泵浦效率。此外, 吸收系数随着 Nd 离子掺杂浓度的增加而增大。用 808nm 附近的 LD 泵浦 Nd: YVO_4 晶体, 将粒子从基态激发至能级 $^4F_{5/2}$, 因为该能级的寿命极短 (约为 0.1ns), 所以粒子通过无辐射跃迁快速地弛豫至能级 $^4F_{3/2}$ 。能级 $^4F_{3/2}$ 是一个亚稳态 (寿命约为 0.1ms), 因此易于实现布居反转。处于能级 $^4F_{3/2}$ 的粒子跃迁至低能级, 主要有四条发射谱线: $^4F_{3/2} \rightarrow ^4I_{15/2}$ 、 $^4F_{3/2} \rightarrow ^4I_{13/2}$ 、 $^4F_{3/2} \rightarrow ^4I_{11/2}$ 、

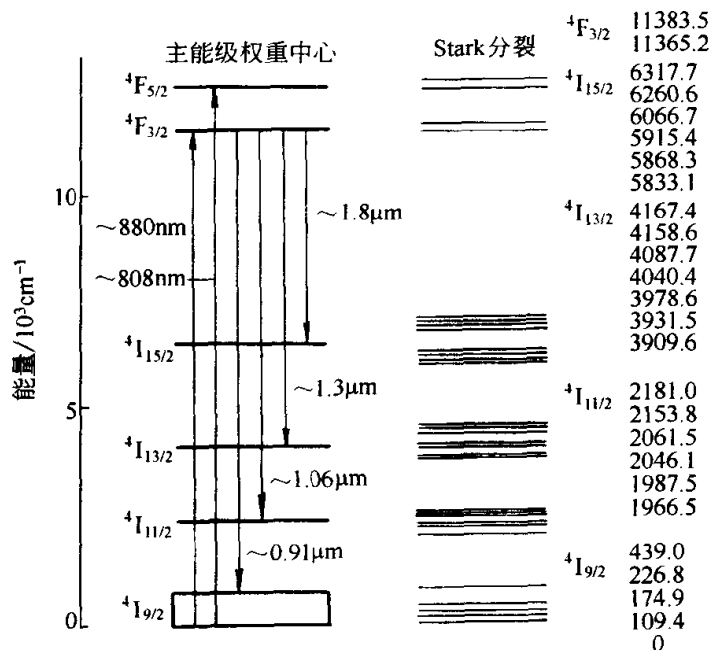


图 13-13 Nd: YVO_4 晶体能级图

${}^4F_{3/2} \rightarrow {}^4I_{9/2}$, 相应的发射波长分别为 1839nm、1342nm、1064nm、914nm。其中 1064nm 谱线的发射截面 ($20 \times 10^{-19} \text{cm}^2$, 约 Nd: YAG 相应波长的 7~8 倍) 最大, 是最强的一条谱线, 占绝对优势, 因而增益最高, 不过其谱线宽度为 0.8nm 略宽于 Nd: YAG 的 0.6nm; 1342nm 次之 (发射截面为 $6 \times 10^{-19} \text{cm}^2$); 1839nm 和 914nm 最弱。能级图的右侧给出了主能级的 Stark 分裂的各个能级大小, 精确的辐射跃迁发生在各个不同的子能级之间, 如 914nm 对应于从 ${}^4F_{3/2}$ 的 11383.5cm^{-1} 能级到 ${}^4I_{9/2}$ 的 439.0cm^{-1} 。Nd: YVO₄ 晶体基态的 Stark 分裂仅有 439.0cm^{-1} , 相当于室温下热能的约 20%, 因此 914nm 的跃迁可以被看作准三能级系统中的跃迁。相反其他三条谱线的跃迁则是典型的四能级系统。同时其光-光转换效率高, 最高的斜效率达 72%。虽然 Nd: YVO₄ 晶体具有许多优点, 但最大的缺点是其热传导性较差。综合以上因素, Nd: YVO₄ 晶体很适合于中小功率激光器。我国生产的激光晶体大批出口到国外。令人高兴的是我们生长出许多高质量的晶体, 但令人遗憾的是我们自己并未研制成在国际上有影响的激光器系统, 出口的仅仅是高科技“原料”或“配件”。

(2) 半导体激光器 就基本原理而言, 半导体激光器和其他类型的激光器 (如前面刚提到的 Nd: YVO₄ 激光器) 没有本质上的区别。为得到相干的、受激光输出, 也必须满足布居反转 (必要条件) 和阈值条件 (充分条件)。与其他激光器不同的是, 半导体激光器中电子分布在不同能带的不同能量状态中, 而其他激光器中的粒子则是分布在有源介质的不同能级上, 因此布居反转条件的具体表达形式也有差异。

半导体中的电子在价带与导带的不同能态上的正常分布遵从费米-狄拉克统计规律, $f(E) = \{1 + \exp[(E - E_F)/k_B T]\}^{-1}$, E_F 为量子系统的费米能级。若 $E_F - E \ll k_B T$, 该能级基本上被电子所占据; 若 $E - E_F \gg k_B T$, 电子则遵从与一般固体或气体的粒子数布居相同的玻尔兹曼分布规律。利用外部激励可以使固体、液体或气体激光增益介质实现布居反转 (即反玻尔兹曼分布)。但在具有能带结构的半导体有源介质中沿用上述布居反转条件则难以理解, 因为价带电子态密度比导带高得多。难以想像能利用某种方法使导带中的电子数多于价带 (即布居反转)。但伯纳德-杜拉夫格从电子的能带间的跃迁速率出发, 得到了半导体有源介质中的布居反转条件。在能量密度为 $\rho(\nu)$ 的光作用下, 当忽略自发辐射时, 净的受激发射速率必须大于零

$$B_{21}f_v\rho_c\rho_v(1-f_v)\rho(\nu) - B_{12}f_c\rho_v\rho_c(1-f_c)\rho(\nu) > 0 \quad (13-46)$$

上式左侧的两项分别为电子从价带至导带和从导带至价带的受激发射跃迁几率, ρ_c 和 ρ_v 以及 f_c 和 f_v 分别为导带和价带的电子态密度以及相应态的电子

占据几率。根据爱因斯坦系数关系式(13-35)，并考虑费米-狄拉克分布，得到半导体激光器布居反转条件

$$E_{F_c} - E_{F_v} > h\nu \geq E_g \quad (13-47)$$

在此， E_{F_c} 和 E_{F_v} 为导带和价带的费米能级，而 E_g 为禁带宽度。上式意味着为实现半导体有源介质中的布居反转，导带与价带的费米能级差必须大于禁带宽度，如图13-14(a)。在同质 p-n 结激光器中，要通过重掺杂使 E_{F_c} 进入导带，或者 E_{F_c} 和 E_{F_v} 同时分别进入其导带和价带。但对双异质结半导体激光器可以利用异质结势垒很好地将注入的载流子限制在有源区中而得到高的非平衡电子浓度，无需重掺杂即可满足布居反转条件。从图13-14(a)可以发现，由于价带准费米能级处于价带顶之上，价带内参与受激发射的能级很少被空穴所占据，因此为实现布居反转必须提高注入载流子的密度，势必增大阈值电流。如图13-14(b)，通过能带工程设计可以使价带的有效质量接近导带（ E_{F_c} 和 E_{F_v} 同时分别进入其导带和价带），不仅可以降低阈值电流，而且也可以减少俄歇复合进而提高激光器的稳定性；这可以通过应变超晶格实现。

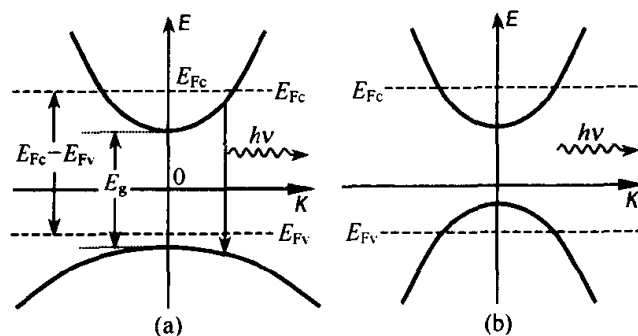


图 13-14 半导体能带结构

(a) 伯纳德-杜拉夫格布居反转条件；

(b) 理想的满足布居反转条件的能带结构

与其他激光器一样，实现布居反转仅仅满足激光振荡的必要条件，还必须满足阈值条件。在半导体激光器中的阈值条件同样是通过光学谐振腔提供正反馈来实现。谐振腔可以是解理面构成的 FP 腔也可以是单频腔，后者又分为分布反馈(DFB)、分布布拉格反射(DBR)、外部闪耀光栅和光纤光栅等。半导体激光器与其他激光器的不同主要在于产生布居反转的机制，但阈值条件和谐振条件则无任何区别。半导体激光器的有源层有许多类型：常规的厚有源层、量子阱(单量子阱和多量子阱)、量子线和量子点等。概括地讲，半导体激光器是在直接带隙半导体 p-n 结中，注入载流子实现伯纳德-杜拉夫格条件控制的布居反转；通过高度简并的电子和空穴复合产生的受激

辐射经谐振腔放大，最终实现相干辐射输出。

由于信息、通信和激光技术等领域的应用需求以及微结构制备技术（如 MBE 和 MOCVD）的发展与完善，特别是由于能带工程，使得器件性能实现质的飞跃。迄今，已开发出满足各种不同应用所需波长和功率的半导体激光器材料体系，波长已从当初的近红外拓宽至现在的可见蓝光和近 $7\mu\text{m}$ 的红外光。目前，研究热点之一就是 III 族氮化物（特别是 GaN）半导体激光器。最主要的原因是其具有大的带隙，可以覆盖整个可见波段；其次，与其他高带隙半导体相比，氮化物的主要优点是具有强的化学稳定性，使得材料即使在高电流和强光照射下仍然非常稳定。氮化物也存在明显的缺点，因其强的化学键要求高的生长温度而限制衬底的选择，缺少生长薄膜所需的合适衬底。两个最常用的衬底材料是 $\alpha\text{-Al}_2\text{O}_3$ 和 6H-SiC 。现在，有关氮化物最主要的研究集中在实现室温连续工作的电流注入激光器。目前，氮化物激光器的阈值电流仍然很高，最具挑战性的工作就是降低阈值电流和增加寿命。

13.2.5 一种新的激光发射机理与材料——纳米硅发光

激光的发展经历了 40 年，取得了辉煌的成就，但是对新型激光器和新技术、新的发光机理、新材料的研究从未停止过。众所周知，如零阈值光子晶体激光器、低维纳米有序硅基发光，纳米硅晶发光是最近的研究热点。将硅片赋予发光功能是材料研究最具挑战性的问题。硅在微电子领域是一种极为重要的半导体材料。但已被认为对光电子应用是不适的材料，是因为它的间接带隙限制了作为光发射器的效率之缘故。在硅中的光放大是非常困难的。原因之一，硅对有效的自由载流子的吸收，降低了净增益；之二，在高功率下存在有显著的发光强度 Auger 饱和；之三，在硅纳米结构中存在明显的辐射能的尺寸依赖性，导致了在系统中大的非均匀展宽和显著的光损耗。不过，令人振奋的是最近的研究表明，当硅以低维系统的形式或者选择激活掺杂和(或)新相嵌入硅晶格时，室温硅的光发射是可能的。对于低维系统中的光致发光的物理机理主要是纳米尺度晶体结构的激发量子限制效应。最近，Paresl 等在他们的论文中，报道了在硅纳米结构中所做的受激发射和光放大的实验结果，并验证了单程配置中的光增益，并声称这一发现为实现硅基激光找到了一条新的途径。现在扼要地阐述该领域的最新进展。

硅纳米晶是采用负离子注入超纯石英衬底或者在硅衬底上生长的 SiO_2 层，并进行高温热退火制备的。实验上，一个峰值波长位于 800nm 的宽发射带被观测到，这正是硅纳米结构中的载流子的辐射再复合特征，吸收测量在近红外区域出现一个吸收带和在短波长区域的上升吸收边。上升边源于纳米晶的量子受限态的吸收，相反近红外吸收带则属于 $\text{Si}-\text{O}$ 界面态的特征。发射带和界面态的吸收带的重合，断定在硅纳米晶上的光发射通过与纳米晶-氧

界面相联系的辐射态而出现的。ps 脉冲激发下的时间分辨发光研究表明，具有很快的上升时间（约 ns 量级）。发光的衰减时间依赖于激发能量，大致为 μs 量级。

对光放大最直接的证据，可以在泵浦条件下探测信号的透射。为了达到放大所需的布居反转，用波长为 390nm 的强激光泵浦样品。波长为 800nm 的弱探测信号单次通过激活层。光增益测量的实验结果表明当探测光束通过被激发的纳米硅晶时被放大。

为了解释在纳米硅晶中的光放大（光增益）的实验结果，提出了一个三能级分子轨道模型，如图 13-15。两个能级分别对应于最低空闲分子轨道（或者导带底）和最高被占据分子轨道（或者价带顶），第三个能级是处于导带和价带中间的辐射界面态。光激励使得电子填充导带，相反使价带倒空，而后电子从导带快速地（约 ns 量级）弛豫到界面态。电子在界面态具有较长的寿命（约 μs 量级），这样使得在界面态和价带之间的布居反转成为可能。该模型可以解释为什么其他硅基系统中自由载流子吸收导致的损耗通常超过受激发射。虽然在硅纳米晶中用 390nm 超短脉冲激发，在 800nm 波长实现了光放大，但是尚未实现真正意义上的激光振荡。还有许多理论和技术上的问题需要解决。不过，令人欣慰的是毕竟为实现硅基激光器的研究开辟了一条新途径，也许距最终实现激光振荡仅一步之遥。

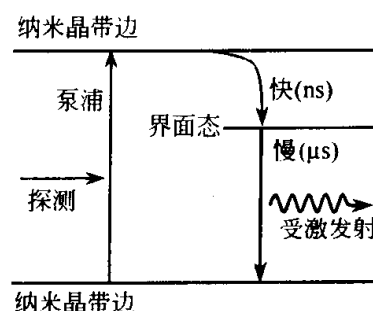


图 13-15 硅纳米晶能级模型

13.3 激光频率变换

本章第 1 节已讨论过线性介质中的光波传播问题。线性介质中极化强度正比于极化电场，所以在线性介质中即便有两个或两个以上光波同时存在，相互间也不会发生耦合。本节将处理在非线性介质中光波频率变换，这是与二阶非线性相关的物理过程。

13.3.1 非线性极化

在光波场的作用下，电介质的价电子偏离平衡位置形成偶极子，此现象称为极化作用。因光波场是交变的，当然极化也应是交变的而形成极化波，产生偶极辐射。辐射的性质由极化强度决定。如果极化强度与光场成正比，此种极化为线性极化；但当光波场足够强时，极化强度还含有光场的二次、三次项的贡献，此种为非线性极化。那么可以将极化强度 \mathbf{P} 分为线性和非线性两部分

$$\mathbf{P} = \mathbf{P}_L + \mathbf{P}_{NL} = \mathbf{P}_L + \mathbf{P}^{(2)} + \cdots, \mathbf{P}^{(2)} = \epsilon_0 \chi^{(2)} \mathbf{E} \mathbf{E}^* \quad (13-48)$$

$\chi^{(2)}$ 表示二阶介电响应。总场 \mathbf{E} 表示为其傅里叶分量（即不同激光频率）的叠加，那么非线性极化将由不同频率组合的几项组成。例如，对圆频率为 ω_1 和 ω_2 的两个场，二阶非线性极化应包含圆频率为 $2\omega_1$, $2\omega_2$, $\omega_1 + \omega_2$, $\omega_1 - \omega_2$ 和零频直流分量。考虑所产生的圆频率为 $\omega_3 = \omega_1 + \omega_2$ 的二阶非线性极化，那么其极化复振幅的第 i 个分量可以表示为

$$P_i^{(2)}(\omega_3) = \epsilon_0 D \sum_{jk} \chi_{ijk}^{(2)}(-\omega_3, \omega_1, \omega_2) E_j(\omega_1) E_k(\omega_2) \quad (13-49)$$

对未知场 $D=1$ ，而对可知场 $D=2$ ； $\chi_{ijk}^{(2)}$ 表征介质的二阶非线性，称为二阶非线性极化率。从式 (13-49) 可知，若 $E_j(\omega_1)E_k(\omega_2)$ 乘积的次序互换，并不影响二阶非线性极化强度的第 i 个分量的数值与符号，所以 $\chi_{ijk}^{(2)}$ 具有 $\chi_{ijk}^{(2)}(-\omega_3, \omega_1, \omega_2) = \chi_{ikj}^{(2)}(-\omega_3, \omega_2, \omega_1)$ 这样的固有互换对称性。在非线性光学中，通常 $\chi_{ijk}^{(2)}$ 用非线性光学系数 d ， d 也是一个张量，且有 $d_{ijk} = \chi_{ijk}^{(2)}/2$ 。固有互换对称性使 d 系数的最后两个下标 jk 可缩写为一个 l ，即 $d_{ijk} \rightarrow d_{il}$ ，这样 d 系数就简化为 3×6 矩阵

$$d_{il} = \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \end{bmatrix} \quad (13-50)$$

d 系数的字母与数字替换规则按如下

$$\begin{aligned} i: & x \rightarrow 1 \quad y \rightarrow 2 \quad z \rightarrow 3 \\ jk \rightarrow l: & xx \rightarrow 1, \quad yy \rightarrow 2, \quad zz \rightarrow 3, \quad yz \text{ (zy)} \rightarrow 4, \quad zx \text{ (xz)} \rightarrow 5, \\ & xy \text{ (yx)} \rightarrow 6 \end{aligned} \quad (13-51)$$

例如： $d_{xyz} = d_{xzy} = d_{14}$ ， $d_{zxx} = d_{31}$ ， $d_{yzx} = d_{yxz} = d_{25}$ ， $d_{xxy} = d_{xyx} = d_{36}$ 。对于典型的晶体对称类型的 d_{il} 系数非零矩阵元见表 13-1。在反演对称的晶体中，所有的 d_{il} 或 $\chi_{ijk}^{(2)}$ 必须恒等于零。此乃中心反演对称晶体中不会有二阶非线性效应的原因。

表 13-1 典型的晶体对称类型的 d_{il} 系数非零矩阵元

晶 系	点 群	非 零 矩 阵 元
三斜		均为非零的
单斜	$2//y$	$d_{14}, d_{16}, d_{21}, d_{22}, d_{23}, d_{25}, d_{34}, d_{36}$
	$2//z$	$d_{14}, d_{15}, d_{24}, d_{25}, d_{31}, d_{32}, d_{33}, d_{36}$
	$m(m \perp y)$	$d_{11}, d_{12}, d_{13}, d_{15}, d_{24}, d_{26}, d_{31}, d_{32}, d_{33}, d_{35}$
	$m(m \perp z)$	$d_{11}, d_{12}, d_{13}, d_{16}, d_{21}, d_{22}, d_{23}, d_{26}, d_{34}, d_{35}$
正交 (斜方)	222	d_{14}, d_{25}, d_{36}
	$2mm$	$d_{15}, d_{24}, d_{31}, d_{32}, d_{33}$

续表

晶 系	点 群	非 零 矩 阵 元
正方	4	$d_{14}, d_{15}, d_{24} = d_{15}, d_{25} = -d_{14}, d_{31}, d_{32} = d_{31}, d_{33}$
	$\bar{4}$	$d_{14}, d_{15}, d_{24} = -d_{15}, d_{25} = d_{14}, d_{31}, d_{32} = -d_{31}, d_{36}$
	422	$d_{14}, d_{25} = -d_{14}$
	4mm	$d_{15}, d_{24} = d_{15}, d_{31}, d_{32} = d_{31}, d_{33}$
	$\bar{4}2m$	$d_{14}, d_{25} = d_{14}, d_{36}$
立方	432	均为零
	$\bar{4}3m$ 和 23	$d_{14}, d_{26} = d_{14}, d_{36} = d_{14}$
三角	3	$d_{11}, d_{12} = -d_{11}, d_{14}, d_{15}, d_{16} = -d_{22}, d_{21} = d_{22}, d_{22}, d_{24} = d_{15},$ $d_{25} = -d_{14}, d_{26} = -d_{11}, d_{31}, d_{32} = d_{31}, d_{33}$
	32	$d_{11}, d_{12} = -d_{11}, d_{14}, d_{25} = -d_{14}, d_{26} = -d_{11}$
	3m ($m \perp x$)	$d_{15}, d_{16} = -d_{22}, d_{21} = -d_{22}, d_{22}, d_{24} = d_{15}, d_{31}, d_{32} = d_{31}, d_{33}$
	3m ($m \perp y$)	$d_{11}, d_{12} = -d_{11}, d_{15}, d_{24} = d_{15}, d_{26} = -d_{11}, d_{31}, d_{32} = d_{31}, d_{33}$
六角	6 与正方 4 同	$d_{14}, d_{15}, d_{24} = d_{15}, d_{25} = -d_{14}, d_{31}, d_{32} = d_{31}, d_{33}$
	6mm 与正方 4mm 同	$d_{15}, d_{24} = d_{15}, d_{31}, d_{32} = d_{31}, d_{33}$
	622 与正方 422 同	$d_{14}, d_{25} = -d_{14}$
	$\bar{6}$	$d_{11}, d_{12} = -d_{11}, d_{16} = -d_{22}, d_{21} = -d_{22}, d_{22}, d_{26} = -d_{11}$
	$\bar{6}m2$ ($m \perp x$)	$d_{16} = -d_{22}, d_{21} = -d_{22}, d_{22}$
	$\bar{6}m2$ ($m \perp y$)	$d_{11}, d_{12} = -d_{11}, d_{26} = -d_{11}$

13.3.2 块状晶体中的光频变换

(1) 非线性介质中耦合波方程 强光波在非线性介质中传播时会产生非线性极化，并由此导致不同频率光波间的能量和动量交换，实现光波耦合。基于二阶非线性可以实现二次谐波产生或倍频 (SHG)，和频产生 (SFG)，差频产生 (DFG)，参量产生 (OPG)，参量放大 (OPA) 以及参量振荡 (OPO) 等非线性耦合过程。支配光波耦合过程的是源于麦克斯韦方程的耦合波方程。在非磁性、非导电的介质中，耦合波方程为

$$\nabla^2 \mathbf{E} = \mu_0 \epsilon_0 \partial^2 \mathbf{E} / \partial t^2 + \mu_0 \partial^2 \mathbf{P} / \partial t^2 \quad (13-52)$$

总极化强度由线性和非线性两部分组成。由于在此关心的是基于二阶非线性的光波耦合问题，因此非线性极化只取到二阶。那么，波方程则可改写为

$$\nabla^2 \mathbf{E} = \mu_0 \epsilon_0 \partial^2 (\epsilon \mathbf{E}) / \partial t^2 + 2 \partial^2 (\mathbf{d} \mathbf{E} \mathbf{E}) / \partial t^2 \quad (13-53)$$

该方程右边第一项描述的是光波场在线性介质中传播，其通解为一系列独立的不同频率平面波的线性组合，各频率光波场间无任何形式的能量转移和动量交换；第二项则为非线性相互作用引起的非线性极化，导致在新频率上的光波辐射。

假定参与相互作用频率为 $\omega_1, \omega_2, \omega_3$ 的三个单色光波场，均为沿 z 方向传播，那么总光波场为

$$\mathbf{E} = \sum_m \mathbf{E}_m = \frac{1}{2} \sum_m [\mathbf{e}_m E_m(z) \exp(ik_m z - i\omega_m t) + c.c.], m = 1, 2, 3 \quad (13-54)$$

式中, E_m , \mathbf{e}_m 分别为相应的光波场的复振幅和偏振单位矢量。用新参数 A_m 取代 E_m

$$A_m(z) = \sqrt{\frac{n_m}{\omega_m}} E_m(z), m = 1, 2, 3 \quad (13-55)$$

通常, 研究基于二阶非线性的光频变换, 总是具有特定偏振方向光波作用于晶体上, 并在特定偏振方向上产生非线性极化; 同时考虑到 \mathbf{d} 系数的对称性及能量守恒 $\omega_1 + \omega_2 = \omega_3$ 。将式 (13-54) 代入式 (13-53), 同时考虑到式 (13-55), 在慢变振幅近似下, 则有

$$\begin{aligned} dA_1(z)/dz &= i(\gamma/2)A_2^*(z)A_3(z)\exp(i\Delta kz) \\ dA_2^*(z)/dz &= -i(\gamma/2)A_1(z)A_3^*(z)\exp(-i\Delta kz) \\ dA_3(z)/dz &= i(\gamma/2)A_1(z)A_2(z)\exp(-i\Delta kz) \end{aligned} \quad (13-56)$$

式中, γ 为耦合系数, d_{eff} 为有效非线性系数

$$\gamma = d_{eff} \sqrt{\omega_1 \omega_2 \omega_3 / n_1 n_2 n_3} / c, d_{eff} = \mathbf{e}_3 \cdot \mathbf{d} : \mathbf{e}_1 \mathbf{e}_2 \quad (13-57)$$

Δk 为波矢失配

$$\Delta k = k_3 - k_1 - k_2 \quad (13-58)$$

方程 (13-56) 即为耦合波方程, 是描述基于二阶非线性的三波相互作用的基本方程。

(2) SHG、SFG 和 DFG 若忽略波耦合过程中泵浦光转换为新输出波所致的损耗, 即视泵浦光的强度为常数。对于 SHG, SFG 和 DFG 过程的转换效率可以通过求解上面的耦合波方程得到

$$\eta_{SHG}, \eta_{SFG}, \eta_{DFG} \propto d_{eff}^2 \sin^2(\Delta k l / 2) l^2 I \quad (13-59)$$

式中, l 为介质的长度, I 分别为基波 (对 SHG) 的强度、强泵浦光 (对 SFG 和 DFG) 的强度。

当 $\Delta k \neq 0$, 即位相失配时, 转换效率正比于泵浦光的强度, 因此提高泵浦功率可以提高转换效率; 同时也发现, 转换效率随介质长度呈现周期性变化 (见后面的图 13-16), 而且极大转换效率出现在一系列的 $l = (2m + 1) l_c$ 处, $m = 0, 1, \dots$, $l_c = \pi / \Delta k$, 被称为相干长度。可见, 当位相失配时不能通过增加 l 来提高转换效率, 惟一的途径是增加泵浦光强度。但是强度的增加也是有限的, 因为强度过高会导致介质的永久性损伤。

当 $\Delta k = 0$, 即位相匹配时。从式 (13-59) 可知, 转换效率不仅仍正比于泵浦强度, 更重要的是正比于 l^2 。这样增加 l , 可以大大提高转换效率。不

过, 当转换效率变得比较可观时, 将不再随 l^2 增长; 此时, 必须计及泵浦的耗尽。虽然上述由小信号近似得到的转换效率公式已不再适用, 但却从中却得到了最普适的适用于任何的波耦合过程的重要结论, 即为实现有效的光频变换, 最有效途径就是满足位相匹配条件。

(3) 位相匹配 位相匹配对于实现有效的光频变换是必不可少的条件。在此讨论块状晶体中实现位相匹配的方法。

① 角度位相匹配技术。又分为标量的和矢量的。标量即参与相互作用所有光波是共线传播的, 是最普遍的情况。矢量则是非共线传播的 (在此不予以讨论)。位相匹配意味着, 期望波与驱动它的非线性极化波的相速度匹配。这些条件对于各种光频变换过程概括为

$$\mathbf{k}_\omega + \mathbf{k}'_\omega = \mathbf{k}_{2\omega} \text{ (SHG)}, \quad \mathbf{k}_{p1} + \mathbf{k}_{p2} = \mathbf{k}_s \text{ (SFG)}, \quad \mathbf{k}_{p1} - \mathbf{k}_{p2} = \mathbf{k}_d \text{ (DFG)} \quad (13-60)$$

对于标量的 (即共线的) 情况, 它们相应地简化为

$$n_{2\omega} - (n_\omega + n'_\omega) / 2 = 0 \quad \text{(SHG)} \quad (13-61)$$

$$\omega_{p1}(n_s - n_{p1}) + \omega_{p2}(n_s - n_{p2}) = 0 \quad \text{(SFG)} \quad (13-62)$$

$$\omega_{p1}(n_{p1} - n_d) - \omega_{p2}(n_{p2} - n_d) = 0 \quad \text{(DFG)} \quad (13-63)$$

对于 SHG 过程, $\mathbf{k}_\omega (n_\omega)$ 和 $\mathbf{k}'_\omega (n'_\omega)$ 可以相同, 也可以不同 (例如, 偏振不同但频率相同)。由于色散的存在, 位相匹配条件通常不能自动被满足。但在具有双折射的各向异性介质中, 在特殊的配置情况下, 通过选择不同偏振组合并在适当的传播方向上可实现位相匹配。位相匹配又分为: 第 I 类, 两个泵浦波具有相同的偏振 (如 $o+o \rightarrow e$, $e+e \rightarrow o$); 第 II 类, 两个泵浦波具有正交的偏振 ($o+e \rightarrow o$, $o+e \rightarrow e$, $e+o \rightarrow o$, $e+o \rightarrow e$)。如单轴晶体中的 SHG 过程。对于负单轴晶体 ($n_o > n_e$), 若 $n_e^{2\omega} < n_o^\omega$, 定存在一个角度 θ_{pm} (被称为匹配角) 使得 $n_e^{2\omega}(\theta_{pm}) = n_o^\omega$; 若基波为 o 偏振并沿 θ_{pm} 方向传播, 定会在该传播方向产生 e 偏振的二次谐波 (即 $o+o \rightarrow e$), 而不可能有 $e+e \rightarrow o$ 。将 $n_e^{2\omega}(\theta_{pm}) = n_o^\omega$ 代入式 (13-23) 得到负单轴晶体中 SHG ($o+o \rightarrow e$) 的匹配角 θ_{pm} 为

$$\sin^2 \theta_{pm} = \frac{(n_e^{2\omega})^2 (n_o^{2\omega})^2 - (n_o^\omega)^2}{(n_o^\omega)^2 (n_o^{2\omega})^2 - (n_e^{2\omega})^2} \quad (13-64)$$

对正单轴晶体 ($n_o < n_e$), 若 $n_o^{2\omega} > n_e^\omega$, 也存在某一个角度 θ_{pm} 使得 $n_e^\omega(\theta_{pm}) = n_o^{2\omega}$ 。需注意: 正单轴晶体中只存在 $e+e \rightarrow o$ 配置。容易得到正单轴晶体中 SHG ($e+e \rightarrow o$) 的匹配角 θ_{pm} 为

$$\sin^2 \theta_{pm} = \frac{(n_e^\omega)^2 (n_o^{2\omega})^2 - (n_o^\omega)^2}{(n_o^{2\omega})^2 (n_e^\omega)^2 - (n_o^\omega)^2} \quad (13-65)$$

对角度匹配, 通常 $\theta_{pm} \neq 90^\circ$, 被称为临界位相匹配。在各向异性介质中, 异常波的波矢与其能流方向并非一致。这样对非临界位相匹配, 有限光束口径的寻常和异常光束将不可能在非线性介质中总是完全重叠, 这一效应被称为走离效应 (walk-off effect)。在物理上, 不重叠的光束将不可能发生相互作用, 因此走离效应对转换效率起了严重的限制作用。如果当匹配角 $\theta_{pm} = 90^\circ$ 时, 被称为非临界位相匹配。

② 温度位相匹配技术。在有些晶体中, 折射率显著地依赖于温度。所以, 可以通过改变温度, 实现前面提到的非临界位相匹配。那么式 (13-60) 中的波矢匹配条件, 对非临界位相匹配的情况, 可写为

$$n_{2\omega}(T_{pm}) = \frac{1}{2} [n_{\omega}(T_{pm}) + n'_{\omega}(T_{pm})] \quad (\text{SHG}) \quad (13-66)$$

$$\omega_s n_s(T_{pm}) = \omega_{p1} n_{p1}(T_{pm}) + \omega_{p2} n_{p2}(T_{pm}) \quad (\text{SFG}) \quad (13-67)$$

$$\omega_d n_d(T_{pm}) = \omega_{p1} n_{p1}(T_{pm}) - \omega_{p2} n_{p2}(T_{pm}) \quad (\text{DFG}) \quad (13-68)$$

位相匹配温度 T_{pm} 可以通过折射率与波长和温度的变化关系, 求解找到。非临界位相匹配的优点是: 对光束的发散性不敏感的; 走离角为零。

(4) 光参量过程 前面讨论了 SHG、SFG 和 DFG 过程, 可以用于产生新频率的相干辐射, 但波长不能调谐。然而, 在诸如光谱学、半导体材料表征、光化学、遥感等领域, 则需要可调谐相干光源。非线性光学参量器件恰可满足此要求。光参量过程包括: 光参量产生 (OPG)、光参量放大 (OPA) 和光参量振荡 (OPO)。

① 光参量产生 (OPG)。也被称为参量荧光。OPG 的物理图像就是: 一个频率为 ω_p 的光子入射在非线性介质上并与之相互作用, 分裂为频率分别为 ω_s 和 ω_i 的信号和闲频两个低频光子 ($\omega_p = \omega_s + \omega_i$)。严格地讲, OPG 是量子效应, 并首次被 Louisell 等预言。最初产生的信号和闲频辐射总是试图出现在所有可能的方向, 惟一的限制是能量守恒。虽然能量守恒允许信号光出现在任何方向和从 0 到 ω_p 范围内的任一频率, 但是信号功率的峰一定出现在满足位相匹配条件的频率和方向上。通常, 参量荧光的带宽 (或线宽) 和发射方向都是较宽的, 因为它毕竟属于自发发射。

② 光参量放大 (OPA)。OPA 与 OPG 不同之处在于: 前者需要一个强泵浦光和一个弱信号光同时作用于非线性介质上, 而后者则仅有一个泵浦光。由于强泵浦光和弱信号光在晶体内相互作用, 将产生闲频光。实质上, 这是一个 DFG 过程, $\omega_i = \omega_p - \omega_s$; 新产生的 ω_i 又会与原泵浦光 ω_p 相互作用出现 $\omega_p - \omega_i = \omega_s$ 过程, 即另一种 DFG 过程 $\omega_s = \omega_p - \omega_i$ 。这样的正反馈过程使得泵浦光能量不断转换为信号光和闲频光, 原来入射的弱信号光被放大, 此过程被称为 OPA。OPA 虽然实际上也是一种差频过程, 但是与前面

讨论过的差频过程还是有所区别的。OPA 关心的是被放大的信号光；而差频过程关心的则是另一个被产生的差频光，同时差频过程的两个输入泵浦光的强度也接近。位相匹配对 OPA 过程同样重要。此时位相匹配条件 $k_p = k_i + k_s$ ，与 DFG 实际上完全相同。

③ 光参量振荡 (OPO)。在 OPA 系统中参量光是单程通过非线性晶体，将泵浦光功率转化为信号光功率和闲频光功率。若将非线性晶体置于光学谐振腔内，当光参量增益超过光损耗时，振荡将出现，这样的器件被称为 OPO。显然，可以视其为一种新型激光器，其输出的信号光和闲频光都是具有高度相干性的激光。但是，在机理上与上面提到的激光器有着本质的区别。对于 OPO，泵浦光必须是激光，而且输出的信号光和闲频光的特性与泵浦的性质密切相关，而与介质的线宽和能级结构无关。泵浦光的单色性越好，相干性和方向越佳，得到的信号光和闲频光的单色性和方向性也越好；原因是信号光和闲频光的源是泵浦光与晶体非线性相互作用，产生的参量荧光。激光器则不同，激光增益介质吸收泵浦光的能量，实现激光能级间的布居反转而获得增益，激光器输出的源是激光上下能级跃迁产生的辐射，因此辐射只与激光增益介质的能级结构有关，而与泵浦光的性质无关；激光器的泵浦光可以是非激光的光源。OPO 又分为双共振和单共振两种。双共振，即信号光和闲频光同时在腔内形成振荡；其优点是振荡阈值较低，其缺点是温度或声振动引起的腔长微小的起伏将导致输出频率和功率的不稳定（原因在于能量守恒要求的信号光和闲频光的频率与腔模所要求的振荡模的频率不匹配所致）。单共振，仅信号光在腔内共振；单共振可以克服双共振时不稳定的弱点（因为只有信号光频率是共振的，而对闲频光频率无任何限制，通过能量守恒闲频光的频率可以自由地连续变化，微小腔长的起伏不会引起输出的信号光和闲频光频率以及功率大的变化）。

对非线性光学晶体，我国的研究者开发出如 BBO、LBO、LAP 等具有重要科学价值的晶体，对激光器件与技术的发展作用做出了极为重要的贡献。

13.3.3 光学超晶格中的光频变换

至此所阐述的都是在常规块状晶体中的非线性光频变换过程，利用材料的双折射实现有效光频变换所要求的位相匹配，但其存在一些缺点。例如不能在某些方向和在其整个透明波段实现位相匹配；此外，由于位相匹配要求参与相互作用的光波必须具有特定的偏振方向，使得介质的最大非线性系数张量元难以被利用，而限制转换效率。在此，将讨论基于准位相匹配原理的光学超晶格中的激光频率变换。

准位相匹配 (quasi-phase matching, QPM) 理论虽然早在 1962 年就已提出，但直至 1980 年才在实验上首次被南京大学研究组验证。他们生长

出当时被称为聚片多畴的 LiNbO_3 晶体。而后，由于室温电场极化技术上的重大突破，并借助于光刻技术，使得可以按预先设计的图案制备反转畴结构成为可能。Yamada 等人于 1993 年利用该技术成功的制备出周期极化 LiNbO_3 。而后，南京大学研究组，利用该技术成功制备出周期极化 LiTaO_3 晶体。由于 LiTaO_3 短波吸收边比 LiNbO_3 更短，所以在短波段前者更具竞争力。最近，已利用其他铁电晶体，如 KTP, SBN, RAT (RbTiOAsO) 成功制备出周期极化样品。由于这种周期极化的人工微结构材料与半导体超晶格和金属超晶格（由于它们的调制周期在纳米尺度，也称为纳米超晶格）在结构上的类似性，同时此种微结构材料用于经典光波的非线性频率变换，且其周期具有光波长量级，所以称其为光学超晶格。

QPM 概念的基本物理思想是容易理解的，波矢匹配是通过一种全新的非临界匹配方法（借助于倒格矢—reciprocal vectors）实现的。如前所述，在波矢失配时二次谐波场正比于 $d_{\text{eff}} \sin c(\Delta k l / 2) l$ ，所以二次谐波场随晶体长度的增加而呈现周期变化（其周期为相干长度 $l_c = \pi / |\Delta k|$ 的两倍），而非单调增长；即在此情况下，增加晶体长度并不能使二次谐波得到增大。如图 13-16，在第一个相干长度内，二次谐波场随长度 l 的增加单调增长，而进入第二个相干长度范围则下降。可以想像，如果能通过某一方法使得在第二个相干长度内，有效线性系数 d_{eff} 为负的，那么二次谐波也同样随长度 l 的增加而增长。以此类推，如果将所有的第偶数次相干长度范围内的畴反转使 d_{eff} 为负的，而保持所有的第奇数次相干长度范围内的 d_{eff} 符号不变；这样就构成周期为两倍相干长度的正负交替的周期性畴反转微结构材料—周期光学超晶格。自然地可以维持二次谐波随晶体长度的增加而持续单调增长，这

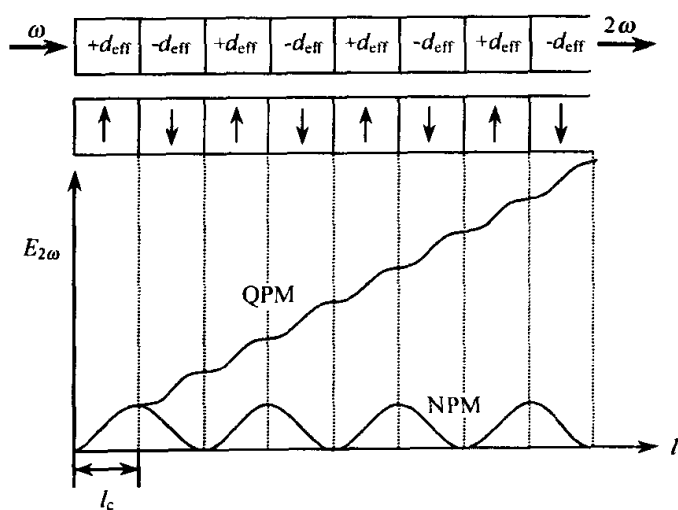


图 13-16 准位相匹配概念

就是 QPM 概念的直观物理解释。

光波在周期光学超晶格中的传播类似于电子在周期势调制的晶体中的运动, 所以可以用倒格矢概念来处理光波在这样的微结构材料中的传播。QPM 就是利用这种畴被周期调制的微结构材料提供的某一倒格矢 G_m 去补偿二次谐波与基波间的波矢失配, 即

$$k_{2\omega} - k_{\omega} - G_m = 0 \quad (13-69)$$

这与利用双折射实现的位相匹配原理有着本质区别。从傅里叶分析的观点可知: 实空间的一维周期函数与倒格矢空间的一维函数是一对傅里叶变换对; 一维的周期结构可以提供一系列等间隔的分立倒格矢, 这些倒格矢中的任何一个都是基元倒格矢的整数倍。这样一个周期结构提供的所有倒格矢仅由一个整数和一个结构参数 (即周期) 决定, 可调参数实质上只有结构参数一个。所以, 在一维周期光学超晶格中, 通常只能提供一个有效倒格矢, 换句话说通常仅有一个有效的光频交换过程。最初, 对光学超晶格的研究主要集中在周期结构。准晶的发现在凝聚态物理开辟了一个新的领域, 引起广泛关注。一维周期结构的光学超晶格概念被扩展到一维准周期的光学超晶格 (即准晶在一维的投影), 并用于激光频率变换, 使准晶找到了有实际应用。进而又扩展到双周期、三周期乃至非周期结构, 使得光学超晶格概念更普适, 赋予了更丰富的物理内涵并拓宽了其实际应用。同时发展了新的设计方法, 完善了制备技术, 许多新效应被验证。一维准周期光学超晶格在空间对称性上低于周期结构, 但高于无序结构; 其倒格矢由两个整数和两个结构参数决定, 不像在一维周期光学超晶格中仅由一个整数和一个结构参数决定。这样在结构设计上具有更多的自由度, 所以准周期光学超晶格可提供更多的有效倒格矢实现多个光频变换过程。典型的准周期光学超晶格是按 Fibonacci 序列构筑的, 两个基本组元 A 和 B , 并按规则 $S_{j+2} = S_{j+1}S_j$ ($j \geq 1$, $S_1 = A$, $S_2 = AB$) 排列。每个基本组元由方向相反的正负畴组成, A 和 B 的宽度为 l_A 和 l_B , 且 $l_A = l_{A+} + l_{A-}$ 和 $l_B = l_{B+} + l_{B-}$ 。假定 A 和 B 的正畴宽度有 $l_{A+} = l_{B+} = l$, 负畴 $l_{A-} = l(1 + \eta)$ 和 $l_{B-} = l(1 - \tau\eta)$ 。 l 和 η 是两个可调参数, τ 为黄金分割比。图 13-17 (a) 为 Fibonacci 准周期光学超晶格结构, 图 13-17 (b) 为一个 Fibonacci 超晶格晶体样品的畴照片。

在这样的准周期光学超晶格中, 提供的倒格矢和相应的系数大小为

$$G_{m,n} = 2\pi(m + \tau n)/D, f_{m,n} = 2\text{sinc}(G_{m,n}l/2)\text{sinc}(Z_{m,n}) \quad (13-70)$$

在此, m 和 n 是两个整数, $D = \tau l_A + l_B$ 为平均结构参量, $Z_{m,n} = \pi D^{-1}(1 + \tau)(ml_A - nl_B)$ 。可见, 倒格矢由两个结构参量决定。因此, 该准周期结构至少可以提供两个有效倒格矢满足两个耦合过程所要求的 QPM 条件。

利用一个准周期光学超晶格可以通过其中的倍频与和频两个耦合过程实

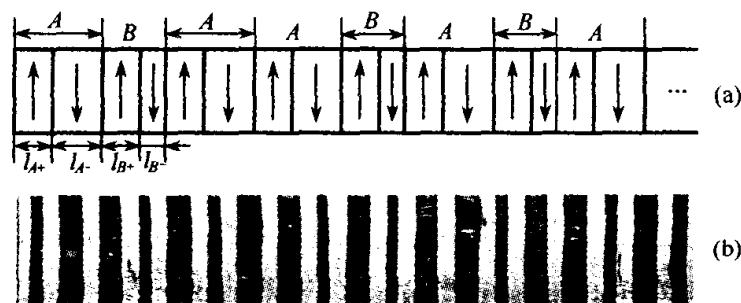


图 13-17 Fibonacci 序列准周期光学超晶格

(a) Fibonacci 序列准周期光学超晶格结构；(b) 相应的畴照片

现直接 THG。例如，采用电场极化技术制备出 Fibonacci 序列的 LiTaO_3 准周期光学超晶格，其结构参数为 $l = 10.7\mu\text{m}$ ， $\eta = 0.23$ 。对基本组元 A 和 B， $l_A = l_{A+} + l_{A-} = 10.7 + 13.2 = 23.9\mu\text{m}$ 和 $l_B = l_{B+} + l_{B-} = 10.7 + 6.7 = 17.4\mu\text{m}$ ，总长 $\approx 8\text{mm}$ ，厚度 $\approx 0.5\text{mm}$ ，其畴照片如图 13-17 (b)。在此设计中倒格矢 $G_{1,1}$ 被用于基波的倍频 (SHG)， $G_{2,3}$ 被用于基波与倍频的和频 (即 THG)，两个过程间的相互耦合使能量连续地从基波转移到 SHG 再到 THG。这样一个准周期光学超晶格可以直接实现高效 THG。实验中基波来自于重复频率 10Hz、脉宽约 8ns、线宽约 1.5nm 的 OPO，其偏振方向沿畴方向；超晶格的设计使偏振方向沿畴方向的 SHG 和 THG 同时满足 QPM，目的是利用 LiTaO_3 最大的非线性系数张量元 d_{33} 。当基波调谐至 1570nm 时，观测到稳定的 785nm 的 SHG 和 523nm 的 THG，THG 的转换效率达 23%。

准周期光学超晶格决不仅限于 Fibonacci 序列；除准周期结构之外，还有如双周期，多周期，乃至非周期结构。对于实际应用，超晶格的设计是以满足所要求光频变换过程为前提，尽可能选择结构简单、易于制备、效率高为原则。在光学超晶格中不仅可以实现直接 THG，也可以实现直接 FHG、OPO 以及各种级联过程。在光学超晶格发展中，最重要的是突破了周期结构的束缚，使得光学超晶格的设计自由度更多，为光频变换提供了一条非常有效的新途径。由于设计方法的完善和制备技术的成熟，光学超晶格已从最初的原理验证，发展为现在的原型器件，不久的将来将走向实用化，有望成为有竞争力的新型光频变换系统。这主要是因为光学超晶格具有如下优点：几乎在非线性材料的整个透明波段均可实现 QPM；可以在弱双折射材料或各向同性的材料中实现 QPM；可以随意选择相互作用光波的偏振方向；利用材料的最大非线性系数张量元；可以同时实现多个期望的光频变换过程；根据实际需要设计光学超晶格的微结构；无走离效应。

13.4 非线性折射率与非线性吸收

前一节介绍了与材料的二阶光学非线性相关的光频转换。本节将阐述三阶光学非线性及与其密切相关的现象。三阶光学非线性，与二阶截然不同，不受材料固有对称性的制约；原理上，任何材料均具有三阶非线性，只是其强弱因材料而异。材料的三阶光学非线性可以用宏观量——三阶非线性极化率 $\chi^{(3)}$ 和微观量——二阶超极化率表征。非线性折射率与 $\chi^{(3)}$ 的实部相关，而非线性吸收则与 $\chi^{(3)}$ 的虚部相联系。

13.4.1 非线性折射率

(1) 非线性折射率的一般描述 介质折射率与光波场的依赖关系一般可表达为

$$n(\mathbf{r}, t) = n_0(\mathbf{r}, t) + \Delta n[I(\mathbf{r}, t)] \quad (13-71)$$

折射率的变化 Δn 是光强 $I(\mathbf{r}, t)$ 的函数，这一函数关系对不同的非线性介质和物理机制具有不同的具体表达式。

① 类克尔介质。在这种介质中非线性折射率可以写为 $\Delta n = n_2 \langle \mathbf{E} \cdot \mathbf{E} \rangle$ ，其中 n_2 为非线性折射率系数，尖括号表示至少在一个光振动周期内的时间平均， \mathbf{E} 是一个快速变化的总实光波场。由于 $I \propto |\mathbf{E}|^2$ ，折射率的变化也可以用光强表示为 $\Delta n = n_2^I I$ ，其中 n_2^I 为非线性折射率强度系数。从各向同性介质的对称性考虑，非零的独立的 $\chi^{(3)}$ 矩阵元仅有 $\chi_{iiii}^{(3)}$ ， $\chi_{ijij}^{(3)}$ ， $\chi_{ijji}^{(3)}$ ；而且对角矩阵元为 $\chi_{iiii}^{(3)} = \chi_{ijij}^{(3)} + \chi_{ijji}^{(3)} + \chi_{ijji}^{(3)}$ 。对于一个线偏振光或两个平行线偏振光作用于非线性介质时， n_2 正比于 $\chi_{iiii}^{(3)}$ 。若 $\chi^{(3)}$ 存在非对角矩阵元，表明在各向同性介质中通过三阶非线性极化诱导双折射的可能。这将在椭圆偏振光或两个非平行线偏振光的情况出现。 n_2 、 n_2^I 与 $\chi^{(3)}$ 三者关系如表 13-2。

表 13-2 n_2 、 n_2^I 与 $\chi^{(3)}$ 的关系

单光束线偏振 $\Delta n = 2n_2 \mathbf{E} ^2$	单光束圆偏振 $\Delta n_{\pm} = 2n_2^{cir} \mathbf{E}_{\pm} ^2$	单光束椭圆偏振
$n_2 = \frac{3}{4n_0} \chi_{xxxx}^{(3)}(-\omega; \omega, \omega, -\omega)$ $= \frac{3}{4n_0} (2\chi_{xxyy}^{(3)} + \chi_{xyyx}^{(3)})$	$n_2^{cir} = \frac{3}{2n_0} \chi_{xxyy}^{(3)}(-\omega; \omega, \omega, -\omega)$	$\Delta n_{\pm} = \frac{3}{n_0} [\chi_{xxyy}^{(3)} \mathbf{E}_{\pm} ^2 + (\chi_{xxyy}^{(3)} + \chi_{xyyx}^{(3)}) \mathbf{E}_{\mp} ^2]$
单光束强度系数	双光束(强泵浦 ω 和弱探测 ω')平行线偏振	
$\Delta n = n_2^I I, n_2^I = \frac{n_2}{\epsilon_0 n_0 c}$	$\Delta n(\omega') = 2n_2^{(weak)} \mathbf{E} ^2, n_2^{(weak)} = \frac{3}{2n_0} \chi_{xxxx}^{(3)}(-\omega'; \omega', \omega, -\omega)$	

② 非类克尔介质。在非类克尔介质中非线性折射率变化依赖于光强，但依赖关系不是简单的线性关系，而是光强的非线性函数，即 $\Delta n = f(I)$ 。

对于足够小的光强, 可以将其展开为光强 I 的幂级数, 其中主要项是 I 的线性项。对足够小的光强, 介质将是类克尔的, 而且非线性折射率变化仍然可以用 n_2 表示。对于大的光强则不可能。

③ 准类克尔介质。在此类材料中, 非线性折射率变化为光强的线性函数的表达式, 一个例子就是上面刚提到低光强时的非类克尔介质。另一个例子是介质中的物理变化, 即借助于来自光波的累积能量导致折射率的变化。在一些情况下, 可以将折射率的变化表示为光强的线性函数, 这样的例子如热的非线性和级联二阶非线性效应。

④ 强度调制效应介质。在此情况, 非线性折射率的变化与上面提到的情况完全不同, 与绝对光强无直接关系, 相反与光强的相对调制有关。折射率变化可以表示为 $\Delta n(\mathbf{r}) \propto \delta I(\mathbf{r})/I_0$, 其中 I_0 为总的空间平均光场强度, 而 δI 为光强的空间调制部分, 即 $I = I_0 + \delta I(\mathbf{r})$ 。最典型的例子就是光折变介质。

(2) 非线性折射率的物理机制 对非线性折射率的贡献主要有如下几个物理机制。

① 电子极化。这一机制涉及光波场引起的原子或分子的电子云畸变。如果原子或分子是高度可极化的, 那么它可以呈现巨大的电子非线性。非线性折射率与三阶光学非线性的实部相关联, 而后者又源于二阶超极化率 γ 且与其成正比。对于具有中心对称的材料, 三阶是最低阶的非线性。从量子光学的微扰理论, 可知二阶超极化率的表达式共有 48 项; 但是当光频接近于一个单光子或双光子共振时, 只有其中的几项对其贡献起主要作用。当光频 ω 小于系统的任一共振, 双光子共振项试图增加正的 n_2 ; 而另一方面单光子共振试图增加负的 n_2 。对于透明介质的非共振的电子过程, 二阶超极化率的典型值为 γ 约为 10^{-36} esu ($10^{-61} \text{C} \cdot \text{m}^4/\text{V}^3$), 但有机大分子可以有 γ 约为 (100~1000 倍) 10^{-36} esu。典型的 $\chi^{(3)}$ 约为 10^{-14} esu (约为 $10^{-22} \text{m}^2/\text{V}^2$) 和 $n_2 \sim 10^{-13}$ esu (约为 $10^{-22} \sim 10^{-21} \text{m}^2/\text{V}^2$)。一些聚合物系统, 非共振的 $\chi^{(3)}$ 约为 $10^{-12} \sim 10^{-10}$ esu (约为 $10^{-20} \sim 10^{-18} \text{m}^2/\text{V}^2$)。非共振的过程是很快, 典型时间响应约为 10^{-15} s。

② 喇曼诱导克尔效应。此效应与受激喇曼散射有关, 频率为 ω_L 强光束入射在喇曼激活介质上, 同时也有另一频率为 ω_S 弱光束入射在同一介质上, 此弱光束被称为斯托克斯光束 ($\omega_S < \omega_L$)。线性喇曼散射是一个双光子过程, 一光子被散射至一个低频的斯托克斯光子, 下移频率等于介质的喇曼激活模的频率。受激喇曼散射则是一个四光子过程。在此, 入射在喇曼激活介质上的弱斯托克斯光束被放大, 其增益来自于强光束。拉曼增益因子与拉曼极化率 $\chi_{\text{Raman}}^{(3)}$ 虚部相关 $G(\omega_S) = -(6\omega_S/n_S c) \text{Im}[\chi_{\text{Raman}}^{(3)}(-\omega_S; \omega_S, \omega_L,$

$-\omega_L)]|E_L|^2$ ，其中 n_S 为斯托克斯频率 ω_S 处的线性折射率。增益的频率依赖性显示出拉曼共振

$$G(\omega_S) = \frac{(\Gamma/2)^2 G_{\max}}{[\omega_S - (\omega_L - \omega_V)]^2 + (\Gamma/2)^2} \quad (13-72)$$

其中 G_{\max} 为峰值处的最大增益， Γ 为拉曼线宽， $\omega_L - \omega_V$ 为拉曼共振。非线性折射率系数可以用增益系数 g_S (g_S 为单位光强的增益因子) 表示为 $|n_2| \simeq \epsilon_0 n_0 c \lambda_S g_S / 8\pi$ 。需要注意的是：当斯克托斯频率通过共振频率时， $\chi_{\text{Raman}}^{(3)}$ 的实部将改变符号。对于典型的数值， g_S 约为 $3 \times 10^{-11} \text{ m/W}$ ， $n_L \sim n_S$ 约为 1.5， λ_S 约为 10^{-6} m ， $|\text{Re}[\chi_{\text{Raman}}^{(3)}]|$ 约为 $10^{-20} \text{ m}^2/\text{V}^2$ 。这比前面讨论过的典型的非共振电子极化率大两个量级。其响应时间是喇曼诱导材料激发的寿命量级，约为 10^{-12} s 。

③ 分子取向效应。各向异性分子构成的系统总是试图呈现宏观上的光学各向同性，即无序的系统（分子的取向随机分布的），如 CS_2 、液晶等。当一强电场加到这样的系统上，感应的分子偶极矩感受到一扭矩，此扭矩试图使分子克服热起伏而沿着最可极化轴向排列。当外加电场是直流电场时，这效应称为克尔效应。类似地，场可以来自于强光波，系统对快速变化的瞬态场无响应，但是对其时间平均场响应，此效应被称为光克尔效应。当取向光波场和探测光波场为同一光波时（即简并分子重新取向过程），这样可用 n_2 表示的自诱导非线性折射率变化，且 n_2 可以用 $\chi_{iiii}^{(3)}(-\omega; \omega, \omega, -\omega)$ 表示。强泵浦诱导的弱探测情况（即非简并分子重新取向过程），当泵浦和探测同偏振时，对于弱探测的非线性折射率系数为单光束自诱导情况的 2 倍 $[n_2^{(\text{weak})}]_{ii} = 2n_2$ ；相反，对于交叉偏振的泵浦和探测情况，则有 $[n_2^{(\text{weak})}]_{ij} = -[n_2^{(\text{weak})}]_{ii}/2$ 。

具有小的各向异性分子，典型的非线性为 $\chi^{(3)}$ 约为 $10^{-13} \sim 10^{-12} \text{ esu}$ ($10^{-22} \sim 10^{-21} \text{ m}^2/\text{V}^2$)， n_2 约为 $10^{-12} \sim 10^{-11} \text{ esu}$ ($10^{-21} \sim 10^{-20} \text{ m}^2/\text{V}^2$)，响应时间 10^{-12} s 。液晶具有大的非线性，但响应时间慢， n_2^I 约为 $5 \times 10^{-4} \text{ cm}^2/\text{W}$ （约比 CS_2 大 10 个数量级），响应时间 $10^{-3} \sim 10^{-2} \text{ s}$ 。

④ 电致伸缩。电致伸缩是与非均匀强度（即空间变化电场）的存在有关的物理现象。例如，这一现象会出现在如下两个情况：相干光波叠加形成的明暗相间条件的干涉图案，或者沿着高斯光束传播方向的横向。这样的非均匀场在分子或原子组成的系统中产生电致伸缩力。电致伸缩力正比于电场模的梯度，其比例常数 $\gamma_e/2$ 由电致伸缩系数 $\gamma_e = \rho \partial(\epsilon_0 n_0^2)/\partial \rho$ 给出，其中 ρ 为质量密度。介质中感应偶极子在非均匀的电场中受到平移力的作用（在均匀场中偶极子将不受这样力的作用），电致伸缩力使偶极子移到高强度区

域, 这导致了由于局域密度的增加产生局域折射率的增加。实际上, 即便在各同性分子系统中也可以是电致伸缩的。源于电致伸缩的非线性折射率系数为

$$n_2 = \frac{3}{4n_0} \chi_{xxxx}^{(3)} = \frac{\epsilon_0 (n_0^2 + 2)^2 (n_0^2 - 1)^2}{36 n_0 \rho v_a^2} \quad (13-73)$$

其中 v_a 为介质中声速。典型情况下, v_a 约为 10^4 m/s , ρ 约为 10^3 kg/m^3 , n_0 约为 1.4, 那么电致介质伸缩对非线性折射率系数的贡献为 n_2 约为 $10^{-19} \text{ m}^2/\text{V}^2$ 。这与分子重新取向的贡献大致相同量级, 或者略高些。但是响应时间比分子重新取向慢近三个量级, 约 10^{-9} s 。

⑤ 布居重新分布。当入射光频接近原子或分子系统的共振跃迁频率时那么将导致实际跃迁的发生。这意味着电子可以在有限的时间内占有(填充)激发态, 此即布居重新分布。由于光极化通常由处于基态(对于低强度光)原子和分子总数确定, 此布居重新分布导致了折射率的变化。这一效应可以在诸如原子蒸气, 分子气体, 有机分子溶液, 金属离子或色心掺杂的介电体和半导体中被观测到。近共振相互作用可以被描述为两个能级的情况。这是由于基态和激发态间的共振相互作用强于其他非共振作用。对于具有有限价带和导带的半导体略有不同, 依赖于相互作用的时间尺度。

两能级系统中的两个能级一个是基态(g)而另一个是激发态(e), 与两能级相关的共振跃迁频率定义为 ω_{eg} 。系统可以用跃迁偶极矩 μ_{eg} 来表征, 平衡布居差为 $\Delta N = N_{g0} - N_{e0}$, N_{g0} 和 N_{e0} 为平衡时处于基态和激发态的粒子数密度。特征时间 T_1 和 T_2 分别代表非平衡布居差衰减到平衡值的时间和跃迁偶极矩的特征迟相寿命时间。对近共振情况, $\omega = \omega_{eg} + \Delta$ (Δ 为频率失谐), 极化率将为复量, 其实部与折射率相关, 而虚部与吸收系数有关, 极化率正比于非平衡布居数。强度依赖的极化率为

$$\chi = \frac{c\alpha_{00}}{\omega_{eg}} \frac{i - T_2\Delta}{1 + (T_2\Delta)^2 + |E|^2/|E_{so}|^2} \quad (13-74)$$

其中 α_{00} 为线性吸收系数, E_{so} 为饱和场

$$\alpha_{00} = \omega_{eg} |\mu_{eg}|^2 T_2 \Delta N_0 / \epsilon_0 c \hbar, |E_{so}|^2 = \hbar^2 / 4 |\mu_{eg}|^2 T_1 T_2 \quad (13-75)$$

方程(13-75)仅适用于折射率近似为1的原子或分子气体, 对于如溶液或掺杂的固体等, 必须引入适当的局域场因子。当外加场较大, 但仍然小于饱和场时, 式(13-75)所示的极化率可以展开为量 $|E|^2/|E_{so}|^2$ 的二项式。前一项为线性极化率 $\chi^{(1)}$; 而后一项为含有 $|E|^2$ 项, 非线性极化率 $\chi^{(3)}$ 为相应的系数 $1/3$, 即有

$$\chi^{(1)} = (c\alpha_{00}/\omega_{eg})(i - T_2\Delta)/(1 + T_2^2\Delta^2), \chi^{(3)} = -\chi^{(1)}/3(1 + T_2^2\Delta^2)|E_{so}|^2 \quad (13-76)$$

那么, 非线性折射率系数则为 $n_2 = 3\text{Re}[\chi^{(3)}]/4n_0$ 。

对钠蒸气的 $3s \rightarrow 3p$ (λ 约为 589.3nm) 跃迁, $|\mu_{eg}| = 2.5ea_0 = 2.0 \times 10^{-29}\text{C}\cdot\text{m}$, 取 $N = 10^{20}\text{m}^{-3}$, $T_2 = 2T_1 = 32\text{ns}$, $a_{00} = 1.5 \times 10^7\text{m}^{-1}$ 。光频失谐 $\Delta = -6\pi \times 10^{10}\text{rad/s}$ 。得到 $\chi^{(1)} \sim 2.3 \times 10^{-4}$, 而线性折射率 n_0 约为 1, 忽略局域场因子, 最终得到 $\chi^{(3)}$ 约为 $-1.6 \times 10^{-16}\text{m}^2/\text{V}^2$, n_2 约为 $-1.2 \times 10^{-16}\text{m}^2/\text{V}^2$ 。

半导体是非常重要的非线性光学材料, 它的一些效应源于电子布居重新分布。半导体材料与透明电介质的主要不同是: 半导体在热平衡时可以有少的自由载流子布居。这些载流子, 可以吸收电磁辐射, 导致折射率直接依赖于等离子体频率。线性吸收能量为 $\hbar\omega$ 的光子 ($\hbar\omega > E_g$, E_g 为能隙), 激发电子至导带, 增加电子布居密度, 导致强度依赖折射率变化。那么, 非线性折射率强度系数为 $n_2^I = -e^2\tau_R\alpha_0/2m_e\epsilon_0n_0\hbar\omega^3$, m_e 为电子有效质量, τ_R 为电子-空穴再复合时间 (约 10ns), α_0 对应于能隙的线性吸收系数。 n_2^I 对 ω^{-3} 的依赖性使得这些效应在红外波段更强, 典型值 n_2^I 约 $10^{-6}\text{cm}^2/\text{W}$ 。上面的讨论仅限于体半导体材料。大的非线性效应已经在人工制备的半导体材料 (如多量子阱和量子点) 中被观测到, 量子阱材料是由不同种类的半导体材料交替变化而成的多层膜结构。在薄的维度, 量子限制效应导致比相应的体半导体更尖锐的激发共振, 而且更易于饱和, 最终呈现大的光学非线性, n_2^I 约为 $-2 \times 10^{-4}\text{cm}^2/\text{W}$ 量级的非线性折射率强度系数已被观测到。量子点是具有几个纳米大小的半导体颗粒, 测得的典型值 n_2^I 约为 $10^{-12} \sim 10^{-11}\text{cm}^2/\text{W}$ 。

⑥ 热贡献。稳态时源于热贡献的非线性折射率强度系数为 $n_2^I = (\alpha\tau/\rho c)dn/dT$, 因温度上升引起的折射率变化可以表示为 $\Delta n = \Delta T dn/dT$, dn/dT 为热光系数。在多数液体和固体中, 折射率的变化源于密度随温度的变化。膨胀导致密度的下降, 同时折射率又正比于密度, 因此对 dn/dT 的贡献通常是负的。但是一些半导体材料中, 吸收边随温度红移 (Franz-Keldish 效应), 这导致一个正的热光系数。 α 为线性吸收系数, 是材料的本征或杂质吸收。但是, 具有可观的吸收任何材料, 在稳态时均将呈现热的非线性折射率。时间常数 τ 表示稳态时的热扩散时间。量 $\alpha\tau I$ 为材料单位体积吸收的能量而且转换为热。吸收的能量通过比热容 c 与密度 ρ 的积, 转化为温度变化 ΔT 。 $|dn/dT|$ 的典型值在 $10^{-6} \sim 10^{-3}/\text{K}$ 的范围内。如果 α 约为 $10^2/\text{m}$, τ 约为 $10^{-4}/\text{s}$, ρ 约为 $10^3\text{kg}/\text{m}^3$, c 约为 $10^3\text{J}/(\text{kg}\cdot\text{K})$, dn/dT 约为 $-10^{-4}/\text{K}$, 那么非线性折射率强度系数 n_2^I 约为 $-10^{-12}\text{m}^2/\text{W}$ 。

⑦ 光折变效应 (photorefractive effect)。光折变效应也是强度诱导的折

射率变化现象，但与前面描述的非线性折射率效应完全不同；折射率的变化与绝对光强度无关，仅与光强度的空间变化（即光强度梯度）有关。概括的讲，光折变效应是由于光强度的空间分布引起折射率变化的现象，所以光折变效应是一种非局域响应。在材料中，最大的折射率变化既非瞬态地出现也不需要局域的光强度最大。非局域响应是因为载流子的移动和在光折变材料中的积累，使得它可以在时间上积累存储光激励，时间上的积累显著地降低了引起显著折射率变化所需要的瞬态光强度。

超低强度特性是光折变效应的主要优点之一。多数的光学材料要求大的光强度（一般 $> \text{kW}/\text{cm}^2$ ），以产生显著的折射率变化。光折变材料则相反，所需强度仅 mW/cm^2 甚至 $\mu\text{W}/\text{cm}^2$ 量级。因其累积特性使得其响应速度较慢，但是不同的光折变材料和工作条件提供了较大范围的响应速度，可以从纳秒或皮秒到小时或更长。光折变效应的另一个特点是具有记忆（或存储）功能。快速响应的材料适合于实时光信息处理，而慢响应的材料对光存储有益。

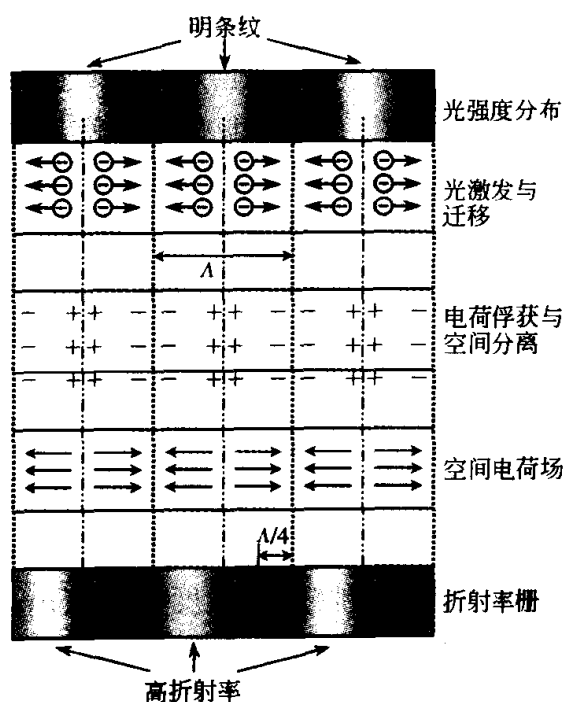


图 13-18 扩散主导迁移的光折变过程
(折射率栅相对于干涉条纹存在四分之一
周期的空间移动，即有 $\pi/2$ 的相移)

联合使其移动（迁移）。当无外加电场时，载流子从强度最大处对称地向它处扩散。当有外加电场时，漂移将主宰迁移过程。

光生载流子的产生和迁移之后的光折变过程是载流子的俘获，在材料中

光折变效应是一个典型的线性与二阶非线性效应级联所致的非线性折射率变化现象。其物理过程或机制可以概括为如下五个过程：a. 杂质的光离子化和电荷载流子的生成；b. 电荷载流子的迁移；c. 电荷载流子的俘获和空间电荷分离；d. 光诱导空间电荷场的建立；e. 通过普克尔（线性电光）效应而引起折射率变化。在图 13-18 中示出了无外加电场时扩散主导迁移的光折变过程。详细地，首先空间变化的光场（如明暗相间的干涉条纹）照射在晶体上，激发产生光生载流子，生成速率在强度最大处（明条纹处）局域最大。非平衡光生载流子通过扩散或漂移或两者的

形成的半永久性的空间电荷密度。捕获发生在缺陷位置，因为那里是空的，可以有效地捕获载流子。被分离和俘获的空间电荷形成空间电荷场。在有外加电场下，漂移主导迁移过程时，暗条纹处电场最大，明条纹处则被屏蔽。另一方面，如果扩散主导迁移过程，电荷场相对于明条纹移动四分之一条纹间距；在材料内空间电荷场相对于干涉条纹的空间位移，对光折变非线性光混频起了最基本的作用。

光折变折射率栅形成过程的最终一步就是空间电荷场通过线性电光（普克尔）效应改变材料的折射率。折射率相对光激励的四分之一条纹间距位移（对于扩散主导的迁移），在光波混频过程中，起了特殊作用。它允许相互干涉的两个光束，一个被放大，另一个被倒空；这被称为不可逆能量转移，或光折变增益，是许多光折变现象和应用的基础。

光折变效应首先在 LiNbO_3 晶体中被意外发现，现已在许多材料中，如： BaTiO_3 ， $\text{Bi}_{12}\text{SiO}_{12}$ (BSO)， KTaNbO_3 (KTN)， $\text{Sr}_{1-x}\text{Ba}_x\text{Nb}_2\text{O}_6$ (SBN)， $\text{K}_{1-x}\text{Na}_x\text{Sr}_{1-x}\text{Ba}_x\text{Nb}_5\text{O}_{15}$ (KNSBN)， KNbO_3 ， $\text{Bi}_{12}\text{GeO}_{20}$ (BGO)，GaAs，InP，CdTe，CdS 等观测到光折变效应。有实用价值的光折变材料必须具有适当的材料特性。例如，光折变材料必须是绝缘体或半绝缘的半导体，否则剩余自由载流子将屏蔽被俘获的空间电荷。此外，光折变材料也必须有适当的光电导，允许电荷分离和形成空间电荷场。对于光折变过程，足够浓度的缺陷态也是必需的，因为它们为俘获空间电荷提供了空位。电光效应最终决定光诱导折射率栅的大小。铁电氧化物，如： BaTiO_3 ， KNbO_3 ， LiNbO_3 和 SrBaNbO_3 等，具有大的线性电光系数，使得它们非常重要。此外，大的电光效应也会出现在体半导体或量子阱结构的共振或近共振光激发情况。

13.4.2 非线性吸收

所谓的非线性吸收指的是材料的透过率随光强变化的效应或现象。对足够高的光强，材料同时吸收一个以上光子的几率可以被大大增强。不仅双光子吸收，多光子吸收在许多材料中也被观测到。强激光诱导的布居重新分布导致许多有趣的现象，如受激发射和吸收，在复杂分子系统中的复杂能量跃迁，以及在固体中的自由载流子产生。这些现象在光学上表现为降低（饱和）吸收和增加（反饱和）吸收。

(1) 双光子吸收 双光子吸收 (two-photon absorption, TPA) 涉及从系统的基态到激发态的跃迁，但该跃迁必须同时吸收来自入射光波的两个光子。双光子吸收跃迁与单光子吸收跃迁遵从不同的选择定则，所以双光子吸收谱补充了线性吸收谱在研究系统激发态上的不足。图 13-19 示出了两种可能情况的双光子吸收概念图，①两个被吸收光子来自同一频率 ν （圆频率 ω ）的光波场；②两个被吸收光子来自两个不同频率 ν_e 和 ν_p （圆频率 ω_e 和

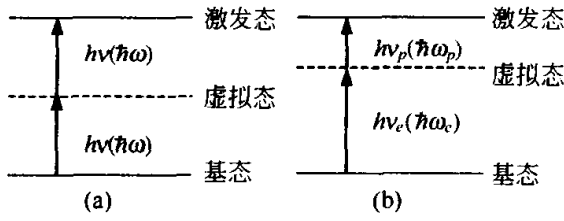


图 13-19 双光子吸收概念图

(a) 自 (简并) 双光子吸收;

(b) 泵浦-探测 (非简并) 双光子吸收

ω_p) 的光波场, 下标 e 表示泵浦或激发光束, 下标 p 表示探测光束。图中的虚线表示的中间态在系统中不是实际存在的。所以, 系统必须同时吸收两个光子。涉及单光子泵浦的中间态二步 (two-step) 吸收, 将在后面作为激发态吸收予以讨论。

(2) 简并双光子吸收 此类非线性吸收正比于瞬态强度的平方, 光损耗的微分方程为

$$dI/dz = -\alpha I - \beta I^2 \quad (13-77)$$

其中 α 为杂质引起的线性吸收, β 为双光子吸收系数。双光子吸收系数是表征材料的一个宏观物理量。通常, 感兴趣的是单个分子的双光子吸收特性, 即由双光子吸收界面 σ_2 来描述, 而且宏观量 β 和微观量 σ_2 间关系为 $\sigma_2 = \beta \hbar \omega / N$, N 为系统的分子数密度, $\hbar \omega$ 为入射光场的光子能量。双光子吸收系数 β 与三阶非线性极化率 $\chi^{(3)}$ 的虚部相关联

$$\beta = (3\pi/\epsilon_0 n^2 c \lambda) \text{Im}[\chi_{xxxx}^{(3)}(-\omega; \omega, \omega, -\omega)] \quad (13-78)$$

通常, 入射光为线偏振的, 而且介质为中心对称的, 因此 β 仅与 $\chi^{(3)}$ 的 $xxxx$ 分量有关。

(3) 非简并双光子吸收 此过程, 如图 13-19 (b), 涉及两个不同频率光子的同时吸收。它们分别被称为泵浦或激发频率 ω_e 与探测频率 ω_p 。两个光束的强度原则上没有限制, 但通常是 $I_e \gg I_p$ 。此过程的双光子吸收系数当然仍与 $\chi^{(3)}$ 有关, 并考虑到对称性

$$\beta_{ep} = \beta_{pe} = (3\pi/\epsilon_0 n_e n_p c \sqrt{\lambda_e \lambda_p}) \text{Im}(\chi_{ep}^{(3)}) \quad (13-79)$$

$$\chi_{ep}^{(3)} = \chi_{ijij}^{(3)}(-\omega_e, \omega_e, \omega_p, -\omega_p) = \chi_{jiii}^{(3)}(-\omega_p; \omega_p, \omega_e, -\omega_e) \quad (13-80)$$

其中 i, j 分别为泵浦和探测光束的偏振方向。实际上, 泵浦和探测光束的偏振可以是正交的, 意味着此非线性过程可以导致材料的感应二向色性。描述非简并双光子吸收过程中的光强衰减的微分方程为

$$\begin{aligned} dI_e/dz &= -\alpha_e I_e - 2(\omega_e/\omega_p)^{1/2} \beta_{ep} I_e I_p - \beta_{ee} I_e^2 \\ dI_p/dz &= -\alpha_p I_p - 2(\omega_p/\omega_e)^{1/2} \beta_{pe} I_p I_e - \beta_{pp} I_p^2 \end{aligned} \quad (13-81)$$

方程组中的第一项和最后一项分别描述线性吸收和自诱导双光子吸收, 而中间项代表诱导二向色性。现在考虑两种情况。首先, 泵浦光子能量小于允许的双光子跃迁能量的一半, 这样含有 β_{ee} 项可以被忽略, 因为不满足双光子

吸收的共振跃迁条件；且含有 β_{pp} 的自诱导吸收项也可以被忽略，因为探测相对于泵浦已被假定为弱的。其次，泵浦光束的自诱导双光子吸收被允许，且由探测诱导的二向色性可以被忽略；这意味着泵浦光束可以激发由自身（在 $2\omega_e$ ）以及与探测的联合（在 $\omega_e + \omega_p$ ）的允许双光子态；当材料具有大的双光子允许态密度时（如在半导体的导带和多原子分子的激发态），此种情况是可能出现的。

(4) 三光子吸收 共振三光子吸收现象如图 13-20。包括三种情形，单一光波场，两个光波场和三个光波场。在此仅讨论单一光波场情况，假定杂质引起一些背景线性吸收，描述单光束三光子吸收衰减的微分方程为

$$dI/dz = -\alpha I - \gamma I^3 \quad (13-82)$$

其中 γ 为三光子吸收系数。三光子吸收是一个五阶非线性过程，假定光束是线偏振的且介质为中心对称的。 γ 与五阶非线性极化率的虚部关系为

$$\gamma = (5\pi/\epsilon_0^2 n^3 c^2 \lambda) \text{Im}[\chi_{xxxxx}^{(5)}(-\omega; \omega, \omega, \omega, -\omega, -\omega)] \quad (13-83)$$

(5) 多光子吸收 双光子和三光子的结果可以扩展到任意数量的光子吸收。多光子吸收表现为从单光束或多光束同时吸收多个光子。同样假定材料存在杂质引起的线性吸收，单光束 n 光子吸收衰减方程为

$$dI/dz = -[\alpha + \gamma^{(n)} I^{(n-1)}] I \quad (13-84)$$

$\gamma^{(n)}$ 为 n 光子吸收系数。非共振多光子吸收通常非常弱，但在半导体中可以是比较显著的。

(6) 激发态吸收 当入射光的强度很强时，那么激发态布居变得显著。在诸如多原子分子和半导体系统中，存在高密度的激发态，被激发电子在它跃迁返回基态之前，可以快速跃迁到这些高密度态中的一个。所以，电子完全弛豫回基态之前，它可以再次经历吸收，使其进入更高的能态，此过程被称为激发态吸收。当入射光的强度足以显著地倒空基态时，可以观测到明显的激发态吸收。

当激发态的吸收截面小于基态时，如果系统是高度被激发的，系统的透过率将增加，此过程称为饱和吸收。它类似于简单的二能级系统的饱和吸收，但它更复杂。当激发态的吸收截面大于基态时，如果被激发，那么系统的透射将减小，这给出与饱和吸收相反的结果，则被称为反饱和吸收。

在半导体中，一个能量比带隙大的光子的吸收将电子激发至导带，在那

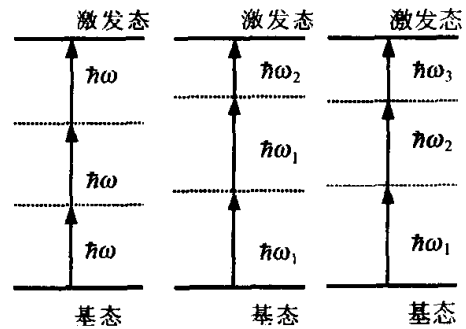


图 13-20 三光子吸收概念图

(a) 简并三光子吸收；(b) 两光束三光子吸收；

(c) 三光束三光子吸收

里它是一个自由载流子而且可以对电流有贡献，被激发的电子将快速的热弛豫至导带底。从那里经一特征复合时间之后，它将与处在价带上的激发空穴再复合。但是，在足够高强度，它可以吸收其他光子，而仍然处在导带中。这一过程被称为自由载流子吸收，定性地与反饱和吸收类似。

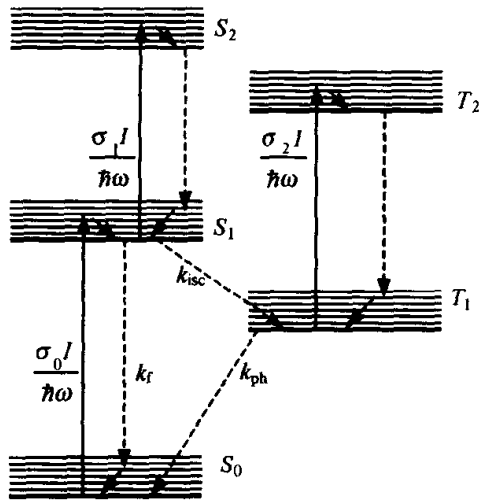


图 13-21 五能级系统图

(7) 饱和吸收 多原子分子的简化能级图在图 13-21 示出。通常称为分子的五能级模型，涉及五个固有电子态，成功地解释了在较宽入射强度范围内的非线性吸收。在每个电子态，存在许多非常致密的振-转能级。当电子从一个电子态被激发到另一个电子态时，它将跃迁到这些振-转能态之一。但是，碰撞将快速地使电子加热，它将跃迁至大量电子态中最低的振-转能级。从该态，它或者吸收其他光子或弛豫到致密低能态中的任何一个。基电子态称为单重态，这些态有一反平行自旋的电子对。将这对

反平行自旋中的一个反转为平行自旋的电子对，选择定则不允许任何的辐射跃迁。自旋反转的实现可以借助于外部过程（如与顺磁离子的碰撞）或内部过程（如强旋-轨耦合）。在这样条件下，首次被激发的电子态可以通过无辐射跃迁至较低的三重态（即具有平行自旋电子对的态），根据选择定则，从这一态允许的辐射跃迁，只有至另一三重态。所以，在这样的分子系统，只涉及单重态-单重态和三重态-三重态的辐射跃迁。

在五能级系统中的吸收过程如下：吸收一个入射光子而激发一个电子至第一激发单重态。在该态可以发生三种可能中的一个过程：电子可以通过辐射或非辐射跃迁弛豫回基态，该跃迁的净速率常数由参量 k_f 给出；另一个可能是电子遭受自旋反转跃迁至三重态，此过程称为系统内交叉，且有一速率常数 k_{isc} 。第三个可能是分子可以吸收其他光子，激发电子至更高单重态，而后弛豫回第一激发单重态。对处于最低三重态的电子，有两种可能：它可以通过其他的自旋反转跃迁至基态，伴随有磷光过程的发生（具有速率常数 k_{ph} ）；另一种可能是分子吸收其他光子，激发电子至更高的三重态，而后电子弛豫回最低三重态。从更高的单重和三重态的弛豫速率通常非常大。一般地假定这些速率如此之大，以至于这些态上的布居密度 S_2 和 T_2 可以忽略不计。激发态的受激发射也可以忽略。对于三个最重要的态， S_0 、 S_1 和 T_1 的布居密度速率方程容易写出。引入无量纲量 $S_0 = N_{S_0}/N$ ， $S_1 =$

N_{S1}/N 和 $T_1 = N_{T1}/N$, N_{S0} , N_{S1} 和 N_{T1} 分别为相应态的布居密度。那么

$$\partial S_0/\partial t = -(\sigma_0 I/\hbar \omega) S_0 + k_f S_1 + k_{ph} T_1 \quad (13-85)$$

$$\partial S_1/\partial t = (\sigma_0 I/\hbar \omega) S_0 - k_f S_1 + k_{isc} T_1 \quad (13-86)$$

$$\partial T_1/\partial t = k_{isc} S_1 - k_{ph} T_1 \quad (13-87)$$

入射光的衰减可以描述为

$$\partial I/\partial z = -(\sigma_0 S_0 + \sigma_1 S_1 + \sigma_2 T_1) NI \quad (13-88)$$

在此, $I(z, t)$ 为光强度, σ_0 , σ_1 和 σ_2 分别为基态, 第一激发单重态和第一激发三重态的吸收截面。由于密度守恒, $S_0 + S_1 = T_1$ 。

在稳态情况下, 式 (13-85) ~ 式 (13-88) 等于零。从式 (13-88) 可知 $T_1 = g S_1$ ($g = k_{isc}/k_{ph}$)。一般地, 系统内交叉速率比磷光速率大几个量级 (如, k_{isc} 约为 10^9s^{-1} , k_{ph} 约为 10^6s^{-1}), g 是很大数, 表明三重态成为电子的陷阱 ($T_1 \gg S_1$)。所以, 在稳态的非线性吸收主要由三重态截面确定。通常, 不希望 σ_1 和 σ_2 的大小相差太大。但是, 三重态增强因子 g 使得 σ_2 成为最重要的参量。如果 $\sigma_2 < 0$, 系统在稳态时, 显现饱和吸收。饱和强度 $I_s = \hbar \omega (k_f + k_{isc})/\sigma_0$ 由第一激发单重态的特性 (即 $S_1 \rightarrow S_0$ 截面) 和 S_1 态的总弛豫时间 $\tau_1 = (k_f + k_{isc})^{-1}$ 确定。

在脉冲激发情况下, 若入射光脉宽 Δt 为第一激发单重态弛豫时间量级, 那么非线性吸收将更显著地依赖于 $S_0 \rightarrow S_1$ 吸收截面。但是, 若脉宽 $\Delta t \ll \tau_1$, 单重态间跃迁将主导非线性吸收; 当脉宽 $\Delta t \gg \tau_1$, 三重态间吸收占主导。当入射光的强度达到某一强度值时, 即使再增加强度, 透过率不再增加达到较高的饱和值。定义参量 $\delta_1 = (\sigma_1 - \sigma_0)/\sigma_0$ 和 $\delta_2 = (\sigma_2 - \sigma_0)/\sigma_0$ 为相对吸收截面差。特别地当 $\delta_2 = -1$ 时, 饱和透过率可接近 1, 即没有损耗。

(8) 反饱和吸收 在呈现反饱和吸收的材料中, 激发态的吸收截面大于基态的吸收截面。这样所有描述饱和吸收的表达式 (13-85) ~ 式 (13-88) 对反饱和吸收都适用, 只需注意反饱和吸收时 $\delta_1, \delta_2 > 0$ (饱和吸收 $\delta_2 < 0$)。C₆₀ 就是典型的反饱和吸收体。

需要注意的是如果 δ_1 和 δ_2 的符号相反, 饱和吸收或反饱和吸收都可出现, 取决于那个能级占主导地位。此外, 当激光的强度足够高时, 更高的单重—单重, 三重态—三重态的激发态吸收也会出现。

参 考 文 献

- 1 Shen Y. R. The principles of Nonlinear Optics. New York: John Wiley & Sons Inc., 1984
- 2 雅里夫 A. 著. 李宗琦译. 光电子学导论. 北京: 科学出版社, 1983

- 3 Hummel R E. Electronic properties of materials. 2nd ed. Tokyo: Spriger-Verlag, 1992
- 4 Nolte D. D. Photorefractive effects and materials. London: Kluwer Academic Publishers, 1995
- 5 Born M. Wolf E. Principles of optics. 6th ed. Cambridge: Cambridge University Press, 1997
- 6 Pavels L et al. Nature. 2000, 408: 440
- 7 Zhu Y. Y. et al. Optical and Quantum Electronics. 1999, 31: 1093
- 8 Joannopoulos J. D. et al. Nature. 1997, 386: 143
- 9 Ponce F. A. Bour D. P. Nature. 1997, 386: 351
- 10 Richard L. Sutherland. Handbook of Nonlinear Optics. New York: Manel Dekker Inc. 1996

第 14 章 力学性质

撰稿人 冯 端 陈延峰 杨 彬

材料的力学性质是指材料在外应力作用下的行为。最常用的力学性质包括：弹性模量 E ，描述应力和应变之间的比例关系；屈服强度 σ ，指材料发

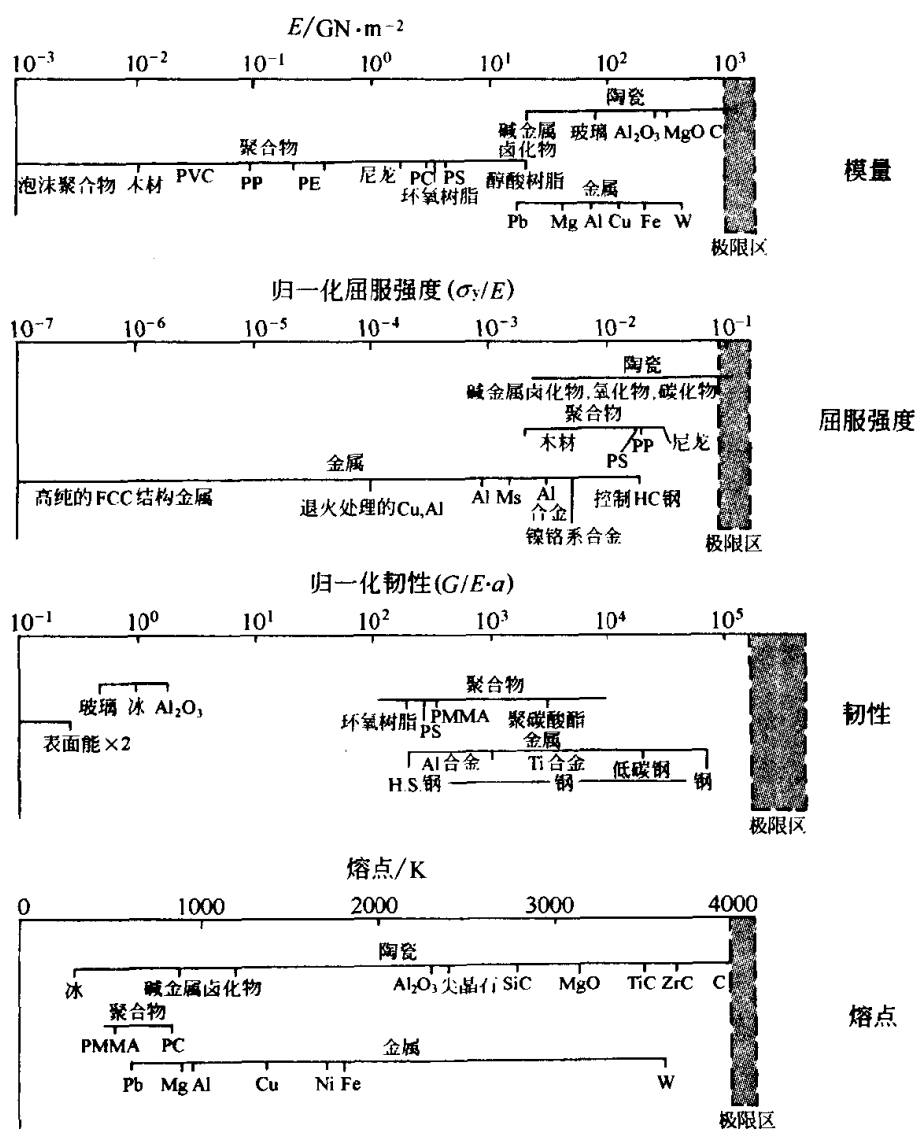


图 14-1 工程材料的金属、陶瓷和聚合物重要力学特性

其中对于归一化的韧性， G 的单位为 J/m^2 ，

E 为相应单位的杨氏模量， a 为原子半径

(引自 M. F. Ashby. 发展中的材料研究)

生塑性形变的最小应力；硬度 G ，描述材料软硬的程度等。

人类最早学会利用的材料性质便是力学性质，如石器时代利用天然岩石的强度和硬度；青铜器时代利用铜的高塑性和高强度及加工硬化性能；而在铁器时代更是利用 Fe-C 合金的高强度、硬度和塑性。尽管如此，人类真正认识和开始系统地理解材料的力学性质起始于 19 世纪中叶，人们利用金相显微镜对材料细微组织进行了研究。在我们身边注意到的经验是金属具有延展性，陶瓷硬而脆，橡胶具有很大的弹性形变，如此等等。图 14-1 比较了金属、陶瓷和高分子力学性质。迥然不同的力学行为是由其基本结构决定的。金属与陶瓷材料的晶体结构（包括键合类型）、缺陷（主要是位错）是理解和描述其力学性质的核心概念；而在高分子材料中，却是分子链的构形，交联与缠结起了关键的作用。本章我们将研究这些概念如何使强度、塑性、断裂和高温行为得以定性和定量的描述和理解。

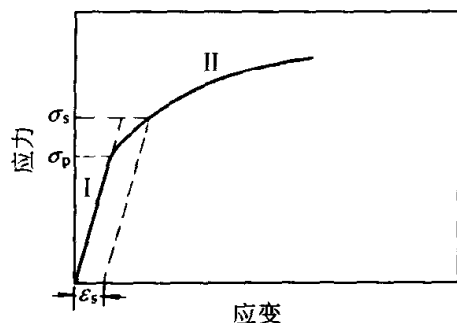


图 14-2 典型的应力-应变曲线

σ_p —屈服强度； σ_s —屈服应力；

ϵ_s —卸载后剩余范性应变值

图 14-2 是一个典型的应力-应变曲线，是由应变速率恒定的拉伸机测量所获得的。曲线共分为两个区域。I 是弹性区，其基本特征是应力应变间满足虎克定律，其比例系数定义为弹性模量，同时，当外力释放后，材料的变形能够恢复原来的状态；II 是塑性变形区，其特征是应力应变间不一定满足正比关系，其特征系数远小于 E ，外力释放之后再回复不到初始材料的长度。划分 I-II 区域的应力定义为材料的屈服强度，标志着塑性变形的开始，

是工程应用中最为关心的材料参数之一。II 区的终点是断裂强度，是材料失效的最小值。本章的主要内容就是对这一现象的规律进行描述和解释。

14.1 弹性与广义弹性

14.1.1 弹性参量

(1) 应力 为了描述在外力作用下，处于平衡状态的任意形状的介质内任一点处的应力状态，我们取通过该点的一个无限小的面元 dS （其法线矢量为 \mathbf{n} ），设通过该面元的作用力为 $d\mathbf{F}$ 。则对应于该面元（或方向 \mathbf{n} ）的应力矢量 \mathbf{T} 定义为：

$$d\mathbf{F} = \mathbf{T}dS \quad (14-1)$$

显然应力矢量 \mathbf{T} 和 \mathbf{n} 的方向不一定相同，也就是说通过该点的应力与所考虑的平面的方向有关。想要全面的描述介质中的应力状态，我们就应该知道

过介质中每一点的任意截面上的应力。为了达到这个目的，一般是在该点附近取一个无限小的体积元（图 14-3），只要求出该立方体积元六个面上的应力，就可以知道过该点任意截面上的应力。

如图 14-3 所示，我们引入笛卡儿坐标 x, y, z ， O 为坐标原点。作用于 O 点的应力，在体积元的六个面上都作用有应力。显然每个面上相应的应力总可以分为平行于面法向和垂直于面法向的两类分量：平行于面法线的应力分量叫做正应力，垂直于面法线的应力分量叫切应力。对应于坐标轴向的应力矢量 T 的分量分别用 T_x, T_y, T_z 来表示，写为分量的形式即：

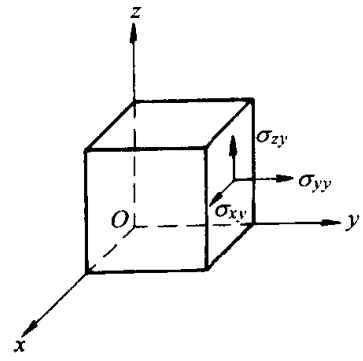


图 14-3 应力张量
诸分量的说明

$$\begin{aligned} T_x &= \sigma_{xx}i + \sigma_{yx}j + \sigma_{zx}k \\ T_y &= \sigma_{xy}i + \sigma_{yy}j + \sigma_{zy}k \\ T_z &= \sigma_{xz}i + \sigma_{yz}j + \sigma_{zz}k \end{aligned} \quad (14-2)$$

式中， σ_{xy} 表示 T_y 的 x 分量，余者类同。 σ_{ij} 即构成了应力张量 σ 。在 σ 的诸分量中， $i=j$ 的是正应力分量，而 $i \neq j$ 的是切应力分量。

当应力张量 σ 的各个分量 σ_{ij} 均已知时，对应于任意方向 n 的应力矢量 T 即可求出。用张量 σ 与矢量的并矢积形式可表示为：

$$T = \sigma \cdot n \quad (14-3)$$

由于无限小体积元是处于平衡状态的，因此根据介质中体元所受的力偶矩平衡条件可导出：

$$\sigma_{ij} = \sigma_{ji} \quad (14-4)$$

这表明应力张量是对称张量，只有六个独立分量，即三个正应力和三个切应力。知道了这六个应力分量就知道了该体积元的应力状态。

(2) 应变 在应力作用下，固体的形状和体积会发生变化。应变就是用来描述固体在应力作用下内部各点相互位置的改变的参量。介质中任意一点形变前后的位置可以分别用矢径矢量 r 和 r' 来表示，则代表变化的位移矢量，也就是位置的函数可定义为：

$$u = r - r' \quad (14-5)$$

如果知道了介质内任意一点的位移矢量，就能够完全地描述其形变状态。在无限小形变的情形下，若忽略刚体式位移，则相邻两点之间的相对位移 du 可表示为：

$$\begin{aligned}
d\mathbf{u} = & \left(\frac{\partial u_x}{\partial x} dx + \frac{\partial u_x}{\partial y} dy + \frac{\partial u_x}{\partial z} dz \right) \mathbf{i} \\
& + \left(\frac{\partial u_y}{\partial x} dx + \frac{\partial u_y}{\partial y} dy + \frac{\partial u_y}{\partial z} dz \right) \mathbf{j} \\
& + \left(\frac{\partial u_z}{\partial x} dx + \frac{\partial u_z}{\partial y} dy + \frac{\partial u_z}{\partial z} dz \right) \mathbf{k}
\end{aligned} \quad (14-6)$$

这里的系数是形变张量 β 的分量。形变张量是非对称的，因此我们可以将它写为对称张量和非对称张量之和

$$\beta_{ij} = e_{ij} + w_{ij} \quad (14-7)$$

其中

$$\begin{aligned}
e_{ij} &= \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) = e_{ji} \\
w_{ij} &= \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i} \right) = -w_{ji}
\end{aligned} \quad (14-8)$$

相对位移 $\sum w_{ij} dx_j$ 与刚体的旋转相当，使介质内相邻两点间的距离和夹角关系保持不变；而相对位移 $\sum e_{ij} dx_j$ 则使体元的形状与大小均发生变化。因而对称张量 e 称为应变张量，其中 $i=j$ 的分量称为正应变分量， $i \neq j$ 的分量称为切应变分量。

(3) 虎克定律 应力和应变是伴生的，他们之间的关系很复杂。只有理想的弹性体应力和应变之间才有最简单的线性关系。但对一般物体而言，在弹性范围内，作为一级近似，特别是在小形变时，应力与应变也满足广义虎克定律：

$$\sigma_{ij} = \sum_{k,l} c_{ijkl} e_{kl} \quad (14-9)$$

c_{ijkl} 构成一个四阶张量，称为介质的弹性模量张量，又称弹性刚度张量，可见弹性模量是材料发生单位应变时的应力，它表征材料抵抗形变能力（即刚度）的大小。 c 越大，越不容易变形，表示材料的刚度越大。弹性模量的单位与应力的单位相同 (N/m^2)

由于应力和应变都是对称张量，所以 $\sigma_{ij} = \sigma_{ji}$ ， $e_{kl} = e_{lk}$ ，故有：

$$c_{ijkl} = c_{jikl} = c_{ijlk} = c_{jilk} \quad (14-10)$$

这意味着弹性模量张量的 81 个分量中只有 21 个独立的分量。而根据晶体不同的对称性，独立分量的数目还会进一步降低。例如三斜晶系有 18 个独立的分量，单斜晶系有 12 个，正交晶系有 9 个，四方和菱面体晶系有 6 个，六角晶系有 5 个，立方晶系有 3 个，而在各项同性的介质中，弹性模量只有两个独立的分量。

下面只讨论各向同性介质中的情况，它有三种不同应变类型的弹性模量，分别称为杨氏模量 E 、切变模量 μ 和体积模量 B ，即

$$E = \frac{\sigma}{\epsilon} = \frac{F}{A_0} \frac{\Delta l}{l_0} \quad (14-11)$$

$$\mu = \frac{\tau}{\gamma} = \frac{F}{A_0 \tan \theta} \quad (14-12)$$

$$B = \frac{P}{\frac{\Delta V}{V_0}} = \frac{PV_0}{\Delta V} \quad (14-13)$$

对于各向同性材料，它们之间存在如下关系：

$$E = 2\mu(1 + \nu) = 3B(1 - 2\nu) \quad (14-14)$$

式中， ν 是泊松比 (Poisson ratio)，定义为在拉伸试验中，材料横向单位面积的减少与纵向单位长度的增加之比值，即

$$\nu = \frac{-\Delta A / A_0}{\Delta l / l_0} = \frac{-\epsilon_l}{\epsilon} \quad (14-15)$$

式中， A_0 为拉伸前材料的横截面积， ΔA 为 A_0 的减少量， ϵ_l 为横向应变。在 E ， μ ， B 和 ν 四个参数中只有两个是独立的，只要知道其中的两个参数就足以描述各向同性材料的弹性力学行为。

弹性模量是固体原子之间结合强度的标志之一，原子半径和离子半径越小，原子价越高的物质，其弹性模量和硬度就越大。通常，弹性模量大小的顺序是碳化物 (400 ~ 700 GPa) > 硼化物、氮化物 > 氧化物 (150 ~ 300 GPa)。表 14-1 列出了一些典型陶瓷材料的弹性模量。烧结形成的陶瓷材料内部存在气孔，导致了其弹性模量随气孔率的增大而降低。

表 14-1 几种典型陶瓷材料的弹性模量

材 料	$E / \text{N} \cdot \text{m}^{-2}$	材 料	$E / \text{N} \cdot \text{m}^{-2}$
刚玉晶体	38×10^{10}	烧结 MgAl_2O_4 (气孔率 5%)	23.8×10^{10}
烧结氧化铝 (气孔率 5%)	36.6×10^{10}	致密 SiC (气孔率 5%)	46.7×10^{10}
高铝瓷 (90% ~ 95% Al_2O_3)	36.6×10^{10}	烧结 TiC (气孔率 5%)	31×10^{10}
烧结氧化铍 (气孔率 5%)	31×10^{10}	烧结稳定化 ZrO_2 (气孔率 5%)	15×10^{10}
热压 BN (气孔率 5%)	8.3×10^{10}	SiO_2 玻璃	7.2×10^{10}
热压 B_4C (气孔率 5%)	29×10^{10}	莫来石瓷	6.9×10^{10}
石墨 (气孔率 20%)	0.9×10^{10}	滑石瓷	6.9×10^{10}
烧结 MgO (气孔率 5%)	21×10^{10}	镁砖	17×10^{10}
烧结 MoSiO_2 (气孔率 5%)	40.7×10^{10}		

14.1.2 常规弹性的物理本质

我们知道一个系统的平衡状态由其自由能 $F = U - TS$ 的极小值所决定。如果加上外力使一个固体拉伸，就会偏离其平衡态，使之自由能增大。通常固体作弹性拉伸时，其原子间距增大，因而外力对抗了原子间作用力作了功，导致固体的内能 U 增加，从而使自由能增大。因此常规弹性来源于

内能增加引起的自由能增加。

在简单的情况下两个固体原子之间相互作用的 Lennard-Jones 势可表示为：

$$\Phi(r) = \frac{pq}{p-q} \epsilon_b \left[\frac{1}{p} \left(\frac{a_0}{r} \right)^p - \frac{1}{q} \left(\frac{a_0}{r} \right)^q \right] \quad (14-16)$$

式中, ϵ_b 是势能极小值, p 和 q 反映势能变化的趋势。对于惰性元素 (如 Ar) 固体和金属, $p=12$, $q=6$, 上式简化为:

$$\Phi(r) = \epsilon_b \left[\left(\frac{a_0}{r} \right)^{12} - 2 \left(\frac{a_0}{r} \right)^6 \right] \quad (14-17)$$

势能曲线见图 14-4。原子间距离为 a_0 时, 势能最低。当固体受压变应力时, 固体势能增加, 原子之间受排斥力; 当受拉应力时, 原子之间受吸引力。力的大小为:

$$F_{\text{外力}} = + \frac{\partial \Phi}{\partial r} \quad (14-18)$$

根据式(14-17) 和式(14-18)有:

$$F_{\text{外力}} = \frac{60\epsilon_b}{91a_0} \cdot \frac{\Delta r}{a_0} = \frac{60\epsilon_b}{91a_0} \cdot \epsilon \quad (14-19)$$

结合式 (14-11), 有:

$$E = \frac{60\epsilon_b}{91a_0} \quad (14-20)$$

因而, 势能最小值越低, 则势阱越深, 改变原子之间的相对距离所作的功越大, 弹性模量越大。

实验上观察到, 多数多晶金属材料, 不管它们是否遵循 Lennard-Jones 势, 其弹性限度仅为 0.2%, 超过这个范围便发生塑性变形。其中的原因是因为金属中总有大量的位错存在, 由于金属键使得位错滑移很容易发生, 从而大大降低了其理论强度。

陶瓷材料的特征是硬而脆, 也就是其弹性模量很高 (通常是金属的 10 倍), 但其变形量

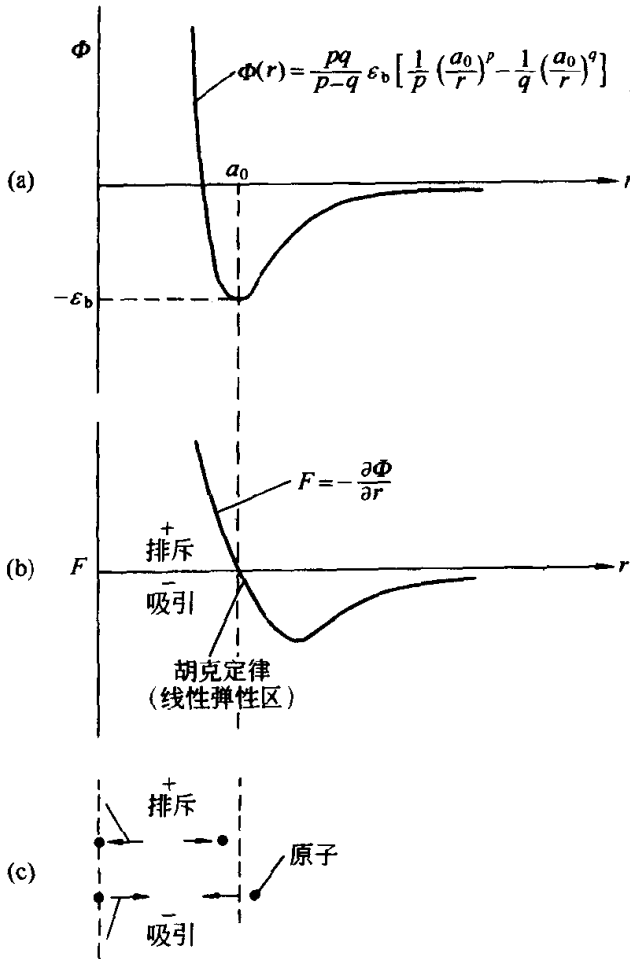


图 14-4 两固体原子之间相互作用的 Lennard-Jones 势
(a) 原子间的相互作用势函数 Φ 与原子间距 r 的关系, 相互吸引的相互作用势是短距的; (b) 原子间的相互作用力与 r 的关系; (c) 惯用的相互作用力方向和符号

很小，以至于很难利用拉伸实验获得弹性模量的数据。这是因为陶瓷通常的键合是离子键或共价键，原子之间的相互作用力很强，相互之间键角十分固定，以至于很难变形，另一方面，材料内部的微观缺陷（如位错，空位，晶界和微裂纹）也显著地降低了理论强度，而且由于键合的特点，使得陶瓷的应力释放以裂纹的扩展为主，而不像金属那样依靠位错的滑移而进行。

14.1.3 高弹性的物理本质

高弹性指物体可以伸长很多倍的性质，人们非常熟悉的橡皮筋就属于这类特别的弹性体。它们具有两个特点，一是宏观变形量特别大，很容易发生大的弹性变形，形变量甚至可以达到百分之几百，相对于金属和陶瓷而言，它可谓是巨弹性体；二是弹性模量很小。如果基于上面讨论的弹性模型，这一现象很难得到解释，不可能想像两个原子之间拉伸了几个原子间距还能不发生断裂分离，并且能够在应力释放之后回复到初始的位置。

一般的固体在张力作用下会产生弹性的伸长，当外力撤去，又恢复原状。通常伸长到 1% 左右就到了弹性极限。而一块高弹性材料则可以弹性地拉伸到原来长度的 10 倍，两者差异是如此之大，就应归源于机制不同。

根据自由能的表达式我们知道，系统自由能是由系统的内能和熵两部分组成的，因此增加内能或者减少熵都可以使系统的自由能增大。系统内能的增加带来自由能的增加导致了常规弹性的产生，而系统熵的减小所引起的自由能的增加则是高弹性产生的根本原因（如图 14-5）。

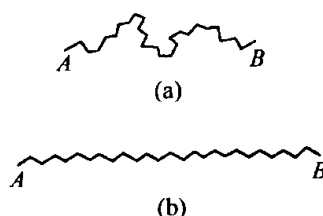


图 14-5 在交联点 A, B 之间的分子值
(a) 松塌状态;
(b) 完全张开状态

首先来考虑一维柔性长链分子的情形。按第 3 章中柔性长链分子的无规行走模型，长链分子一端到另一端距离为 R 的配分函数是 $P(R)$ ，它给出了链构型与所有具有 R 的相似链组成的系综的比例。 $P(R)$ 具有正态高斯分布形式，当 $R=0$ 时，显然长链分子具有最大可几率的空间构型。换句话说，在 dR 体积内发现 R 矢量长度的长链分子的几率应具有高斯函数形式：

$$P(R)dR = Ae^{-b^2 R^2} dR \quad (14-21)$$

其中 A 为归一化因子。

由上式得到 $R \rightarrow R'$ 引起的熵变：

$$\begin{aligned} \Delta S &= k_B \ln P(R') - k_B \ln P(R) \\ &= k_B \ln A e^{-b^2 R'^2} - k_B \ln A e^{-b^2 R^2} \\ &= -k_B b^2 (R'^2 - R^2) \end{aligned} \quad (14-22)$$

对应于自由能变化:

$$\begin{aligned}\Delta F &= \Delta H - T\Delta S \\ &= k_B T b^2 (R'^2 - R^2)\end{aligned}\quad (14-23)$$

考虑沿 x 轴的一维拉伸, 若两端在 x 轴向的间距为 R_x^* (另外两维不受限制), 则作用力为:

$$f = -T \frac{\partial \Delta S}{\partial R_x^*} \quad (14-24)$$

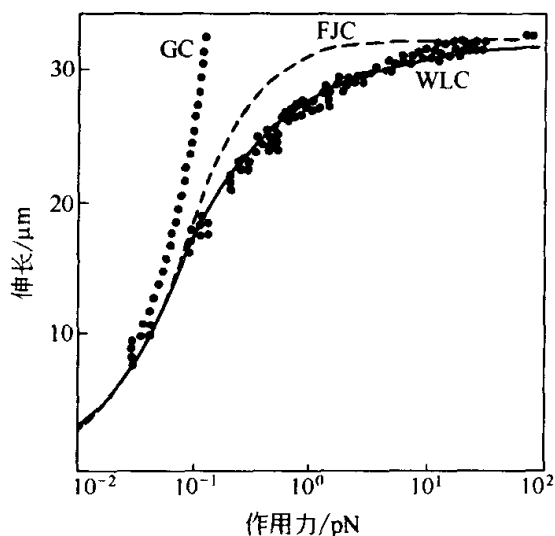


图 14-6 外力作用下的 DNA 分子的弹性性质

图中曲线 GC 为高斯链模型的计算值,

FJC 为自由连接链的计算值,

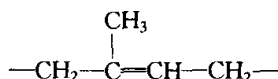
WLC 为蠕虫链的计算值

在作用力不大的情况就会得到胡克定律。

以前由于技术手段的限制, 要对单根柔性分子链进行拉伸试验是困难的。但近年来随着光镊、微探针等新技术的发展, 人们成功的实现了对单根 DNA 分子链进行拉伸实验。其结果如图 14-6 所示。在形变初期, 曲线与上述高斯链 (GC) 模型的结果大体吻合。当形变量较大时就有些偏离, 这时就需要考虑到作用力对链段的顺向排列作用 (类似于磁矩在磁场中排列起来的顺磁性效应), 这样就提出了自由连接链 (FJC) 模型, 即将长链分子视为用枢点连接起来的一段段刚性短棒。

其结果与实验在中形变区吻合的很好。再进一步假设枢点连续的分布在链上, 就得到了蠕虫链 (WLC) 模型, 这个模型在大形变区域能很好的说明实验结果。

DNA 分子是一维的柔性链状结构, 掺硫的橡胶则是三维的网状结构, 同样也具有高弹性。天然橡胶接近于液态, 产自热带的橡胶树, 其长链分子的单体为



在橡胶中掺入了 0.5% ~ 3% 的硫所起的作用是在长链分子间引入了许多交联点 (cross-links), 如图 14-7 中所示的黑点。

橡胶的拉伸使交联点间的分子线段变直 (图 14-7), 但基本上不影响分子中的原子间距。将弯曲的分子线团拉直, 导致分子线段的位形熵减小, 也

就是说拉伸结果是使有序度增加,因而外力的作功会使熵减小,从而增大了自由能,橡胶的弹性形变是和熵联系在一起的。基于熵变化的弹性理论,可以求出橡胶的弹性模量作为交联点数目函数的函数(不引入任何经验参量),与实测结果大体相符(准确度约为25%),这充分表明上述橡胶形变和结构变化之间的关系是可靠的。这一结果也说明了

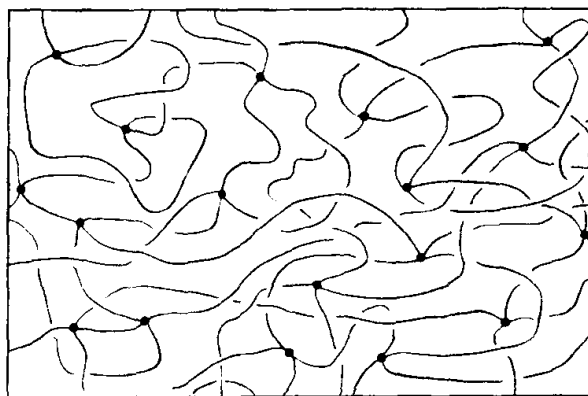


图 14-7 橡胶的交联无规网络模型

橡胶作弹性形变反而导致了有序度的增加, X 射线衍射实验也证实了这一点。有迹象表明形变会导致结晶化。考虑到三维的变形有:

$$\begin{aligned} R' &= (\lambda_1 x, \lambda_2 y, \lambda_3 z) \\ \Delta F &= kTb^2 [(\lambda_1^2 - 1)x^2 + (\lambda_2^2 - 1)y^2 + (\lambda_3^2 - 1)z^2] \\ &= kTb^2 R^2 (\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3) \end{aligned} \quad (14-25)$$

若沿 x 方向拉伸, 则由 $\lambda_1 \lambda_2 \lambda_3 = 1$ 和 $\lambda_2 = \lambda_3$, 可得:

$$\Delta F = kTb^2 R^2 \left(\lambda^2 + \frac{2}{\lambda} - 3 \right) \quad (14-26)$$

由此:

$$f = 2kTb^2 R^2 \left(\lambda - \frac{1}{\lambda^2} \right) \quad (14-27)$$

根据上式, 在比较大的变形情况下 ($\lambda > 1$) 应力-应变关系更接近于线性关系。如果形变量甚大, 高斯链模型也就不适用了, 而需要采用更加接近现实的 FJC 或 WLC 模型。

材料的弹性行为由胡克定律描述, 区分材料弹性特征参数包括两个方面, 一是弹性模量, 二是相对变形的量。陶瓷材料具有很高的硬度, 但相对变形量很小; 橡胶具有很小的弹性模量, 但相对变形量很大; 金属则介于陶瓷和高分子橡胶之间。

14.1.4 粘弹性

实际上自然界中的任何物体均同时具有弹性和粘性这两种性质, 根据外加条件不同 (外载时间和温度), 或主要显示弹性或主要显示粘性。弹性体和粘性体的区别主要是: 在外力作用下的形变的时间依赖关系不同。理想弹性体的形变与应力作用时间无关, 而理想粘性体的形变与应力作用时间呈线性关系, 高分子材料则处于二者之间具有粘弹性。

粘弹性是高聚物材料具有的一个重要的特性。当温度超过流动转变温度 T_f 时, 线性高聚物就开始熔融, 变为流动态。这时所形成的熔体不但会像牛顿流体那样表现出粘性流动, 还会呈现出相当明显的弹性行为。

高聚物的力学性质随时间发生的变化通称为力学松弛, 其中高聚物的蠕变和应力松弛都是重要的力学松弛现象。

蠕变描述的是在一定的温度和应力作用下, 高聚物的形变随时间的变化。图 14-8 是高聚物典型的蠕变曲线。其中 t_1 , t_2 分别代表外应力开始作用和解除作用的时间。在外应力 σ 作用下, 发生蠕变时的总形变 ϵ 可以用下式表示:

$$\epsilon = \epsilon_1 + \epsilon_2 + \epsilon_3 = \frac{\sigma}{E_1} + \frac{\sigma}{E_2} [1 - \exp(-t/\tau^*)] + \frac{\sigma}{\eta_3} t \quad (14-28)$$

式中, ϵ_1 是由分子链内部键长和键角发生变化所引起的弹性形变; ϵ_2 是高弹性部分, 即由于分子链逐渐伸展而带来的弹性性质, 与时间呈指数关系; ϵ_3 为分子间的相对滑移造成的粘性流动所带来的形变。 ϵ_1 、 ϵ_2 是可回复的形变, ϵ_3 是不可回复的形变。另外, 式中 τ^* 是松弛时间, η_3 为粘度系数。

一般来说, 高聚物在玻璃化温度以下时, 分子链段运动的松弛时间 τ^* 很长, 粘度系数 η_3 很大, 因此 ϵ_2 和 ϵ_3 都很小, 这时形变量主要决定于 ϵ_1 ; 当温度高于玻璃化温度时, τ^* 随着温度的升高而减小, 这时 ϵ_1 和 ϵ_2 是形变的主要组成部分; 温度升到粘性流动温度以上时, τ^* 和 η_3 都减小, ϵ_1 , ϵ_2 和 ϵ_3 都比较显著。

在温度和形变保持不变的情况下, 高聚物内部的应力会逐渐衰减, 这就是粘弹性的应力松弛现象。这时, 应力和时间呈指数关系:

$$\sigma = \sigma_0 \exp(-t/\tau^*) \quad (14-29)$$

式中, σ_0 是起始应力, τ^* 是松弛时间。图 14-9 是典型的高聚物的应力松弛曲线, 图中不同种类的高聚物均显示了应力与时间的指数依赖关系。和蠕变一样, 应力松弛也来源于高聚物中分子链段在外应力作用下的运动, 即分子链段会沿着外应力的方向运动并排列, 从而减小或完全消除了内部的应力。

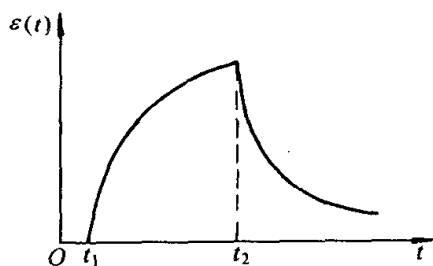


图 14-8 高聚物典型的蠕变曲线

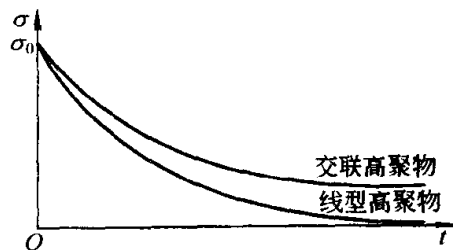


图 14-9 高聚物的应力松弛曲线

由此可见松弛时间 τ^* 和温度是密切相关的。

14.2 塑性形变与粘性流变

屈服是塑性形变的开始。因而探讨塑性形变的物理本质首先需要考虑屈服的问题。

大家知道,有些材料是硬的,而且硬度极高,但非常脆,像金刚石、碳化硅、氧化铝等;而另一些材料却很软,塑性良好,易加工成型,但问题是强度不高,像铜、银、金等纯金属,还有一些材料似乎兼具软硬两种材料的优点,既能加工成型,又有较高的强度,是生产实践中广泛应用的结构材料,像钢铁、硬铝等。

不同材料屈服应力与温度关系的实验结果如图 14-10 所示。为了比较方便起见,我们将图中的温度坐标采用等效温度(即试验温度 T 与试样的熔点温度 T_m 的比值)。根据图中所示的结果,可以大致将晶体材料分为两类:

① 所有温度下屈服强度都不高,表明滑移很容易,塑性良好。一般的面心立方金属如铜、铝等,以及密集六角金属(例如镁)沿基面的滑移,属于这一类。

② 只有在高温 ($T > 0.4 T_m$) 才表现出塑性,在 $T < 0.4 T_m$ 范围内,屈服强度急剧上升,表现出硬而脆。共价键晶体硅和锗、离子晶体 Al_2O_3 、金属间化合物 $NiAl$ 等属于这一类。

另外有一些材料其情况介于①、②两类之间。屈服强度在低温 ($T < 0.4 T_m$) 很高,表现脆性,在室温以上就显著降低,塑性接近典型的金属。体心立方的过渡金属,如铁、钨、钼、铌、钽等都属于这一类。

14.2.1 晶体的理论屈服强度

计算晶体的理论屈服强度 τ_1 实际上是研究使一块完整晶体发生塑性形变需要多大的切应力。

若两层原子面之间发生刚性滑移,相对位移值为 u , 如图 14-11 所示。令 a 表示原子层的间隔, b 表示滑移方向的点阵周期。当两层原子错开时,会引起点阵恢复原状的趋势,相应的切应力 τ 应该是位移 u 的周期函数

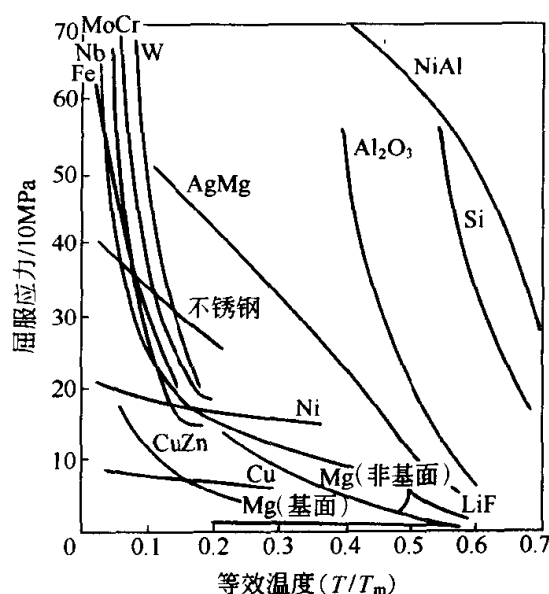
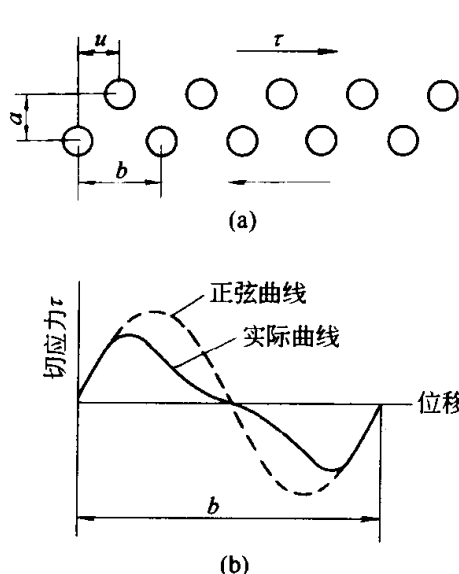


图 14-10 不同材料的屈服强度



(周期等于 b)，可用正弦函数近似地来表示它

$$\tau = k \sin\left(\frac{2\pi u}{b}\right) \quad (14-30)$$

k 是振幅，其数值可以按照胡克定律的条件求出来。当 u 很小时， $\sin\left(\frac{2\pi u}{b}\right) \approx \frac{2\pi u}{b}$ ，而此时 τ 和 u 满足胡克定律（引入切变模量 μ ）

$$k \sin\left(\frac{2\pi u}{b}\right) \approx k \frac{2\pi u}{b} = \mu \frac{u}{a} \quad (14-31)$$

图 14-11 两层原子间刚性滑移示意图 滑移能够持续进行下去的条件是外界所加切应力超过 τ 的振幅，所以不难看出理论屈服强度就等于 τ 的极大值，即

$$\tau_t = k = \frac{\mu}{2\pi} \cdot \frac{b}{a} \quad (14-32)$$

对于简单立方晶体， $b = a$ ， $\tau_t = \mu/6$ ；对于面心立方晶体作 $\{111\} \langle 112 \rangle$ 切变， $b = a_0/\sqrt{6}$ ， $a = a_0/\sqrt{3}$ (a_0 为点阵参量)， $\tau_t = \mu/9$ ；对于层状结构的石墨， $b/a = 0.4$ ， $\tau_t = \mu/15$ 。考虑到滑移面上的晶体结构及原子键合力的影响，人们对于 τ_t 的计算进行了改进，结果列在表 14-2 (a) 中。不难看出，根据式 (14-32) 估计的 τ_t 值，在数量级上和表 14-2 (a) 所列结果是相符合的，只是对面心立方金属的 τ_t 值估计偏高一些。晶体的理论屈服强度和实际屈服强度分别见表 14-2 (a)、(b)、(c)。

表 14-2(a) 理论屈服强度的计算值

材 料	滑移面和方向	切变模量 μ /10GPa	τ_t/μ	τ_t /GPa
铜(10K)	$\{111\} \langle 11 \bar{2} \rangle$	3.25	0.028~0.039	0.91~1.27
铜	$\{111\} \langle 11 \bar{2} \rangle$	3.02	0.028~0.039	0.84~1.18
金	$\{111\} \langle 11 \bar{2} \rangle$	1.86	0.028~0.039	0.52~0.72
银	$\{111\} \langle 11 \bar{2} \rangle$	1.93	0.028~0.039	0.54~0.75
铝	$\{111\} \langle 11 \bar{2} \rangle$	2.25	0.028~0.039	0.64~0.88
铝	$\{111\} \langle 1 \bar{1} 0 \rangle$	2.25	0.114	2.57
锌	$\{0001\} \langle 10 \bar{1} 0 \rangle$	3.7	0.034	2.25
铁	$\{1 \bar{1} 0\} \langle 111 \rangle$	5.9	0.11~0.13	6.5
钨	$\{1 \bar{1} 0\} \langle 111 \rangle$	14.7	0.11~0.13	15.9
金刚石	$\{111\} \langle 1 \bar{1} 0 \rangle$	49.5	0.24	119
硅	$\{111\} \langle 1 \bar{1} 0 \rangle$	5.6	0.24	13.4
石墨	$\{0001\} \langle 10 \bar{1} 0 \rangle$	0.23	0.05	0.113
NaCl	$\{110\} \langle 1 \bar{1} 0 \rangle$	2.32	0.120	2.77
Al ₂ O ₃	$\{0001\} \langle 10 \bar{1} 0 \rangle$	14.4	0.115	16.6

表 14-2(b) 晶须的屈服强度

材料	临界切应力 τ_m /GPa	τ_m/μ	材料	临界切应力 τ_m /GPa	τ_m/μ
铜	0.78	0.022	镉		0.033~0.069
银	0.69	0.031	NaCl	0.53	0.030
铁	3.51	0.060			

表 14-2(c) 高硬度晶体的屈服强度 (根据室温硬度测量数据估计)

材料	硬度 /10MPa	临界切应力 τ_m /GPa	材料	硬度 /10MPa	临界切应力 τ_m /GPa
金刚石	8400	27	BeO	1300	3.5
WC	2100	3.5	AlN	1225	3
TiC	2900	10	Al ₂ O ₃	2600	5.5
B	2500	6.5			

将理论屈服强度和实际晶体屈服强度的实验结果进行对比发现：前者比后者高好几个数量级。从表 14-2 (a) 上可以看出金属的 τ_i 的数量级为 GPa (即 10^3 MPa)，而实际金属单晶体的临界切应力是随试样的纯度、位错密度等的不同而有差异，一般的数量级在 $10^{-1} \sim 1$ MPa 之间 (例如高纯度铜单晶的实验值为 350~650 kPa)。考虑到金属晶体中位错在甚低的应力作用下就能运动，上述结果就表明实际金属单晶的屈服，这不是原子面间发生刚性滑移的结果，而是位错滑移的结果。据此推想，如果将金属晶体中的位错全部清除，并且将表面刻痕、台阶、夹杂物等其他可以引起应力集中的因素也完全去掉，它的屈服强度应该等于在晶体中均匀地形成位错圈所需的切应力，这就接近于 τ_i 。关于金属晶须 (直径为微米级的高度完整的细单晶体) 的实验结果证实了这一点 [参看表 14-2(b)]。但是晶须这种高强度性能是不稳定的，位错一旦产生以后，强度就会大大的下降。至于非金属 (离子键和共价键) 单晶体的实验结果则有两种情况：一种是 NaCl、LiF 等离子尺寸相差不大的离子晶体，与金属晶体相似，实际晶体的屈服强度远低于 τ_i ，说明位错是容易动的；另一类是高硬度的晶体，如金刚石，氧化铝等，室温下很脆，难以进行常规的拉伸试验，但其屈服应力可以间接的从硬度测量结果推出，表 14-2 (c) 列出了一部分实验结果。值得注意，求出的临界切应力和晶体中有无位错或位错的多少并没有关系且很接近于 τ_i ，例如金刚石的结果约为 τ_i 的 1/4。从这个实验结果可以推断，这类晶体中位错是很难运动的，滑移只有在切应力接近于 τ_i 的情况下才会发生。

通过以上的讨论，使我们了解到晶体的理论强度的实际意义：它标志着晶体屈服强度的上限，即晶体处于最硬状态 (也就是说原子键合强度充分发挥作用的状态) 下的屈服强度，它指出了各种强化金属方法的奋斗目标。至

于晶体屈服强度的下限，也就是晶体处于最软状况（即位错最容易运动的情况）下的屈服强度，就相当于晶体中其他缺陷全部被扫除以后使位错滑移所需的临界切应力。这就是位错的点阵阻力。

位错是点阵错排的区域，具有较高的能量，位错在晶体中运动将扫过点阵，显然，原子及原子之间的间隙对位错运动有不同的阻力。根据简化的佩-纳（Peierls-Nabarro）模型导出位错位于两种不同位置之间的能量差的幅值是

$$W_p = \frac{\mu b^2}{\pi(1-\nu)} \exp(-4\pi\xi/b) \quad (14-33)$$

由此得出：

$$\tau_p = \frac{2\mu}{1-\nu} \exp(-4\pi\xi/b) = \frac{2\mu}{1-\nu} \exp[-2\pi a/b(1-\nu)] \quad (14-34)$$

这里的 ν 为材料的泊松比， ξ 表示位错的半宽度。可以看出，位错的宽度极大的影响 τ_p 的取值。宽的位错，因为原子错排区域散开得很开， τ_p 的值就小。而式 (14-34) 中的位错宽度取决于滑移面上的晶体结构，也就是说，取决于参量 $b/2\pi a$ 。根据表 14-2 (a) 所列出的各种材料的 τ_t/μ 值，我们可以得出如下结论：密排金属的 τ_t/μ 值最小， τ_p 应该最小，体心立方金属的 τ_p 高一些，而金刚石、硅这些共价键晶体的 τ_p 应该最高。另一方面，泊松比 ν 的数值对 τ_p 也有影响，一般金属 ν 值较大（约为 0.3~0.4），相应的 τ_p 就小，锗、硅次之（ ν 约为 0.28）； Al_2O_3 、金刚石的 ν 值很小（约为 0.2 以下），相应的这些材料的 τ_p 就应该高。从物理本质上来看，归根结底，上述的这些因素均是键合类型对 τ_p 影响的反映：金属键没有方向性，因此键合能对近邻原子排列方位不敏感，对原子体积却很敏感，所以位错的 τ_p 很小；共价键（及部分的离子键）有强烈的方向性，在位错核心区域正常的价键往往不能保持，再加上价键的晶体不能采取密集结构，具有大的 τ_p 值本是理所当然的。一般说来，佩-纳力高的材料也正是理论屈服强度高的材料。以上关于佩-纳力的讨论限于低温（ $T < 0.4T_m$ ）。在高温（ $T > 0.4T_m$ ），位错的运动可以通过原子的扩散来实现，使得一些方向键晶体的佩-纳力大幅度下降。这可以解释为什么 Si、 Al_2O_3 在高温时屈服强度下降，表现出一定的塑性。

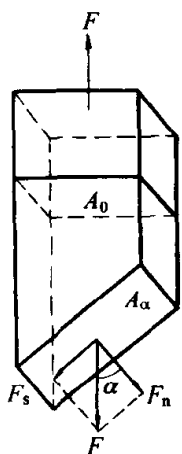


图 14-12 单轴
拉伸应力分析
示意图

综上所述，我们大体上按照键合强度，将晶体分为软硬两类：软的晶体，其位错的佩-纳力远低于理论屈服强度，位错很容易动，塑性良好。金属基本上都属于这一类。硬的晶体，佩-纳力接近于理论屈服强度，在室温位错基本上不

图 14-12 单轴
拉伸应力分析
示意图

能动, 呈现脆性。共价键晶体和一些氧化物晶体属于这一类。

14.2.2 晶体塑性形变的几何学与结晶学关系

(1) 塑性形变的几何学关系 塑性材料在发生形变时, 常可以看到与拉伸方向成大约 45° 角倾斜的切向滑移形变带。不论是高聚物还是金属晶体和非金属晶体, 在发生塑性形变时都有类似的现象, 这说明 45° 角的切向滑移形变带的来源与材料的结晶学结构无关。通过后面的分析, 我们知道实际上这种现象是由材料的几何学关系决定的。

为了简便起见, 我们考虑单轴拉伸的情况, 如图 14-12, A_0 为试样的横截面积, A_α 为试样上与 A_0 面夹角为 α 的任意一个截面。则作用在 A_α 面上的拉力 F 可分解为两个相互垂直的分力, 即沿平面法线的分力 F_n 和沿平面切线的分力 F_s 。显然, 与 A_α 面所对应的法向应力 σ_{an} 和 σ_{as} 分别为:

$$\sigma_{an} = \frac{F_n}{A_\alpha} = \frac{F}{A_0} \cos^2 \alpha \quad (14-35)$$

$$\sigma_{as} = \frac{F_s}{A_\alpha} = \frac{F}{A_0} \frac{\sin 2\alpha}{2} \quad (14-36)$$

由上两式易知, 在外加拉力一定的情况下, 截面 A_α 的法应力与切应力仅决定于截面的仰角 α 。图 14-13 为任意截面上的切应力和法应力与截面仰角 α 的关系图。可以看出, 当仰角 $\alpha = 45^\circ$ 时, 切应力有最大值 $F/2A_0$ 。一般说来, 塑性材料拉伸时, 其斜截面上的最大切应力会首先达到材料的抗切变强度, 于是在试样上就出现了与拉伸方向成 45° 角的切向滑移形变带。

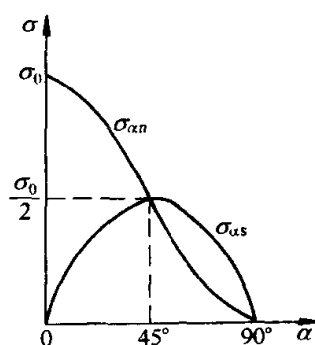


图 14-13 任意截面上的切应力和法应力与截面仰角 α 的关系曲线

对于仰角为 $\alpha + \pi/2$ 的斜截面 A_β , 即与 A_α 垂直的斜截面, 我们可以做完全类似的分析。结果发现, 在两个相互垂直的斜截面上的切应力数值相等, 方向相反。

$$\sigma_{as} = -\sigma_{\beta s} \quad (14-37)$$

这就是塑性材料在发生形变时总是对称形成细颈现象的原因。

与塑性材料不同, 脆性材料的试样断裂时断面通常与拉伸方向垂直。这是因为材料的切应力在达到抗切变强度之前, 正应力已经超过材料的抗张强度, 试样来不及发生屈服就断裂了。

(2) 晶体塑性形变的结晶学关系 非晶体材料的塑性形变受几何学条件的影响, 而晶体材料的塑性形变在受到几何学条件影响的同时, 必然还会受到结晶学条件的限制。例如, 金属具有优良的延展性, 易于加工成形, 这是

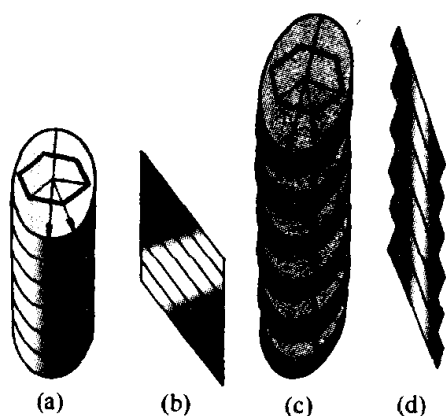


图 14-14 晶体滑移模型

(a) 和 (b) 原始状态；
(c) 和 (d) 拉伸以后

它区别于非金属材料的一个重要特征。其塑性形变的特征就与晶体的结构和金属键性质密切相关。

金属单晶的拉伸实验表明，单晶表面总是出现许多带纹(实际为台阶)，由此可以设想晶体的塑性形变是以滑移方式进行，滑移的意义是晶体的一部分相对于另一部分沿着晶面做相对移动。图 14-14 示出了一个圆柱形六角晶体的滑移模型，滑移面就是六角晶系的基面(0001)，它和拉伸轴斜交，与圆柱面相交截就表现为图中所示的椭圆。图中的短箭代表滑移方向，一般不一定和椭圆长轴重合。图 14-14 (a)、(b)

示出滑移前的形态，图 14-14 (c)、(d) 示出滑移后的形态。可以看出，试样经过滑移之后伸长了，它的横截面(和拉伸轴相正交的截面)也发生了变化，在某一方向上缩小很显著，而在其垂直方向可能略有扩大(扩大是由于滑移方向和滑移椭圆长轴不一致的缘故)，就从原来的圆形横截面变为椭圆形，而总的横截面积产生了收缩。

由于滑移时有些晶体中产生笛塔声音，而且完整滑移线常是在突然间形成的，因而最早关于滑移机制的假想是：滑移面上的原子面如同刚性的整体而彼此滑过。这种滑移过程中原子面内原子必须同时移动，这样滑移所需要的临界切应力，从理论上计算得的数值比实际数值大 1000 倍左右，因而人们设想滑移是原子逐步滑移的过程，即位错运动的结果。这样的设想已为大量实测所证实。

实验表明，滑移通常是沿着一定的结晶面和结晶方向发生，而不受外加负载的影响，仅与晶体结构有关。这一定的结晶面和结晶方向，称为滑移面和滑移方向，一个滑移面和位于该面中的一个滑移方向构成一个滑移系统。一般滑移面总是原子排列最密的面。类似的现象也同样存在于无机非金属晶体材料中。表 14-3 中分别列举了一些重要的金属晶体和非金属无机晶体的滑移系统。

晶体是否产生滑移实际上决定于临界切应力，临界切应力是指晶体在应力作用下开始产生滑移时，该应力在滑移面和滑移方向上的分量。实验证明不同取向的单晶体在不同的拉伸应力作用下开始滑移，但这些应力在滑移面和滑移方向上的分量是完全相同的，即具有相同的临界切应力。图 14-15 是拉伸应力与有效滑移切应力的关系示意图。设 p 为沿着拉伸轴方向的拉力，

表 14-3 一些重要的金属晶体和非金属无机晶体的滑移系统

材 料	点阵类型	温度/℃	滑移面	滑移方向
Al, Cu, Ag, Au, Ni, Pb	面心立方	20	{111}	$\langle 110 \rangle$
Al		$\frac{T}{T_m} > 0.72$	{100}	$\langle 110 \rangle$
α -Fe, Mo, Nb	体心立方		{110} {112} {123}	$\langle 111 \rangle$
α -Fe + 4% Si			{110}	$\langle 111 \rangle$
Zn, Cd, Mg	六角	20	(0001)	$\langle 11\bar{2}0 \rangle$
Mg		$\frac{T}{T_m} > 0.54$	{10 $\bar{1}1$ }	$\langle 11\bar{2}0 \rangle$
Mg + 14% Li		20	(0001){10 $\bar{1}0$ }	$\langle 11\bar{2}0 \rangle$
α -Ti		20	{10 $\bar{1}0$ }{10 $\bar{1}1$ }	$\langle 11\bar{2}0 \rangle$
C(钻石), Si, Ge			{111}	$\langle 1\bar{1}0 \rangle$
NaCl, LiF, MgO, NaF		低温	{110}	$\langle 1\bar{1}0 \rangle$
NaCl, LiF, MgO, NaF		高温	{110} {001} {111}	$\langle 1\bar{1}0 \rangle$ $\langle 1\bar{1}0 \rangle$ $\langle 1\bar{1}0 \rangle$
TiC		高温	{111}	$\langle 1\bar{1}0 \rangle$
PbS, PbTe			{001} {110}	$\langle 1\bar{1}0 \rangle$ $\langle 001 \rangle$
CaF ₂ , VO ₂		低温	{001}	$\langle 1\bar{1}0 \rangle$
CaF ₂ , VO ₂		高温	{001} {110} {111}	$\langle 1\bar{1}0 \rangle$ $\langle 1\bar{1}0 \rangle$ $\langle 1\bar{1}0 \rangle$
C(石墨), Al ₂ O ₃ , BeO			{0001}	$\langle 11\bar{2}0 \rangle$
TiO ₂			{101} {110}	$\langle 10\bar{1} \rangle$ $\langle 001 \rangle$
MgAl ₂ O ₄			{111}	$\langle 1\bar{1}0 \rangle$

p_0 为开始滑移时的拉力，阴影部分为滑移面， T 为滑移方向。作用于晶体上的拉力 p 实际上是要投影到滑移面和滑移方向上去，最后由临界切应力来决定 p_0 的大小。图 14-14 中所示的一叠钱币状的滑移就是由临界切应力决定的，是在临界切应力的作用下位错运动的结果。

$$\tau = \sigma_0 \cos \varphi \cos \lambda \quad (14-38)$$

其中 $\sigma_0 = \frac{p_0}{A}$ ， φ 为滑移面与横截面之间的夹角， λ 为滑移方向与拉伸轴之间的夹角，如图 14-15 所示。

这些观点首先在单晶体材料的实验中得到证实，而后在多晶材料的单个

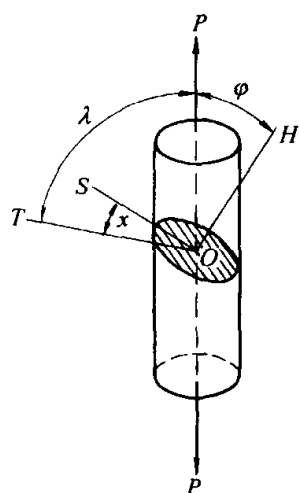


图 14-15 拉伸应力与有效滑移切应力的关系

晶体中也观察到了同样的现象。实际上这是多种滑移系统在相互协调的基础上运动的结果，因而存在平均的取向因子，总的应力-应变的关系是一个比较复杂的平均效应，其中晶界对位错的阻滞也起了一定的作用。

强烈的塑性形变还会引起多晶体中的晶粒取向呈现各向异性的分布，产生织构 (texture)。这个现象在轧制的板材和拉制的线材中特别显著。

14.2.3 塑性流变——从低温到高温

(1) 低温区的塑性形变 晶体塑性形变与温度有密切的关系，在低温区，即 $T < 0.4T_m$ (T_m 为熔点)，位错滑移是主要的形变机制。因而，屈服强度有效滑移切应力的关系

$$\tau_y = \tau_\mu + \tau(\dot{\gamma}, T) \quad (14-39)$$

式中， τ_μ 取决于位错间的相互作用，总的说来和温度关系不大（它和温度的关系是由于切变模量与温度有关而引入的）；另一项 τ 来源于位错越过短程障碍的热激活过程，它依赖于温度 T 和应变速率 $\dot{\gamma}$ ，被认为是热激活项。 τ_y/μ 被称为归一化的屈服强度，它和温度的关系可用图 14-16 来表示。在 0K 时，热激活不起任何作用，因而屈服强度较高。对于佩-纳力较高的材料，如共价键的半导体和 bcc 金属就有屈服强度较高的特征，这使得塑性形变不易产生。

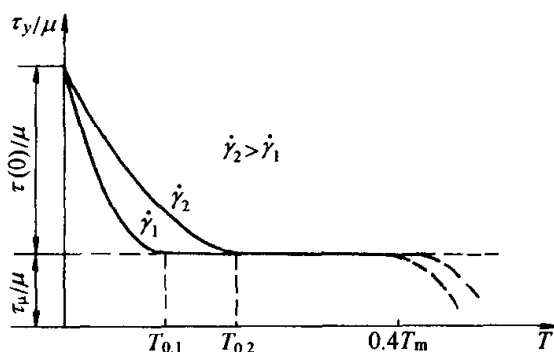


图 14-16 与切变模量 μ 归一化的屈服强度 τ_y 与温度 T 的关系
图中也显示了应变速率的影响

(2) 高温幂律蠕变 在高温区 $T > 0.4T_m$ ，原子的扩散和位错攀移的作用明显地表现了出来，屈服强度急剧下降。这时，蠕变的重要性就突出了。

高温时，金属或非金属晶体在恒定应力（或恒定载荷）下，除发生瞬时形变外，还要发生缓慢而持续的形变。这种形变量随着时间的延长而逐渐增加的现象，称为蠕变 (creep)。与定速的拉伸不同，蠕变描述的是在恒定的应力作用下，应变的速率随时间的变化。金属材料在低温条件下，若受到低于临界切应力的应力作用时，所产生的形变可用胡克定律来描述，该形变在应力不变的情况下，与作用时间的长短无关。但如果在高温条件下，由于蠕

变的发生就会产生完全不同的结果，也就是说其应变将随时间而继续变化。在持续应力的作用下，所有的固体材料高温时都会在一定程度上发生蠕变。

在固体中由于温度升高，位错的交截和交滑移，以及位错的攀移、空位的定向扩散和晶界滑移等，都能迅速进行。故蠕变现象随着温度的升高而越来越显著。在蒸汽轮机、燃气轮机、航空发动机、核动力装置以及石油工业、化学工业、宇航工业等的设备中，很多部件不但处在很高的温度下，而且经常经受着很大的应力。因而蠕变和蠕变断裂便成为材料性能的重要指标。

蠕变试验的结果一般用应变（伸长率）为纵坐标，时间为横坐标的蠕变曲线来表示。金属和非金属材料有相似的蠕变曲线，均可分为三个阶段。如图 14-17 所示即为一条典型的蠕变曲线。曲线的斜率 $\dot{\epsilon}$ 称为蠕变速率。

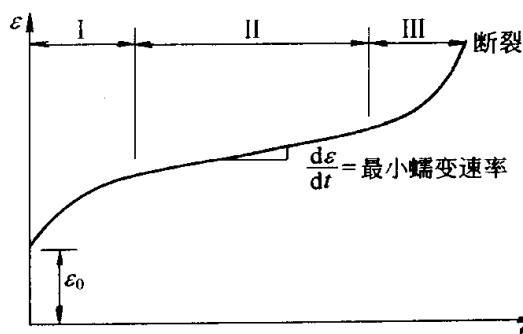


图 14-17 典型的蠕变曲线

由图可见，蠕变曲线可以分为用 I，II，III 来标志的三个阶段：起始的瞬态蠕变（transient creep）；中间的稳态蠕变（steady state creep）和最终导致的断裂。

蠕变的主要机制在于位错的攀移，这需要位错在晶格中的扩散，只有在高温区（ $T > 0.4 T_m$ ）才能实现。

在实验上观察到蠕变遵守一定的时间律，一般情况下蠕变量随着时间的变化可用下式来描述：

$$\dot{\epsilon} = \sum_i A_i t^{-n_i} \quad (14-40)$$

$\dot{\epsilon}$ 为蠕变速率， A_i 和 n_i 是依赖于温度和应力的常数，并且在第 I、II 阶段 $0 \leq n_i \leq 1$ 。

(3) 扩散蠕变 在更高的温度，低的应力下（小于屈服应力），或者在位错能动性差的情况下，位错滑移难以进行时，空位的定向扩散将成为蠕变的主要机制。其一为 Nabarro-Herring 蠕变，如图 14-18 所示。基本机制是：外加拉应力不同取向的晶界，其空位的形成能有了不同的变化，从而导致不同取向的晶界的空位平衡浓度不同，因而形成了在晶粒内部空位扩散。

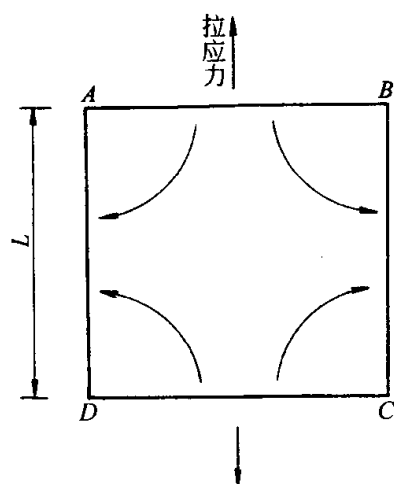


图 14-18 晶粒内应力产生的空位扩散流

物质流动的方向与空位流动方向相反，其结果就晶粒沿拉伸方向发生应变。其蠕变速率为：

$$\dot{\epsilon} = \frac{J}{L} = \alpha \frac{D\sigma b^3}{L^2 kT} \quad (14-41)$$

式中， α 是常数，数量级为 1， D 为自扩散系数， L 为晶粒尺寸。

上式说明，扩散蠕变的速率与晶粒尺寸的平方成反比，与应力的一次方成正比，其激活能与体扩散激活能相同。银、金、铜和铜铈合金的高温低应力蠕变，以及位错能动性很低的以氧化铝和氧化铍为基的陶瓷材料的蠕变，都曾观察到这样的规律。

上述的扩散蠕变机制是假定空位在晶粒内部作定向扩散，形变速率受体扩散控制。在实际情况中还存在另一种扩散蠕变机制，即空位沿着晶界的扩散也可产生蠕变，叫 Coble 蠕变，该机制假定晶粒是球形的。在拉伸应力作用下，晶界球面的极点与赤道处的空位浓度将有差异。此浓度差异导致空位由极点向赤道沿着晶界面扩散，这样便可导出蠕变速率为

$$\dot{\epsilon} = \frac{\alpha\sigma b^3 \delta D_b}{L^3 kT} \quad (14-42)$$

这里的 D_b 是晶界扩散系数， δ 是晶界厚度， $\dot{\epsilon}$ 与晶粒尺寸的三次方成反比。Coble 蠕变对陶瓷材料尤为重要，因为陶瓷材料在高温下的塑性形变通常都是空位沿着晶界扩散而实现的。

图 14-19 所示为处于不同温度和应力状态下多晶纯镍发生塑性形变的机制图。从图中可一目了然各个区域中塑性形变的机制。

(4) 塑性流变的稳定性与超塑性 塑性流变中，通常会由于形变集中在局部，出现细颈而导致断裂。因此拉伸塑性流变是否能够始终均匀地进行，就需要了解流变的稳定性问题。

在低温形变中位错密度的增加导致了加工硬化，其系数 θ 对应变速率是敏感的，有

$$\theta = \frac{1}{\sigma} \left(\frac{\partial \sigma}{\partial \dot{\epsilon}} \right)_\epsilon \quad (14-43)$$

流变速率系数 m 也是如此

$$m = \left(\frac{\ln \sigma}{\ln \dot{\epsilon}} \right)_\epsilon \quad (14-44)$$

因而拉伸塑性形变稳态进行的条件可以表示为：

$$\theta + m \geq 1 \quad (14-45)$$

① 在低温区， $\theta \gg m$ ，稳态流变的判据即变为

16.1 吉布斯相律

吉布斯相律 (Gibbs phase rule) 是处于热力学平衡状态的系统中自由度与组元数和相数之间关系的规律。通常简称相律。由 J.W.Gibbs 于 1876 年导出。相律中有以下几个基本概念:

(1) 相 系统中性质与成分均匀的一部分。相有自身的物理和化学特性, 并且理论上可以机械分离的。相与相之间有界面隔开。相可以是固态、液态或气态。由于气体是互溶的, 平衡系统中的气相数只能为 1。但液相和固相则可能有两种或两种以上。材料中相的种类、大小、形态与分布构成了材料的显微组织。

(2) 相平衡 多相体系中, 所有相的强度性质 (如温度、压强、每个组分在所有相中的化学位等) 均相等, 体系的性质不会自发地随时间变化的状态即相平衡状态。根据热力学第二定律, 具有给定物质与能量的热力学平衡体系 (孤立体系) 的自发过程总是朝熵增加的方向进行, 因此孤立体系位于平衡状态的熵大于处于非平衡状态的熵。根据这一原理, 对于物质一定但与外界有能量交换的体系 (封闭体系), 恒温恒压过程总是朝吉布斯 (Gibbs) 自由能降低的方向进行, 平衡状态吉布斯自由能最低。因此恒温恒压下多相体系中吉布斯自由能最低的状态就是相平衡状态。

(3) 组元 决定各平衡相的成分, 并且可以独立变化的组分 (元素或化合物)。如果系统中各组分之间存在相互约束关系, 例如化学反应等, 那么组元数便小于组分数, 也就是说, 在包含有几种元素或化合物的化学反应中, 不是所有参加反应的组分都是这个系统的组元。不过在许多合金系统中, 组元数往往就等于构成这个系统的化学元素的数目。

(4) 自由度 可以在一定范围内任意改变而不引起任何相的产生与消失的最大变量数, 又称为独立变量数。决定系统平衡状态的变量, 包括系统的组成和各种外界条件, 如温度、压强、电场、磁场和引力场等。但一般情况下除温度和压强外, 其他外界条件对复相平衡的影响甚微, 可以忽略不计。因此, 外界条件通常仅指成分、温度和压强。

相律指出, 在任何热力学平衡系统中, 自由度 F 、组元数 C 和相数 P 之间存在如下关系

$$F = C - P + 2 \quad (16-1)$$

式(16-1)中 2 反映外界温度和压强的影响。若影响系统平衡状态的外界条件有 n 个, 则上式中的 2 改为 n :

$$F = C - P + n$$

譬如, 当研究凝聚态系统时, 如果压强变化不大, 可略去压强这一变

量, 这时 $n = 1$, 得到常压下凝聚态系统的表达式:

$$F = C - P + 1$$

相律可以用以下方法进行推导:

设有一个平衡体系含有 C 个独立组分, P 个相, 分别为 α 、 β 、 γ …… ϕ 那么一共有 C 乘 P 个成分变量, 其中只有 $P(C-1)$ 个是独立的, 加上温度与压力两个外部条件, 系统总变量数为 $P(C-1)+2$ 。而根据相平衡条件, 每一组元在各相中的化学位 μ 相等, 即:

$$\begin{cases} \mu_{1,\alpha} = \mu_{1,\beta} = \mu_{1,\gamma} = \cdots = \mu_{1,\phi} \\ \mu_{2,\alpha} = \mu_{2,\beta} = \mu_{2,\gamma} = \cdots = \mu_{2,\phi} \\ \mu_{3,\alpha} = \mu_{3,\beta} = \mu_{3,\gamma} = \cdots = \mu_{3,\phi} \\ \mu_{C,\alpha} = \mu_{C,\beta} = \mu_{C,\gamma} = \cdots = \mu_{C,\phi} \end{cases} \quad (16-2)$$

共有 $C(P-1)$ 个独立等式, 所以对于一个处于平衡状态的体系, 其独立变数的数目, 即自由度的数目 F 应等于变数的总数与平衡方程式的数目之差, 即

$$F = P(C-1) + 2 - C(P-1) = C - P + 2$$

如果有一个组分在某一相或若干相不存在, 则变数的数目少了一个或若干个, 但相应地也少了一个或若干个化学位等式, 因此体系的总变数与方程式数目之差仍为 $C - P + 2$, Gibbs 相律仍然成立。对平衡体系, 自由度必须大于或等于零。

相律是相图的基本规律之一, 任何相图都必须遵从相律。但应该指出, 相律只是对可能存在的平衡状态的一个定性描述。它可以给出一个相图中可能有些什么点、什么线和什么区, 却不能给出这些点、线、区的具体位置。

16.2 杠杆定律

杠杆定律是质量守恒定律的一种表达形式, 是通过相图来确定各相含量的重要表达式。在 n 元体系中, 如果已知成分为 O 的合金由 α 与 β 两相组成, 那么 O 点的成分必然位于 α 与 β 两相成分点的连线上, 而且任一相的相对含量等于另一相成分点到 O 点的连线的长度与两相成分点连线长度之比。在 n 元体系中, 如果已知成分为 O 的合金由 α 、 β 与 γ 三相组成, 那么 O 点的成分必然位于 α 、 β 与 γ 三相成分点所构成的三角形内, 而且任一相的相对含量等于另两相成分点与 O 点所构成的三角形的面积与 α 、 β 、 γ 三相成分点所构成的三角形面积之比。同样, 在 n 元体系中, 如果已知成分为 O 的合金由 α 、 β 、 γ 、 δ 四相组成, 那么 O 点的成分必然位于 α 、 β 、 γ 、 δ 四相成分点所构成的四面体内, 而且任一相的相对含量等于另三相成分点

与 O 点所构成的四面体的体积与 α 、 β 、 γ 、 δ 四相成分点所构成的四面体体积之比。依次类推。

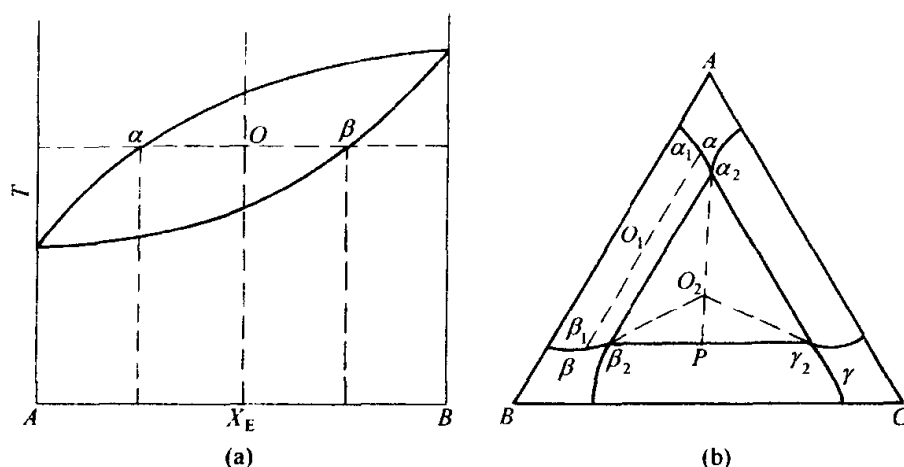


图 16-1 杠杆定律在二元相图与三元相图中的应用

(a) 杠杆定律在二元相图的应用；(b) 杠杆定律在三元相图中的应用

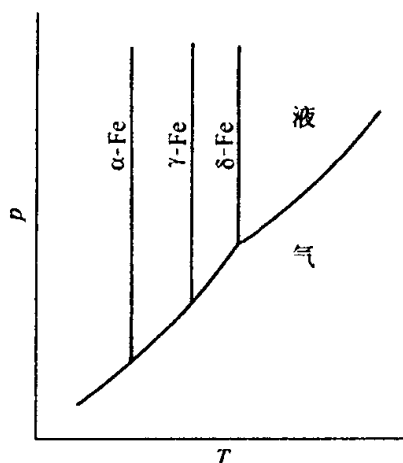
图 16-1 为杠杆定律在二元相图与三元等温截面上应用的示意图。在图 16-1 (a) 所示二元系中，如果成分为 O 的材料在 T_1 温度下处理，将分解为 α 与 β 两相，两相的质量比为： $m^\alpha/m^\beta = O\beta/O\alpha$ 。同样在图 16-1(b) 所示的三元系中，如果成分为 O_1 的材料在等温截面所代表温度下处理，将分解为 α 与 β 两相，两相的质量比为： $m^\alpha/m^\beta = O_1\beta_1/O_1\alpha_1$ ；如果成分为 O_2 的材料在等温截面所代表温度下处理，将分解为 α 、 β 与 γ 三相，三相的质量比为：

$$m^\alpha/m^\beta/m^\gamma = S_{O_2\beta_2\gamma_2}/S_{O_2\alpha_2\gamma_2}/S_{O_2\alpha_2\beta_2} = \frac{O_2P}{Pa_2} \cdot \frac{O_2a_2}{Pa_2} \cdot \frac{P\gamma_2}{\beta_2\gamma_2} \cdot \frac{O_2a_2}{Pa_2} \cdot \frac{P\beta_2}{\beta_2\gamma_2}.$$

16.3 单元体系的温度~压力图 (p - T 图)

由于相图测量过程中温度、压力、成分比其他热力学量更容易控制，因此常见的相图大都以温度 (T)、压力 (p) 与成分 (x_i) 为坐标。

组元数为一的体系称为单元系。单元系相图无成分坐标只有温度和压力坐标。根据相律 ($F = 1 + 2 - P = 3 - P$)，在温度与压力都可变的情况下，单元系最多有三相平衡。在单元系 $p \sim T$ 图上：单相区自由度为二，单相平衡点的集合构成任意形状的区域，温度与压力都独立可变；两相区的自由度为 1，两相平衡点的集合构成一根曲线，在这曲线上，两相的温度与压力都相等，温度与压力只有一个独立可变；三相区的自由度为零，是一个温度与压力都固定的点，在这一点上三相的温度与压力都相等。图 16-2 为纯铁

图 16-2 纯铁的 p - T 图

的 p - T 图, 其中 α -Fe 与 δ -Fe 是体心立方结构, 而 γ -Fe 是面心立方结构。单元体系中两相平衡线的斜率可由克拉珀龙-克劳休斯 (Clapeyron-Clausius) 方程:

$$\frac{dp}{dT} = \frac{\Delta H}{T\Delta V} \quad (16-3)$$

来确定。由于凝聚态平衡转变 ΔV 近似很小, 因此图中液-固与固-固平衡线都近似与纵坐标平行, 而液-气与固-气平衡 ΔV 较大, 因此两相平衡线为斜线。根据克拉珀龙-克劳休斯方程与相交的三条两相平衡线 ΔH 与 ΔV 之间的关系, 我们可以得出结论: 凡代

表两相平衡的曲线通过三相平衡点时, 它必然伸入第三相中, 这就是 180° 原理。

16.4 二元相图

系统组元数为 2 的相图即二元相图 (binary phase diagram)。与材料相关的凝聚态二元相图通常将压力视为恒定而以温度和成分为坐标来表示。根据相律 ($F = 2 + 1 - P = 3 - P$, F 为自由度, P 为相数, $n = 1$), 凝聚态二元系统中允许有无变度三相平衡 ($F = 0$, $P = 3$)、单变度的两相平衡 ($F = 1$, $P = 2$) 和双变度的单相平衡 ($F = 2$, $P = 1$) 存在。它们在相图中具有不同的几何特征: ①单相平衡点的集合构成任意形状的区域, 温度与相的成分都是独立可变的; ②两相平衡点的集合构成两根曲线。一根曲线上的每一点, 可由一水平的直线 (结线) 和另一曲线的共轭点相连, 结线通过的区域为两相区, 结线端点所代表的成分就是两平衡共存相的成分, 结线所示温度即两相平衡温度。两相平衡区由一系列这样的结线组成, 因此温度和相成分只有一个独立可变的; ③三相平衡点为温度相等、成分固定的三点。三点两两相连的结线所限制的水平线段为三相区, 它在这三点分别与三个单相区相连。三点两两相连的结线与三个两相区相连。三相区的温度与各相的成分都是固定不变的。三相平衡有共晶型和包晶型。前者由一相分解为另外两相, 后者由两相生成另一相。

二元凝聚态相图有以下几种基本类型。

16.4.1 匀晶相图

匀晶相图 (isomorphous system) 是指组元在液态与固态完全互溶的相图。图 16-3 (a) 为二元匀晶相图的相区分布图, 该相图存在一个液相区

(L)、一个固相区(S)和一个固液两相区(L+S)。在液相区域内, A 与 B 在原子尺度内相互混合形成均一的液相; 在固相区内, A 与 B 在原子尺度内相互混合形成均一的固溶体相; 固液两相区内为固液两相混合物。液相区与固液两相区的交界线为液相线, 它是两相平衡时的液相成分点的连线, 固相区与固液两相区的交界线为固相线, 它是两相平衡时固相成分点的连线。纯组元 A 和 B 凝固点分别为 T_A^0 和 T_B^0 , 所有成分的 A 和 B 的固溶体都存在一个凝固区间, 从液相线开始, 到固相线结束。当成分一定的固溶体随温度的升高或成分一定的液相随温度的降低进入两相区时, 两相的平衡成分由连接液相线与固相线的水平恒温线(结线)的端点来确定。例如: 图 16-3(b) 中, 成分为 X 的液相当温度降低到 t_1 时, 将开始形成固液两相混合物, 到 t_2 时, 液相的成分为 X_l , 固相的成分为 X_s , 由于析出了富 B 的液相, 因此, 液相中的 B 含量降低。两相区内固液两相的相对含量可由杠杆法则确定, 即 $m_l:m_s=(X_s-X):(X-X_l)$, 式中 m_l 和 m_s 分别为液相与固相的质量。 t_3 时, 凝固完毕, $X_s=X$, $m_s=1$, 形成成分与液相成分完全一致的固

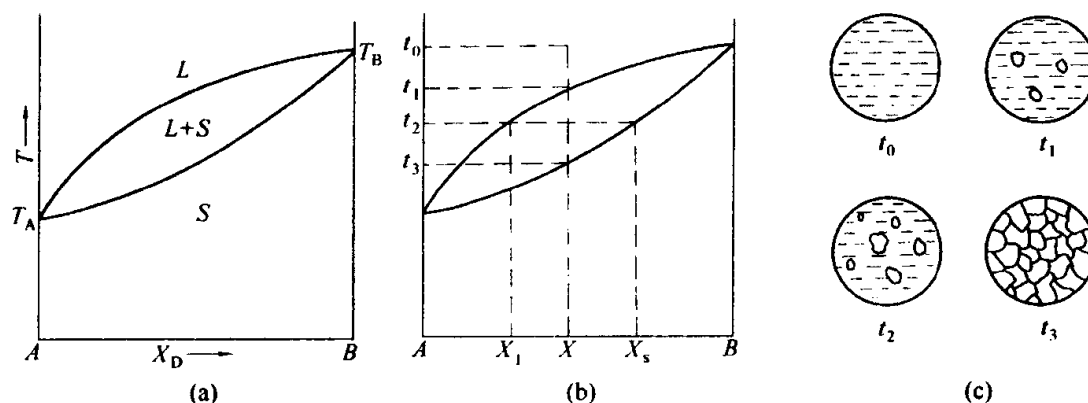


图 16-3 二元匀晶相图及相应二元合金凝固过程与组织演变示意图

(a) 二元匀晶相图的相区分布;

(b)、(c) 匀晶相图中合金凝固过程与组织演变示意图

溶体相。图 16-3(c) 为该成分合金凝固过程的组织演化示意图。有的匀晶相图的固相线与液相线相切, 形成具有最低点[图 16-4(a)]或最高点[图 16-4(b)]的相图。最高点与最低点必定是液相线与固相线的公切点和极值点, 液相线和固相线切线的斜率为零, 平行于成分轴, 这一原理称为克劳瓦劳夫(Knovolov)原理^[30]。在最高点与最低点处固液两相成分相等, 自由度为零。有的匀晶相图虽然在固相线以下固相形成连续固溶体, 但在低温下, 固溶体会分解, 形成结构相同但是成分不同的两相, 这种现象称为溶解度间隙 (miscibility gap), 见图 16-4(c)。

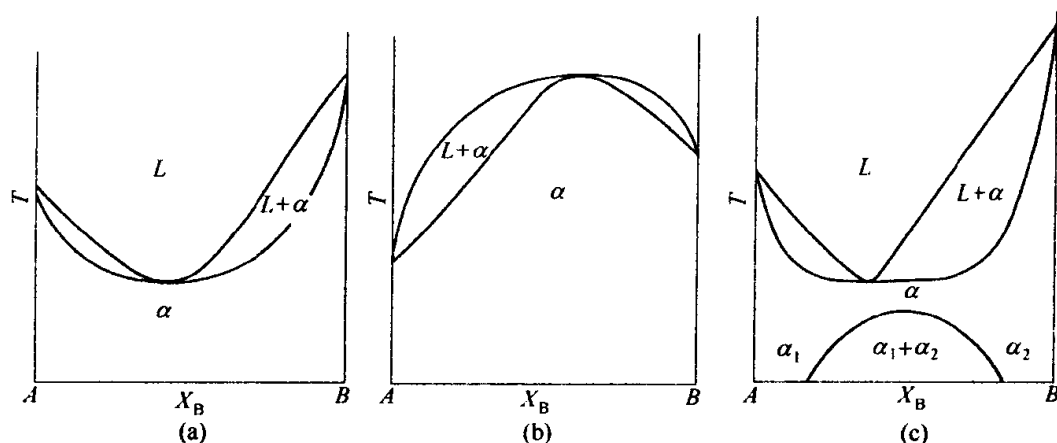


图 16-4 具有最低点、最高点与溶解度间隙的匀晶相图

(a) 具有最低点的匀晶相图；(b) 具有最高点的匀晶相图；

(c) 具有最低点与溶解度间隙的匀晶相图

16.4.2 共晶相图与包晶相图

共晶相图与包晶相图是液态完全互溶固态部分互溶的相图。图 16-5(a)、图 16-6(a) 分别为二元共晶相图与包晶相图的相区分布图。组元 A 和 B 在液相中完全互溶，在固相中仅部分互溶，形成端际固溶体 α 和 β 。相图中除单相区 L 、 α 和 β ，两相区 $L+\alpha$ 、 $L+\beta$ 和 $\alpha+\beta$ 外，还有一根水平恒温线，它表示该温度（共晶温度或包晶温度）下，存在一个零变度三相平衡反应。

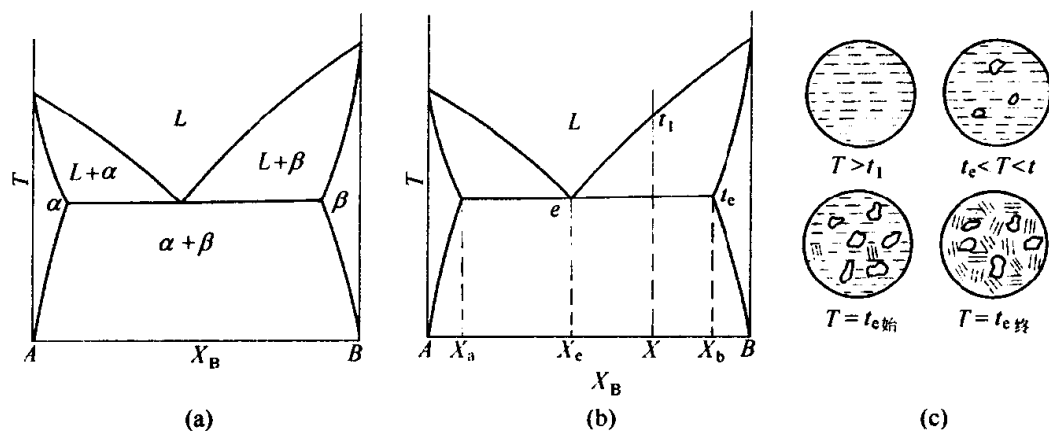


图 16-5 二元共晶相图及相应二元合金凝固过程与组织演变示意图

(a) 二元共晶相图相区分布；(b) 共晶相图中合金凝固过程；(c) 组织演变示意图

对共晶相图，该反应为共晶反应 (eutectic reaction)，其反应温度称为共晶温度。反应式为： $L(X_e) \longrightarrow \alpha(X_a) + \beta(X_b)$ 。参加反应的 L 、 α 和 β 相的成分分别为 X_e 、 X_a 和 X_b 。 e 点为共晶点（最低共熔点）。共晶点成分 (X_e) 的液相冷却至共晶温度时，液相恒温分解为 α 相和 β 相。反应后 α 相

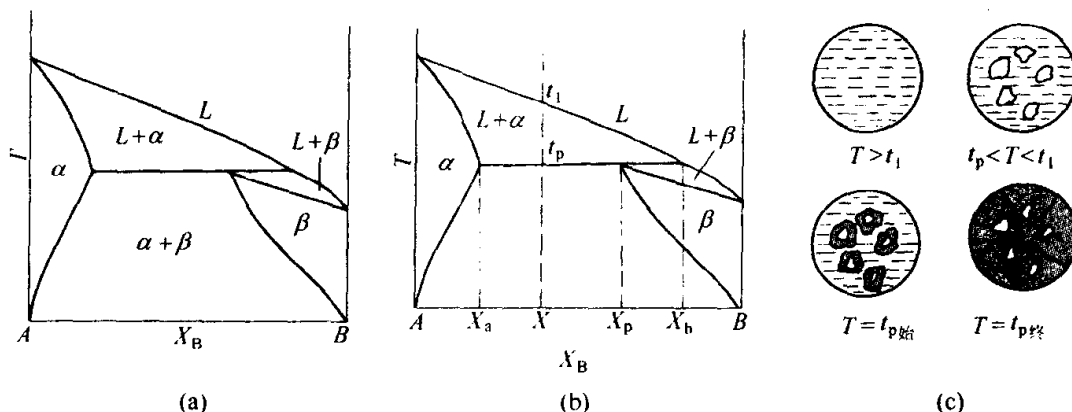


图 16-6 二元包晶相图及相应二元合金凝固过程与组织演变示意图

(a) 二元包晶相图相区分布；(b)、(c) 包晶相图中合金凝固过程与组织演变示意图

和 β 相的相对含量由前述杠杆法则决定。对于共晶范围($X_a \sim X_b$)内的其余成分的液相，例如：图 16-5 (b) 所示的成分 X 的液相，其凝固过程为随温度降低液相逐渐变冷，当液相温度达到两相区的液相线后，液相中开始析出 β 相初晶，温度进一步降低，液相成分沿液相线变化，固相的成分沿固相线变化，直至共晶温度，剩余液相成分为 X_e ，此时固液两相的相对量由杠杆法则 $m_e : m_\beta = (X_b - X) : (X - X_e)$ 确定，接着这部分液相共晶分解转变为 α 和 β 。凝固过程中其组织演化示意图见图 16-5(c)。

对包晶相图，该反应为包晶反应 (peritectic reaction)，其反应温度称为包晶温度。包晶反应反应式为： $L(X_p) + \alpha(X_a) \longrightarrow \beta(X_b)$ ，在包晶反应成分($X_a \sim X_b$)范围内，任一液相的凝固过程为，液相冷却至($L + \alpha$)两相区的液相线温度时，开始析出初晶相 α ，直到包晶温度，然后，余下的液相与初晶相 α 反应，在 α 与 L 的界面上生成 β 相。由于在反应过程中 β 相包裹着 α 相，所以这种类型的反应称为包晶反应。反应前后各相的相对量由杠杆法则确定。图 16-6(b)、图 16-6(c)分别为二元包晶相图中成分为 X 的合金凝固过程及其组织演化示意图。

如果我们将相图的亚稳部分画出来，可以看出，如果两端际固溶体晶体结构不同，则共晶与包晶相图都可以看作由三个匀晶相图组成[图 16-7(a)、(b)]；如果两端际固溶体晶体结构相同，则共晶与包晶相图都可以看作由一个匀晶相图与一个固相溶解度间隙 (miscibility gap) 组成[如图 16-7(c)、(d)]。

图 16-8 为 A-B 二元共晶相图与 A、B 两纯组元 p - T 图的关系图，二元共晶相图温度轴上的稳定与亚稳两相平衡点与 p - T 图上的一大气压下的稳定与亚稳两相平衡温度一一对应。

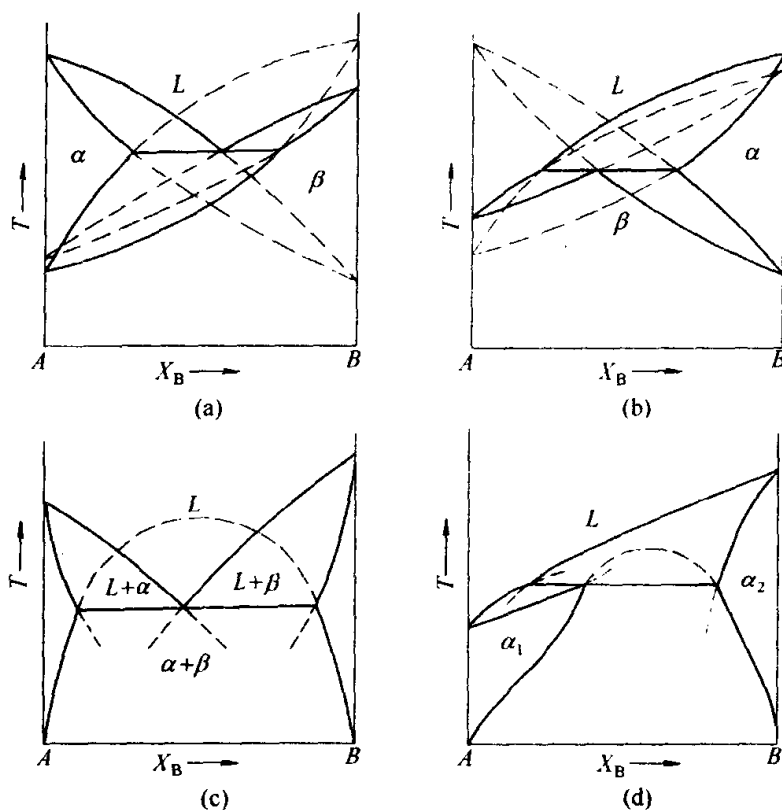
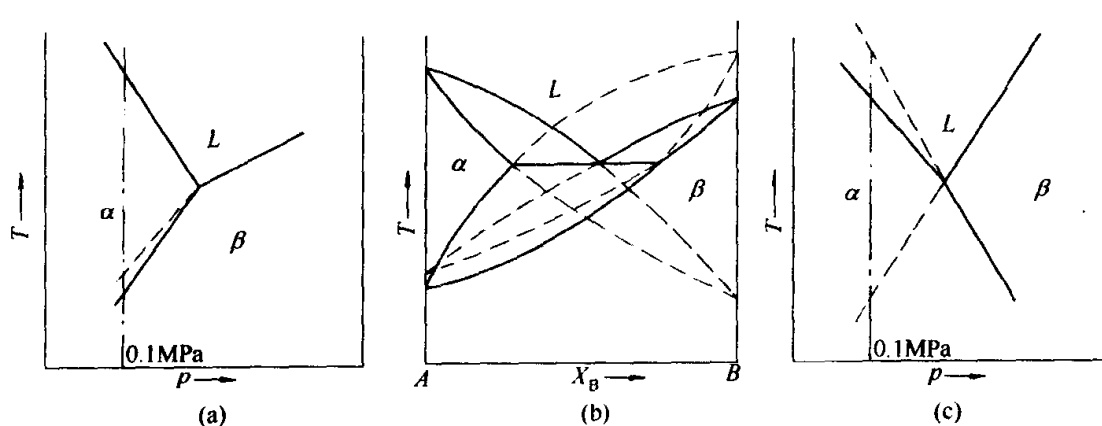


图 16-7 相图的构成

- (a) 两端际固溶体结构不同的情况下共晶相图的构成；
- (b) 两端际固溶体结构不同的情况下包晶相图的构成；
- (c) 两端际固溶体结构相同的情况下共晶相图的构成；
- (d) 两端际固溶体结构相同的情况下包晶相图的构成

图 16-8 二元共晶相图与两纯组元 p - T 图的关系图

- (a) 纯 A 的 p - T 图；(b) 端际固溶体结构不同的 A-B 二元共晶相图；(c) 纯 B 的 p - T 图

16.4.3 具有中间化合物的相图

如图 16-9 所示。A 和 B 组成的二元相图中存在两个中间化合物 β 和 γ 。

β 由液相直接同成分转变而来, 所以称为同成分共熔化合物。而 γ 由包晶反应 $L + \beta \rightarrow \gamma$ 转变而来, 是非同成分共熔化合物。有时我们也简称 β 是同分熔化的, 而 γ 是非同分熔化的。同成分共熔化合物 β 有一定的熔点, 也是 $L + \beta$ 两相区的极高点, 在这点 β 相的溶解度线与液相线相切, 自由度为零。而 γ 相却只有理论熔点, 在加热时未达到这一点便开始熔化, 所以我们有时讲非同成分

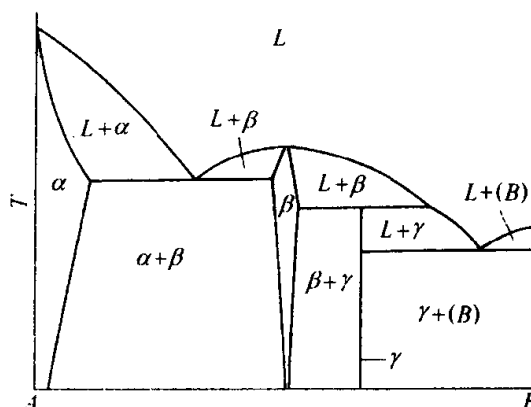


图 16-9 具有化合物的二元相图

分共熔化合物在高温下不稳定, 会发生分解。该图中还有两个共晶反应: $L \rightarrow \alpha + \beta$; $L \rightarrow \gamma + (B)$ 。此外, 为了说明中间化合物可以有成分范围也可以没有成分范围, 在图中 β 相画了成分范围, 而 γ 相没有画成分范围 (并非非同成分共熔化合物一定有成分范围, 而非同成分共熔化合物就一定没有成分范围)。同样为了说明并非所有的端际固溶体都有很明显的成分范围, 在 B 端没画成分范围, 表示 A 在 B 中的溶解度可以忽略不计。实际相图可以没有一个中间化合物, 也可以有七八个中间化合物。如果熔化后液相内不存在与化合物组成一致的原子集团, 中间化合物液相线在化合物熔点处切线斜率为零, 没有尖点, 所以有中间化合物的相图在本质上并非简单二元相图的叠加。

二元相图中, 固液相线在同分熔化处是否出现尖点与固相和液相的结构有关, 见图16-10。在图 16-10(a)所示的情况下, 可以认为固相与液相均由单个的原子混合而成, 固相与液相中无固定的原子集团; 在图16-10(b)所示

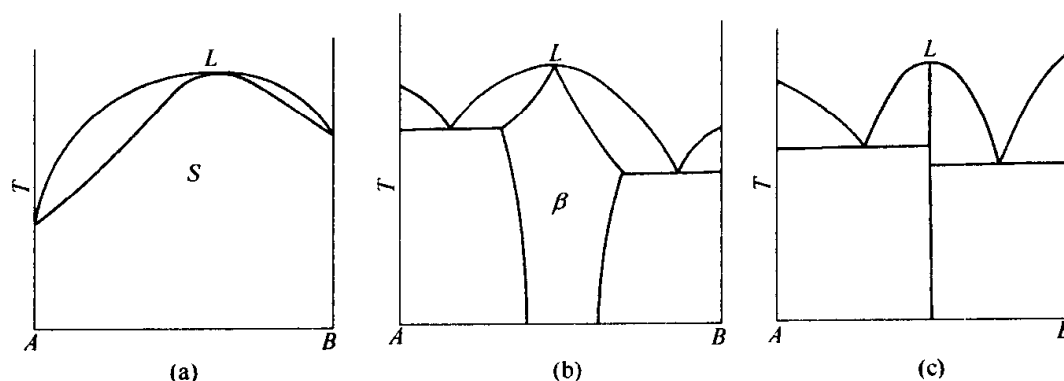


图 16-10 二元相图中同分熔化时固液相线的形状

- (a) 固相线与液相线相切, 都无尖点; (b) 固相线有尖点而液相线无尖点;
(c) 固相线与液相线都有尖点

的情况下, 可以认为固相中有固定的原子集团, 液相中无固定的原子集团; 在图16-10(c)所示的情况下, 可以认为固相与液相中都有固定的原子集团^[13]。

16.4.4 偏晶相图与综晶相图

偏晶相图与综晶相图是指液相不完全互溶(具有溶解度间隙)的相图。在有些情况下, 液相也只能部分互溶, 随温度的降低, 液相会出现分层现象, 由一种液相分解为成分与黏度都不相同的两种液相, 在相图表现为液相存在一个溶解度间隙。常见液相存在溶解度间隙的二元相图有两种: 偏晶相图与综晶相图。

在偏晶相图中[图 16-11(a)], 液相存在由一种液相分解为另一成分的液相与一种固相的三相反应(偏晶反应)。图中的偏晶反应可表示为: $L_1(x_{L_1}) \rightarrow A + L_2(x_{L_2})$ 。此外, 图中较低温度下还具有共晶反应 $L_2(x_e) \rightarrow A + B$ 。

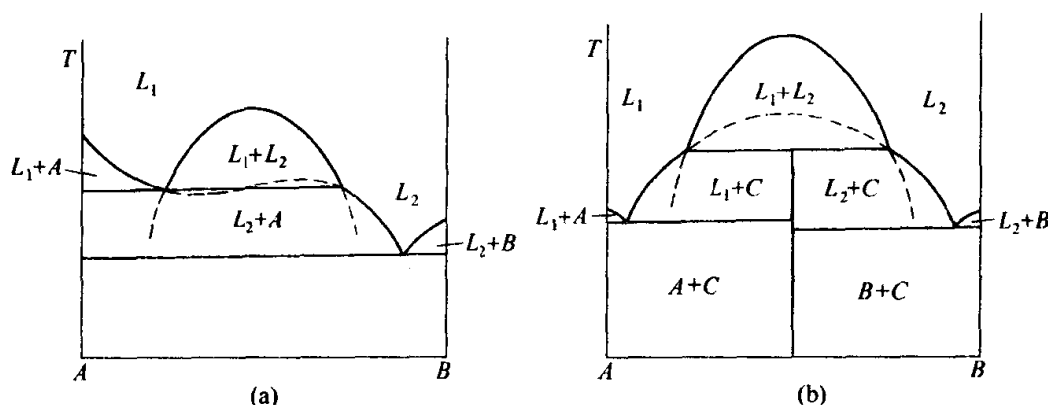


图 16-11 二元偏晶相图与综晶相图

(a) 偏晶相图; (b) 综晶相图

图 16-11 (b)所示的是在综晶相图中, 存在由两个不同成分的液相 (L_1 、 L_2) 转化一个中间化合物 C 的三相反应(综晶反应)。图中的综晶反应可表示为: $L_1(x_{L_1}) + L_2(x_{L_2}) \rightarrow C$, 此外在该相图中还存在另外两个共晶反应, $L_1(x_{e_1}) \rightarrow A + C$, $L_2(x_{e_2}) \rightarrow B + C$ 。

有些相图中, 液相存在亚稳的溶解度间隙, 这时相图对应于溶解度间隙顶点处的液相线一般比较平。

相图中通常还存在一些完全由固相参加的反应——固态反应, 与有液相参加的共晶反应与包晶反应相对应的分别是共析反应(如 $\alpha \rightarrow \gamma + \beta$)与包析反应($\theta + \delta \rightarrow \epsilon$)。当然, 相图中也有可能包含气相。但是无论怎么变化, 二元相图都可参照以上四类基本形式去分析。至于磁性转变、有序无序转变、亚稳分解等可参考相变有关章节。

实际二元相图的形式取决于构成二元系的两组元的性质相近程度。两组元性质愈近,则相图愈简单,如Co与Ni在元素周期表内相邻,元素性质极为相近,在固态能形成连续固溶体,因此,由这两元素组构成的二元相图为匀晶相图。又如Al-Si形成简单的共晶相图,Ag-Pt形成简单的包晶相图。两组元性质相差愈远,则相图愈复杂,通常会有中间化合物,如Fe-O相图。相图是各种可能存在的相在实验条件下热力学上相互竞争的结果,相图相当于一个刚结束战斗的战场,残存下来的只是为数不多的平衡相,实际体系中可能存在大量的亚稳相。近年来,利用机械合金化等亚稳材料制备技术在金属体系中制备出了非晶与亚稳的中间化合物。

16.5 三元相图

16.5.1 三元相图的表示法

(1) 立体图 系统组元数为三的相图为三元相图 (ternary phase diagram)。根据相律,对恒压凝聚态体系, $F = 3 + 1 - P = 4 - P$, 最大自由度为3, 相平衡关系需以3个独立的变量 (温度和任意两组元的浓度) 来表示。

三元系的成分由任意两组元的浓度来确定, 通常以巴基乌斯-洛兹本 (Bakhius-Roozeboom) 等边三角形来表示^[2]。如图16-12所示, 三角形的顶点代表纯组元, 边上的点代表边界二元系的成分, 三元系的成分由三角形内的点来确定。该浓度三角形具有如下性质: 平行于三角形某一边的直线上的任意一点都含有等量的对面顶点的组元; 三角形一顶点与对边上任意点的连线上的所有点所代表的体系中, 其他两组元的含量比值不变。

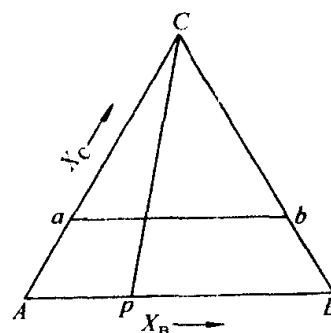


图 16-12 巴基乌斯-洛兹本浓度三角形

在水平的浓度三角形的端点加上垂直于该平面的温度坐标轴, 从而得到棱柱形表示三元系相关系的温度-成分空间立体图。其垂直表面反应边界二元系的相平衡关系, 即为二元相图, 棱柱中的每一点表示三元系中存在的一种平衡状态。由相律可知, 随着平衡共存相数目的增加, 凝聚态三元系相平衡的自由度由3减至0, 反映在相图中, 则具有不同的几何特征。

① 单相平衡。自由度为3。单相平衡点的集合构成任意形状的体积。

② 两相平衡。自由度为2。两相平衡点的集合构成两空间曲面。表面上的每一点可以由一根水平直线 (结线 tie-line) 和另一曲面上的共轭点相连, 结线通过的空间区域为两相区, 结线两端点所代表的成分即两平衡相的成分。

③ 三相平衡。自由度为1。三相平衡点的集合, 构成三条空间曲线。曲

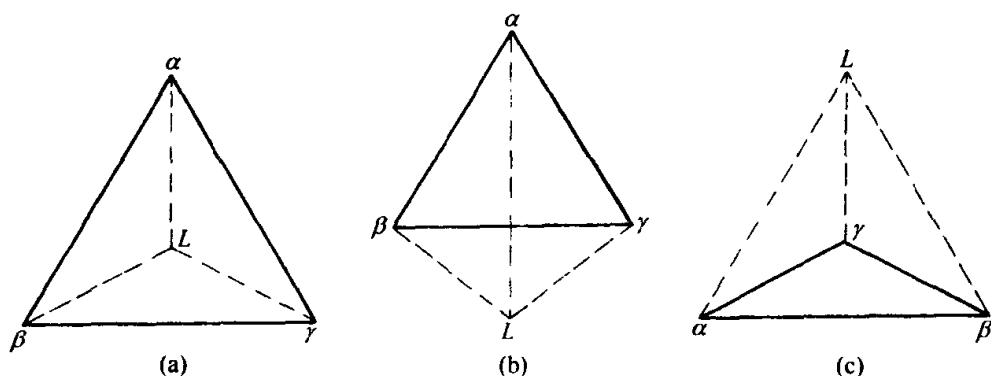


图 16-13 三元系中的四相区

(a) 含共晶转变的四相区；(b) 含包共晶转变的四相区；(c) 含包晶转变的四相区

线上的每一点，由一水平三角形（结三角形）与其他两条曲线上的共轭点相连，所有温度下的结三角构成的空间曲面三棱柱区域为三相区，结三角的端点所代表的成分就是平衡相的成分。

④ 四相平衡。自由度为零。四相平衡点是恒温面上的 4 个成分固定的点。该四点组合成的 4 个结三角所限制的平面区域为四相平衡区，它在 4 个

固定点分别与 4 个单相区相接触，在四个结三角形的六条边上与六个两相区相接触，在四个结三角形处与四个三相区相接触。其形状为三角形或四边形，对应于三角形区域的四相平衡反应为共晶反应 (eutectic reaction) ($L \rightarrow \alpha + \beta + \gamma$) 或包晶反应 (peritectic reaction) ($L + \alpha + \beta \rightarrow \gamma$)，对应于四边形区域的四相平衡反应界于前两者之间的一种反应 ($L + \alpha \rightarrow \beta + \gamma$)，有人称包共晶反应 (monotectic reaction)。如图 16-13 所示。

图 16-14 为三元共晶相图的立体图，图 16-15 为三元共晶相图立体图的分解图。如图所示，在三元共晶相图中包含有 L 、 α 、 β 、 γ 四个单相区， $L + \alpha$ 、 $L + \beta$ 、 $L + \gamma$ 三个固液两相区， $\alpha + \beta$ 、 $\beta +$

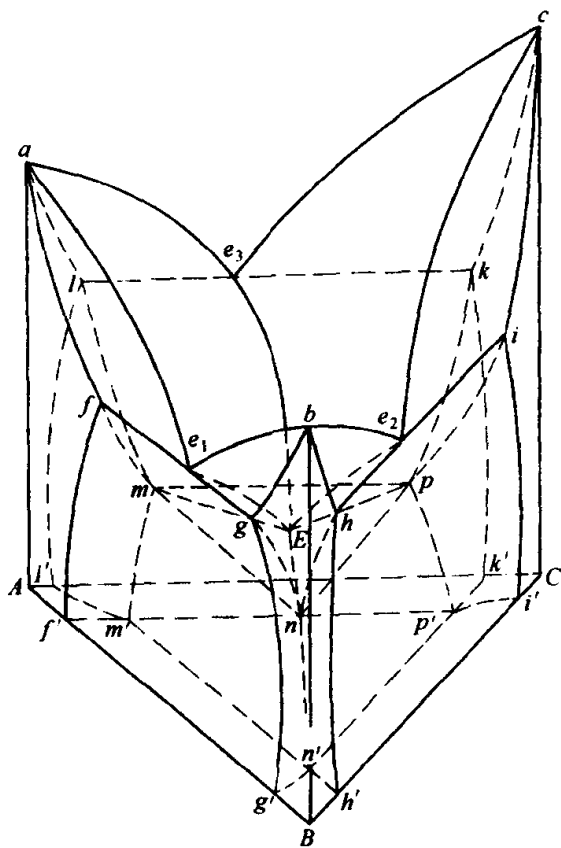


图 16-14 三元共晶相图

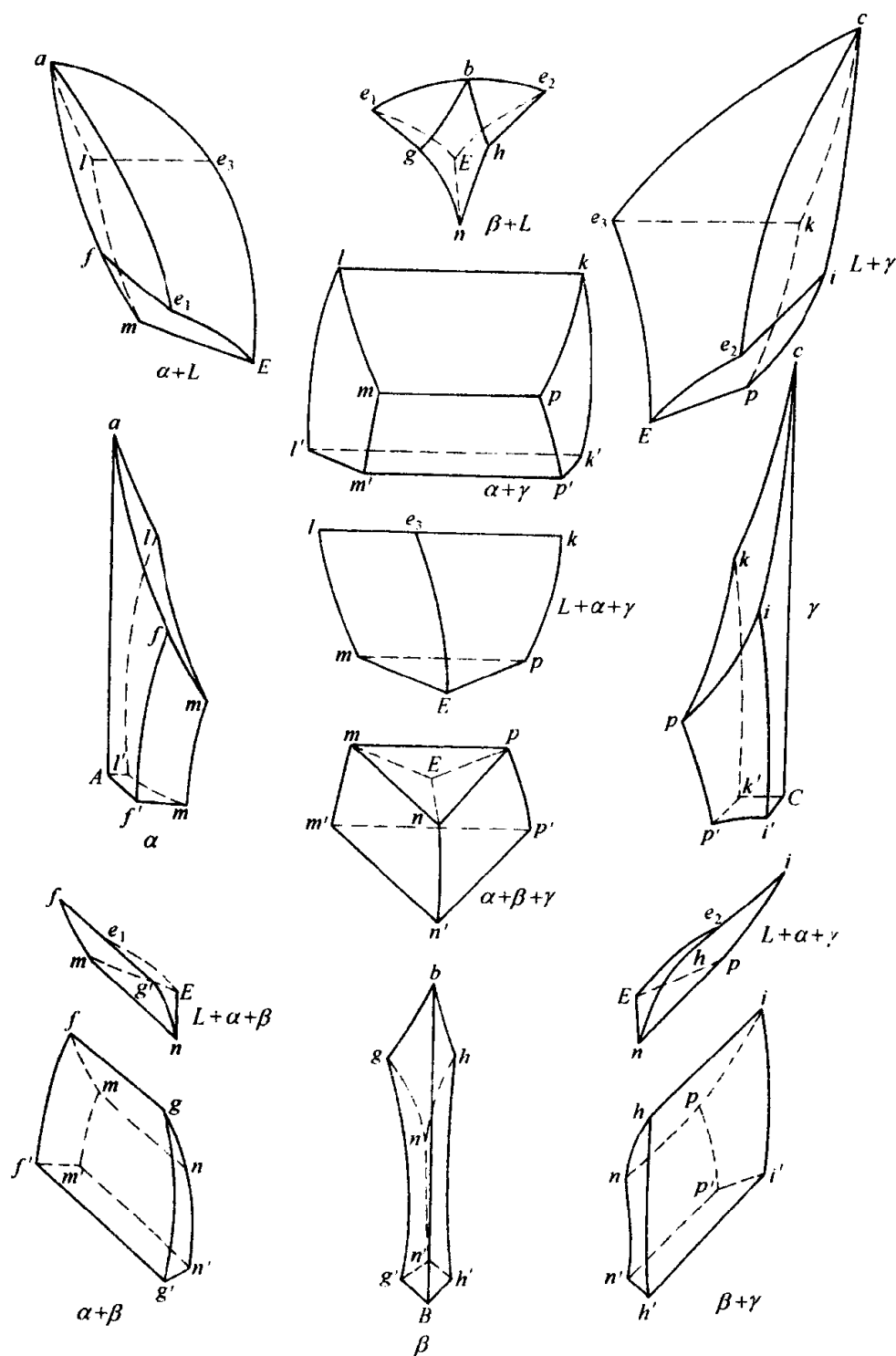


图 16-15 三元共晶相图分解图

γ 、 $\alpha + \gamma$ 三个固固两相区， $L + \alpha + \beta$ 、 $L + \beta + \gamma$ 、 $L + \alpha + \gamma$ 三个固液三相区， $\alpha + \beta + \gamma$ 一个固固三相区，和 $L + \alpha + \beta + \gamma$ 一个四相区。 E 为四相共晶点，在 E 点，液相恒温分解为 $\alpha + \beta + \gamma$ 三个固相。固相的成分即图示

m 、 n 、 p 三点的成分。

(2) 三元相图的截面、投影图 由于绘制与使用立体图非常困难, 使用亦不方便, 所以无论是在测量还是使用过程中, 人们常常用立体图沿水平和垂直方向的若干截面、液相面投影图来表示相平衡关系。

① 等温截面 (isothermal section)。一定温度下平行于浓度平面的截面。在等温截面上, 单相区为任意形状的曲边形, 反映该相的成分范围。两相区为两条直线与两条曲线构成的四边形, 两直线为两相区的两条结线, 两曲线为两平衡相的成分曲线。在两相区内根据两相的平衡成分我们还可以画一系列的结线。对结线上任一点我们都可以用杠杆定律来求平衡两相的成分。由于一个成分的材料在每一相中只有一个平衡成分, 因此, 结线都不会相交。三相区为结三角形, 结三角形的三个顶点即平衡三相成分, 三条边即两两平衡结线。对于结三角形我们也可以用杠杆定律的一种形式 (重心法则) 来求平衡各相含量。在等温截面上, $F = 3 - P$, 最多只有三相平衡, 但是通过四相反应前后等温截面的变化, 我们可以区分不同类型的四相反应。图 16-16 为图示共晶相图的几个等温截面, 在 e_3 温度以上只存在固液两相平衡, 分别在 e_3 、 e_2 、 e_1 处开始出现 $L + \alpha + \gamma$ 、 $L + \beta + \gamma$ 、 $L + \alpha + \beta$ 三个固液三相区, 在三相共晶温度处, 经过共晶反应固液三相区消失, 形成 $\alpha + \beta + \gamma$ 固固三相区。

② 垂直截面 (isopleth)。平行于温度坐标轴的截面。它反映系统中材料

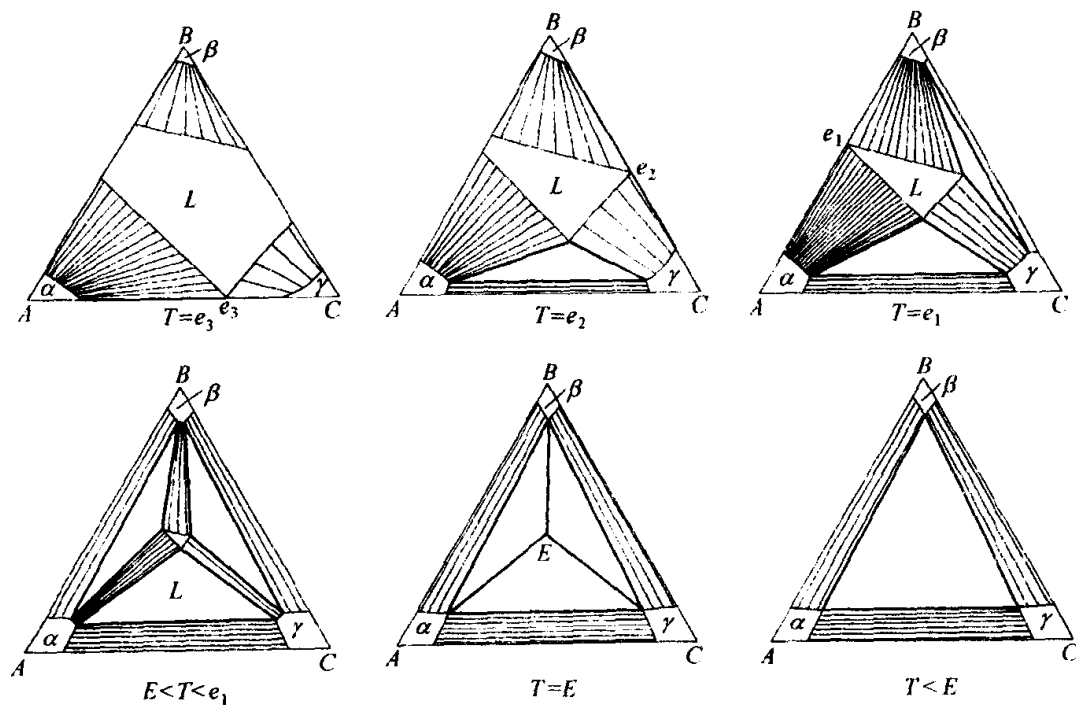


图 16-16 三元共晶相图的等温截面

在凝固与热加工过程中平衡共存相的种类。但一般来说, 因为平衡结线大都不在该截面上, 不能用来确定有关平衡共存相的成分。在垂直截面中, 单相区与两相区均是任意形状的曲边形, 三相区为曲边三角形, 四相区为一水平线, 四相区的类型不同, 截面的位置不同, 水平线上线下的相区分布不同。

③ 液相面投影图。所谓投影图 (projection), 是把空间图中的单变量线 (自由度为 1 的平衡的各相成分点的轨迹线) 或/和液固面的等温线, 沿温度轴方向投影到浓度三角形上的平面图。在投影图上, 所有的相区是相互重叠的, 投影图上的每一点不止代表一种相平衡状态, 从这个意义上讲, 投影图不是真正的相图。但是如果我们熟悉投影图的主要特征, 我们就可以由投影图画出等温截面、垂直截面进而确定整个三元相图的结构。

由三元液相面投影图画等温截面的步骤如下:

- ① 根据各无变度反应温度判断该温度下有哪些单变量线即三相区;
- ② 找出三相平衡液相与两固相成分, 画出共轭三角形;
- ③ 画出该等温截面与边际二元系的交点, 连接这些交点与三相区的顶点, 画出单相区与两相平衡, 即得等温截面。

由三元液相面投影图画垂直截面的步骤如下:

- ① 画出相应的浓度坐标与温度坐标;
- ② 画出垂直截面与固相两相区的交线, 如果固相没有溶解度, 则该交线为平行于温度轴的垂线;
- ③ 在投影图中找出四相反应成分范围, 根据四相反应温度画出四相反应水平线, 及其与相应两相区的交点;
- ④ 根据温度的变化画一系列共轭三角形, 根据这些三角形与截面的交线画出垂直截面的三相区;
- ⑤ 画出截面与二元相图中固相线与液相线的交点, 与单变量线的交点等特殊点, 连接这些点画出垂直截面的两相区与三相区, 即可。

图 16-17(a) 为三元共晶相图的液相面投影图, 与投影图中 1—9 线相对应的垂直截面。投影图中 1—9 线与液相单变量线四相反应时三相区的边界线以及固相三相区边界线的交点在垂直截面中都可以找到对应的点。垂直截面上温度轴上的点由二元相图液相线与共晶反应温度决定。投影图中由三条液相单变量线所划分出来的区域, Ae_1Ee_3 、 Be_1Ee_2 与 Ce_2Ee_3 称为 α 、 β 、 γ 的初晶区, 它是直接从液相析出 α 、 β 、 γ 固相的成分范围。在三元液相面投影图中, 每一个与液相平衡的固相都有自己的初晶区, 同分熔融的固相的初晶区在其成分点的周围, 异分熔融合化合物的初晶区在其成分点以外, 找出初晶区对晶体生长很重要。

根据投影图液相单变量线与相应两固相成分点的连线的关系可以判断固

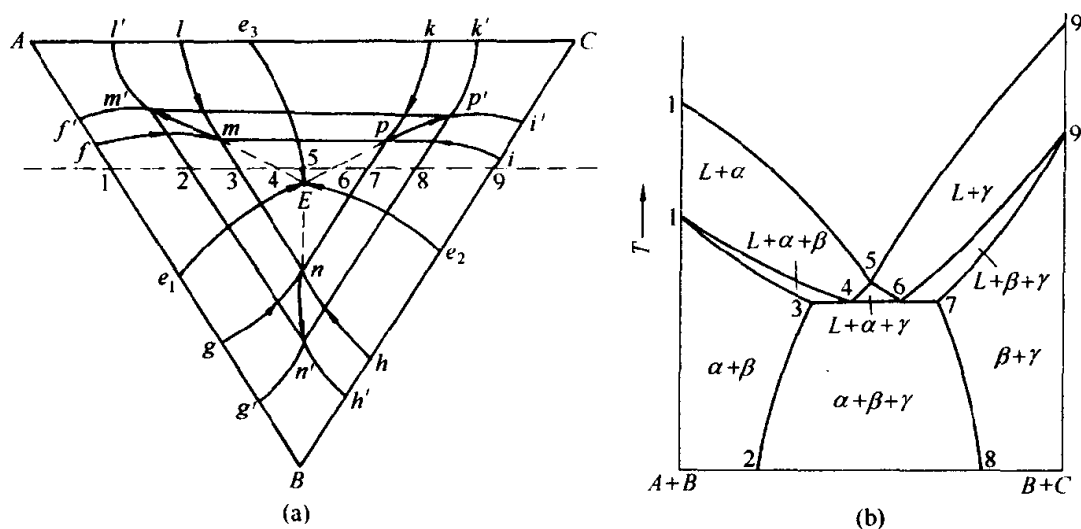
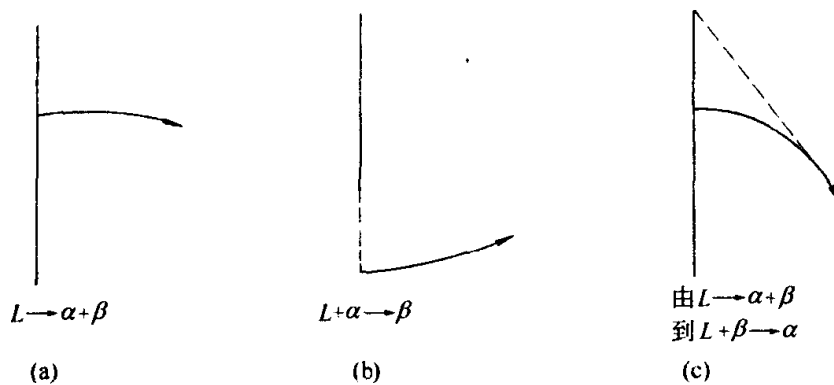


图 16-17 三元共晶相图的投影图与垂直截面

(a) 投影图; (b) 垂直截面

液三相反应的类型。一般情况下，如果单变量线与两固相成分点的连线相交，则为共晶反应；如果单变量线与两固相成分点的连线的延长线相交，则为包晶反应见图16-18(a)、(b)。三元相图中三相区的类型随温度的降低可能发生改变，在某一瞬间三相区是共晶还是包晶可由切线规则确定，如果单变量线的切线与两固相成分点连线相交则为共晶反应，如果单变量线的切线与两固相成分点连线的延长线相交则为包晶反应[图 16-18(c)]。

图 16-18 投影图液相单变量线与
相应两固相成分点的连线的关系^[23]

单变量线只能背向两固相成分点连线不可能指向两固相成分点连线，这就是阿尔克曼德 (Alkemade) 原理^[31]。该原理的完整表述为：在三元系中，如果毗邻两固相初晶区的分界线（或其延长线）与该两固相相组成点的连线（或连线的延长线）相交，则此交点即界限上的最高温度点，并且温度向两

侧下降(图 16-19)。该原理可以用于测量液相面投影图也可以用来分析复杂的三元液相面投影图。

单变量线与成分三角形边界的交点即二元系的三相平衡点。三条单变量线的交点即四相反应的液相成分点，每一个交点对应一个液相面以下的固相三相区与一个四相反应，相交单变量线的走向与对应的关系见图 16-20。

三条单变量线的交点即四相反应的液相成分点，由图 16-20，四相反应的类型也可以由单变量线交点与对应固相三相区的位置关系来确定。由此可见，相图中三元系的主要特征可以用液相面投影图来表示。

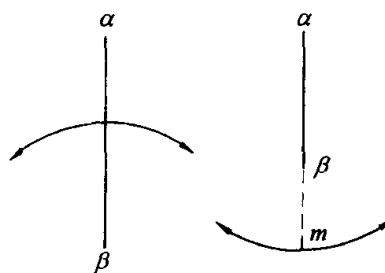


图 16-19 阿尔克曼德原理示意图

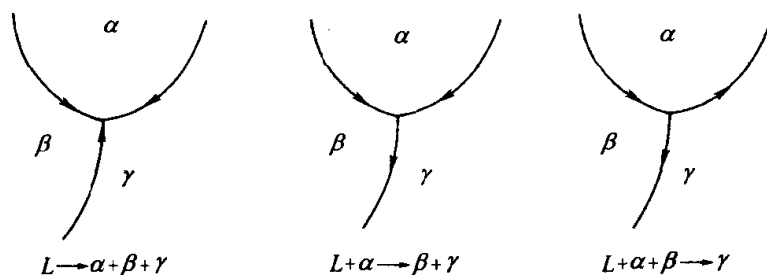


图 16-20 单变量线的走向与四相反应类型的关系

16.5.2 三元相图的分类

为方便讨论可以把三元相图分为以下几种类型。

(1) 最多只有两相平衡的三元相图 在多元系统中，最具代表性的最多只有两相平衡的相图为三元匀晶相图。如图 16-21 所示， ABC 为浓度三角形，三根垂线是温度轴，组成一个三棱柱立体相图。三元匀晶相图存在一个液相区(L)、一个固相区(α) 和一个固液两相区($L + \alpha$)。在液相区域内， A 、 B 、 C 在原子尺度内相互混合形成均一的液相；在固相区内， A 、 B 、 C 在原子尺度内相互混合形成均一的固溶体相；固液两相区内为固液两相混合物。液相区与固液两相区的交界面为液相面，它是两相平衡时的液相成分点的集合，固相区与固液两相区的交界面为固相面，它是两相平衡时固相成分点的集合。组元 A 、 B 、 C 的熔点分别位于三个温度轴上构成三元匀晶相图的三个顶点，液相面与固相面与三棱柱的三个侧面的交界线就是组元间形成的二元匀晶相图的液相线与固相线。所有成分的固溶体都存在一个凝固区间，从液相面开始，到固相面结束。当成分一定的固溶体随温度的升高或成分一定的液相随温度的降低进入两相区时，两相的平衡成分由连接液相面

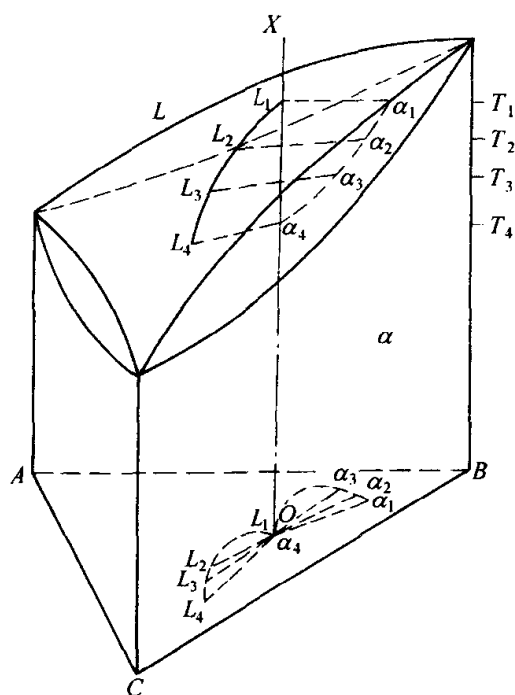


图 16-21 三元匀晶相图

与固相面的水平恒温线（结线）的端点来确定。例如：图 16-21 中成分为 X 的液相当温度降低到 T_1 以下时，将形成固液两相混合物，在 T_2 温度处，液相的成分点为 L_2 ，固相的成分点为 α_2 ，在 T_3 温度处，液相的成分点为 L_3 ，固相的成分点为 α_3 ，固液两相成分点的连线在浓度三角形内的投影形成所谓蝶形轨迹。根据杠杆法则在 T_2 温度处有如下关系： $m_L : m_\alpha = O\alpha_2 : OL_2$ 。式中 m_L 和 m_α 分别为液相与固相的质量。在 T_4 温度处凝固完毕， $X_\alpha = X$ ， $m_\alpha = 1$ ，形成成分与液相成分完全一致的固溶体相。

三元匀晶相图的固相面与液相面也可以相切，形成具有极高点或极低点的相图，在最高点与最低点处固液两相成分相等，自由度为零。有的三元匀晶相图在固相面以下存在溶解度间隙。

（2）最多只有三相平衡的三元相图 图 16-22 为三相平衡共晶相图及其分解图，在该相图的三个边际二元系中， $A-C$ 二元系为匀晶相图， $A-B$ 与

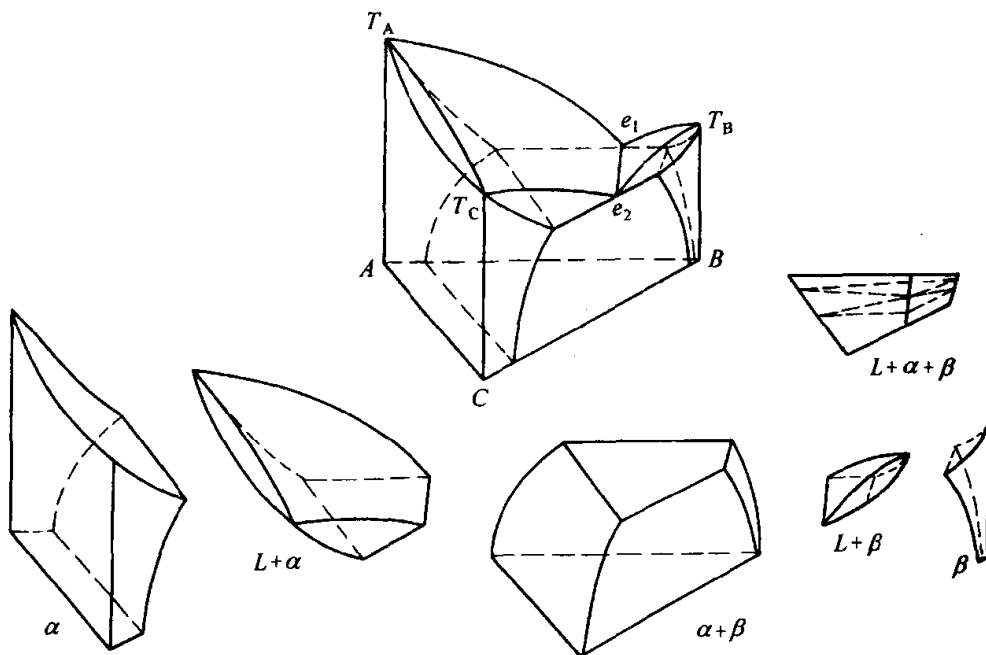


图 16-22 三相平衡共晶相图及其分解图

B-C 二元系为共晶相图，两共晶反应均为： $L \rightarrow \alpha + \beta$ ，在二元边界上，该三相平衡是一条直线，共晶转变在恒定的温度与压力下进行。进入三元系后，上述共晶反应，自由度为 1，三相平衡点的集合，构成三条空间单变量线。在任一温度下，曲线上的每一点，由一水平三角形（结三角形）与其他两条曲线上的共轭点相连，所有温度下的结三角构成一个空间曲面三棱柱区域。随温度降低，单变量线由 A-B 二元系指向 B-C 二元系。

三相平衡共晶相图中合金的凝固过程可以用图 16-23 所示的立体图和投影图来分析。成分位于三相区成分范围内的合金 O 的凝固过程为：在液相面以上，随温度降低液相冷却，到达液相面温度后，开始析出 α 相，液相的成分沿液相面变化，固相的成分沿固相面变化，直至液相成分与固相成分分别落在固液单变量线上，合金成分点落在三相共晶区的 $L + \alpha$ 面上，这时液相开始同时析出 α 与 β 两相，液相成分沿液相单变量线变化， α 与 β 的成分沿各自的单变量线变化，随凝固过程的进一步进行，合金成分点最后到达三相共晶区的 $\alpha + \beta$ 面，凝固过程结束。整个凝固过程中，液相与固相的成分变化如图中箭头所示。

三元系三相平衡相图除共晶相图以外，其余几类的投影图见图 16-24。其中图 16-24(a) 为具有包晶转变的三相平衡相图，A-B、B-C 两个二元系均为包晶相图，三元系三相平衡液相单变量线从 A-B 的包晶点 p_1 指向 B-C 的包晶点 p_2 。

图 16-24(b) 中 A-B 二元系为包晶相图，B-C 二元系为共晶相图，三元系三相平衡液相单变量线从 A-B 的包晶点 p_1 指向 B-C 的共晶点 e_2 ，在凝固过程中三相反应的性质将发生变化。图 16-24(c) 固相溶解度间隙终止于三元系内的三相平衡相图，其中 A-B 二元系为两端际固溶体结构相同的二元共晶相图，A-C、B-C 为匀晶相图，液相单变量线也终止于三元系内。图 16-24(d) 为具有固相岛状溶解的三元相图，三元系三相平衡液相单变量线的始点与终点均在三元系内，三个边际二元系均为匀晶相图。

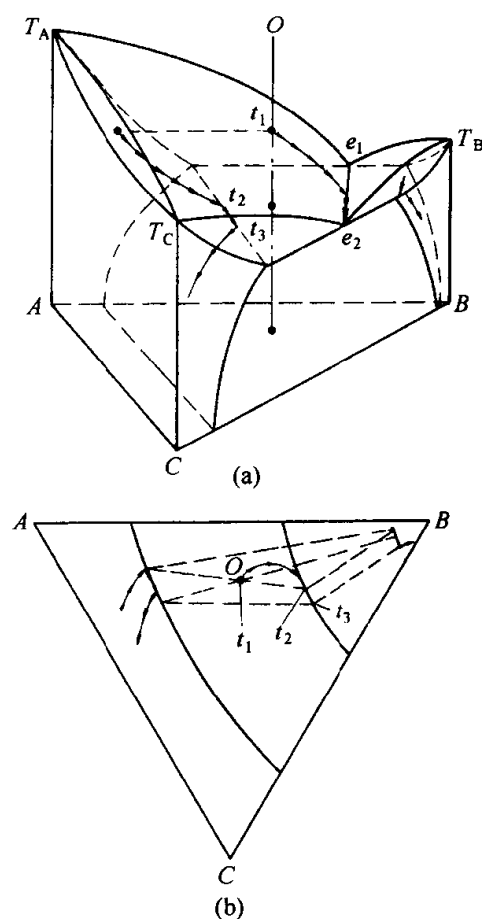


图 16-23 三相平衡共晶相图中合金的凝固过程
(a) 立体图；(b) 投影图

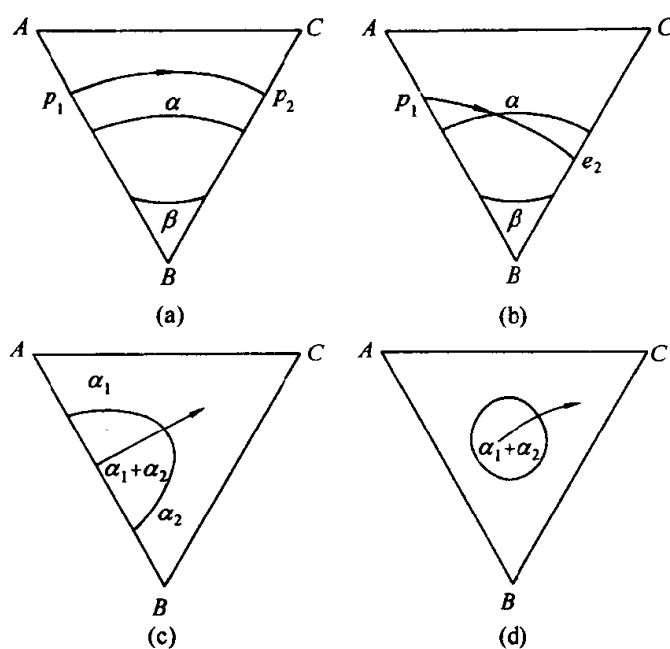


图 16-24 其他类型的三相平衡共晶相图

共晶转变。根据相律，由于第三组元的加入，上述共晶转变就不在恒定成分与温度下进行，而是在一定温度范围内（从二元共晶温度到三元共晶温度）随温度降低而连续的进行，其液相成分对应温度的变化作相应的改变，形成三个由一系列水平三角形构成的曲面三棱柱。三棱柱的一面为固固两相的平衡线的轨迹，在本相图中与二元相图的固相区重合，而另外两面为固液两相平衡线的轨迹，将三相区与固液两相区分开。这样每一个二元三相共晶点就延伸为一条三相共晶转线，其中成分位于 e_1E 线上的液相在冷却过程中将析出 $A+B$ 二元共晶，成分位于 e_2E 线上的液相在冷却过程中将析出 $B+C$ 二元共晶，成分位于 e_3E 线上的液相在冷却过程中将析出 $C+A$ 二元共晶。这三条线将交于一点 E ，此即四相共晶点，表示成分为 E 的液相在 T_E 温度下将发生 $L \rightarrow A+B+C$ 四相共晶转变，生成 $A+B+C$ 三相共晶。四相反应自由度为零，参加

(3) 有四相平衡的三元相图 与四相对应，三元系四相平衡相图也有三类：三元共晶相图、包共晶相图与包晶相图。

三元共晶相图的立体图、等温截面、垂直截面与投影图已在前面进行了说明。这里再讲一个特例，即组元在液态完全互溶，固态完全不溶的情况下的三元共晶相图，见图 16-25，相图的三个侧面是三个固态完全不溶二元共晶相图，分别具有 $Le_1 \rightarrow A+B$ ， $Le_2 \rightarrow B+C$ ， $Le_3 \rightarrow A+C$ 三个不同的

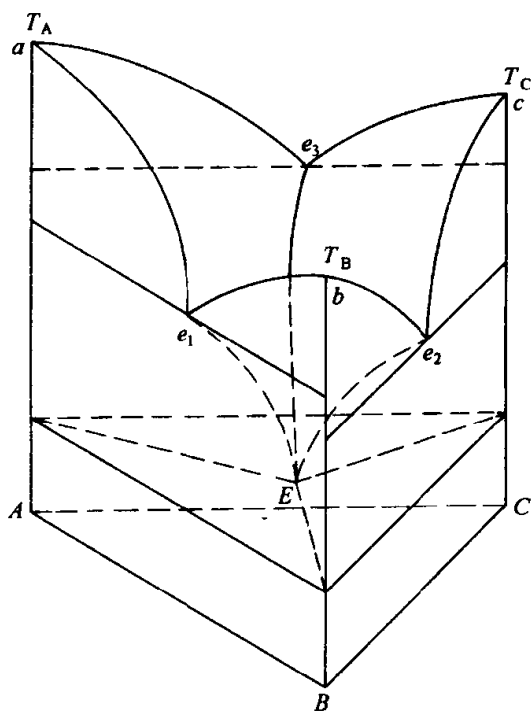


图 16-25 液态完全互溶固态完全不溶的情况下的三元共晶相图

反应的相的温度与成分都不再改变，因此四相共晶区为水平面。由于在这一平面上所有的液相均转变为固相，该水平面又是三元共晶相图的固相面。 ae_1Ee_3a ， be_1Ee_2a ， ce_2Ee_3c 为该相图的液相面。位于液相面以下三相区三棱柱固液两相平衡曲面以上的部分即三元相图的固液两相区，三相共晶线既是三相共晶区液相的成分线(单变量线)，又是固液两相区的分界线，即初晶区的分界线。

图 16-26 为该相图的投影图，根据该图我们可以分析成分为 o 的合金的凝固过程，当温度达到液相面温度后，液相中开始析出 A 的初晶，根据质量守恒定律，液相的成分将沿 Ao 连线的延长线变化，达到 q 点后，将从液相中析出 $A + C$ 两相共晶，液相的成分沿两相共晶线变化，最后到达 E 点，发生四相共晶反应析出 $A + B + C$ 三相共晶。整个凝固组织中初晶的含量为 qo/Aq ，二相共晶的含量为 $(Ao/Aq)(qE/pE)$ ，三相共晶的含量为 $(Ao/Aq)(pq/pE)$ 。

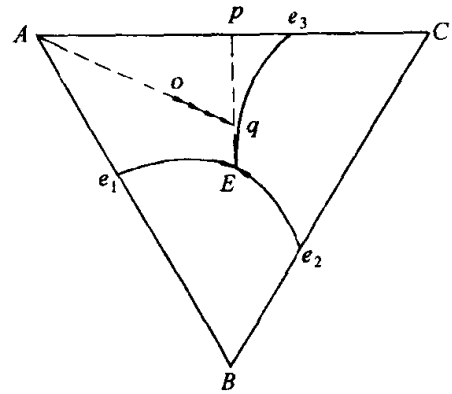


图 16-26 液态完全互溶固态完全不溶三元共晶相图的投影图

三元包晶与包共晶反应的液相面投影图如图 16-27。如果不考虑三元系中三相区的类型随温度降低有可能发生变化，那么四相共晶反应的三个边际二元系都为二元共晶相图，四相包晶反应与四相包共晶反应所对应的三个边际二元系都为包晶与两个包晶，只是包晶反应的温度高于共晶反应的温度，而包共晶反应共晶反应的温度高于包晶反应的温度。如果考虑三元系中三相区的类型随温度降低的变化，则三种四相反应所对应的边际二元系不受

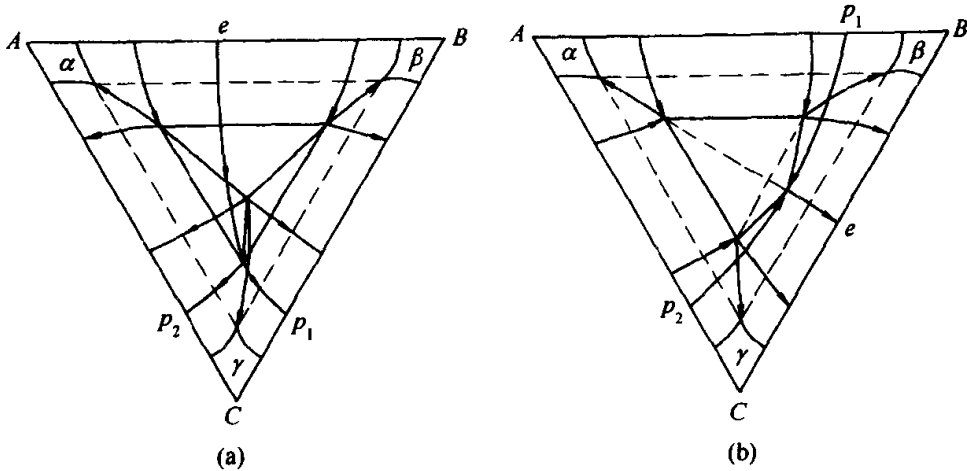


图 16-27 三元包晶相图与包共晶相图的液相面投影图

(a) 三元包晶相图；(b) 三元包共晶相图

上述限制。

(4) 有化合物的三元相图 在实际三元相图中经常出现若干二元化合物与三元化合物, 这些化合物将浓度三角形划分为若干个三角形。如果某一温

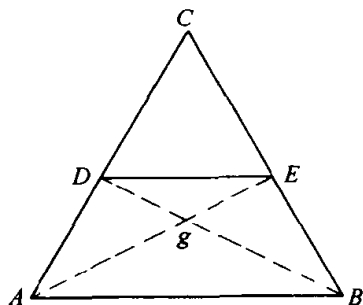


图 16-28 由实验方法确定子三角形^[32]

度下二元化合物的数量为 b , 三元化合物的数量为 t , 那么由这些化合物所划分出的子三角形数量为 $2t + b + 1$ 个。子三角形的划分方法可以用实验加以确定。如图 16-28, A 和 C 形成二元化合物 D, B 和 C 形成二元化合物 E, 连接 AE 或 BD 均可以构成子三角形, 这时我们通常取 AE 与 BD 的交点 g 的成分的配样品, 加热熔化并在给定温度下均匀化退火或在给定温度下烧结, 检查样品中所存在的相, 如果是 A 和 E,

则连接 AE 构成 $\triangle ADE$ 与 $\triangle AEB$ 两个子三角形是正确的, 反之如果析出的是 B 和 D 则连接 $\triangle BDE$ 与 $\triangle ADB$ 两个子三角形是正确的。

在三元液相面投影图中, 同分熔化化合物的成分点在初晶区之内, 异分熔化化合物的成分点在初晶区之外。每一个固液四相反应都对应于一个固相面上的子三角形, 构成子三角形三个顶点的三个相即固液四相反应的三个固相。当固液四相反应为共晶反应时, 液相成分点的位置在子三角形内, 当固液四相反应为包晶或包共晶反应时, 液相成分点的位置在子三角形外。图 16-29(a)~(d) 分别为带一个同分熔化二元化合物、一个异分熔化二元化合物、一个同分熔化三元化合物或一个异分熔化三元化合物的相图的液相面投影图。可以根据前面所讲的方法分析材料的凝固过程, 材料的总成分点落在哪个固相三相区内, 材料最终就由哪三相组成。

(5) 有溶解度间隙的三元相图 图 16-30(a) 为具有独立的液相溶解度间隙的三元相图投影图与相应的边际二元相图。如图所示, 三元系的溶解度间隙是由二元偏晶相图的液相溶解度间隙扩展而来的。溶解度间隙的边界线, 实际上是 $L_1 + L_2 + A$ 三相区的单变量线, 反映三相平衡时液相的成分。当这个液相溶解度间隙扩展到足够大时, 其边界线可能与其他单变量线相交, 形成带四相偏晶反应的相图[图 16-30(b)]。图 16-30(c) 为图 16-30(b) 所示相图沿 $a-a'$ 方向的垂直截面。成分位于液相溶解度间隙内 o 点的材料凝固时先发生液相分层, 然后进入 $L_1 + L_2 + A$ 三相区, 从 L_1 中转变为 $L_2 + A$, L_1 反应完毕, 进入 $L_2 + A$ 两相区, 开始从 L_2 中析出 A, 随 A 的析出液相中 B 的浓度上升, 最后进入 $L_2 + A + B$ 三相区, 开始析出 A + B 二元共晶体, 到达 T_E 温度后, 剩余液相发生四相共晶反应 $L_2 \rightarrow A + B + C$ 全部转化为 A + B + C。

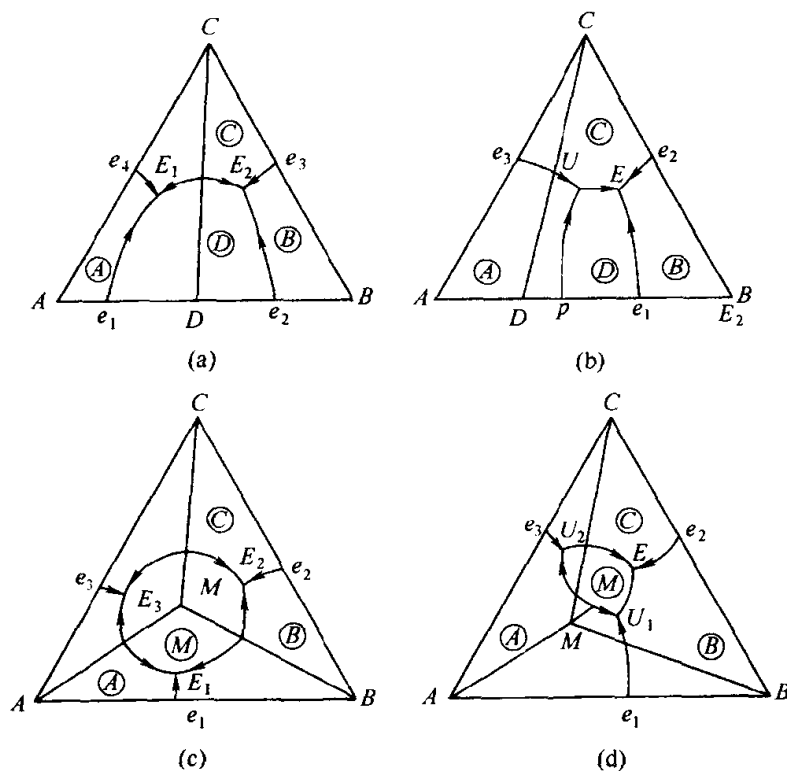


图 16-29 带有化合物的三元相图的投影图

(圆圈内的字母表示初晶区)

- (a) 带一个同分熔化二元化合物；(b) 带一个异分熔化二元化合物；
(c) 带一个同分熔化三元化合物；(d) 带一个异分熔化三元化合物

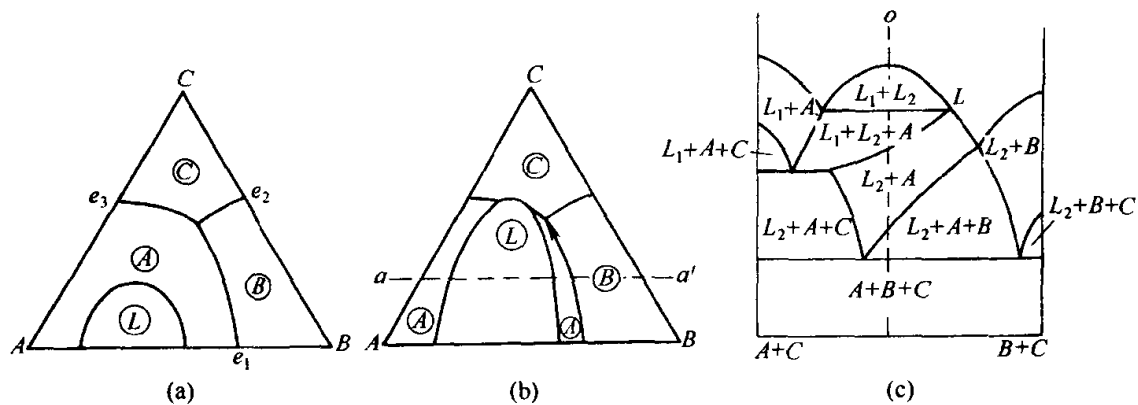


图 16-30 带液相溶解度间隙的三元相图

- (a) 带独立的液相溶解度间隙的相图；(b) 带四相偏晶反应的相图；(c) 沿 $a-a'$ 的垂直截面

参 考 文 献

- 1 Gibbs W. The Collected Works. Vol. 1, Yale University Press, 1957
- 2 Bakhuis Roozeboom, H. W. Die Heterogenen Gleichgewichte vom Standpunkte der Phasenlehre. Braun-

- sxhweig Vol. 1 ~ Vol. 3 1901 ~ 1913
- 3 Tammann G.Z. Anorg. Chem. 1903, 37: 303; 1905, 45: 24; 1905, 47: 287
- 4 Hansen. M. Constitution of Binary Alloys. 2nd ed. New York: McGrawHill Inc, 1966
- 5 Massalski. T. B. Handbook of Binary Alloy Phase Diagrams. ASM International, Materials Park, OH, 1990
- 6 С. Л. Алисова, Л. Ё. Вульф, К. М. Маркович, Л. К. Новик, Л. А. Летрова, Э. М. Рогачевская, Диаграммы Состояний Металлических Систем 1955 ~ , Москва, (1959 ~)
- 7 Pezow, G. Enfferberg. G. Ternary Alloys. Vol. 1 ~ Vol. 8, VCH Verlagsgesellschaft, Weinheim, FGR, 1988 ~ 1993
- 8 何纯孝. 贵金属合金相图. 北京: 冶金工业出版社, 1983
- 9 何纯孝, 周月华, 王文娜. 贵金属合金相图. 第一补篇. 北京: 冶金工业出版社, 1993
- 10 Levin E. M. et al. Phase Diagram for Ceramists. Vol. 1 ~ Vol. 5. Amer. Ceram. Soc. Columbus, Ohio (1964 ~)
- 11 徐祖耀. 金属材料热力学. 北京: 科学出版社, 1981
- 12 唐有祺. 相平衡、化学平衡和热力学. 北京: 科学出版社, 1984
- 13 张圣弼, 李道子. 相图——原理、计算及其在冶金中的应用. 北京: 冶金工业出版社, 1986
- 14 郭祝昆, 林祖襄、严东生. 高温相平衡与相图. 上海: 上海科学技术出版社, 1987
- 15 黄勇, 崔国文. 相图与相变. 北京: 清华大学出版社, 1987
- 16 陆学善. 相图与相变. 合肥: 中国科学技术大学出版社, 1990
- 17 赵慕愚. 相律应用及其进展——相图边界理论. 长春: 吉林科学技术出版社, 1988
- 18 顾菡珍, 叶于蒲. 相平衡与相图基础. 北京: 北京大学出版社, 1991
- 19 梁敬魁. 相图与相结构. 北京: 科学出版社, 1993
- 20 胡德林, 张帆. 三元合金相图. 西安: 西北工业大学出版社, 1995
- 21 Schreinmakers F. A. H., Die Ternaren Gleichgewichte. Braunschweig, 1911
- 22 Findlay Alexander. The Phase Rule and its Applications. 8th ed. London: Longman, 1938
- 23 Ricci, J. E. The Phase Rule and Heterogeneous Equilibrium. New York, London D. Van Nostrand Company Inc. 1951
- 24 Bowden S. T. The Phase Rule and Phase Reactions—The Theoretical and Practical 4th ed. . St. Martin's Press Inc, 1954
- 25 Rhines, F. N. Phase Diagram in Metallurgy; their Development and Application New York: McGraw-Hill Book, 1956
- 26 Platinic, L. S. A. I. Landau, Phase Equilibrium in Multi-component System, Holt, Rinehart and Winston, 1964
- 27 Alperedt. A. M. Phase Diagram—Material Science and Technology, vol 1-vol 6 New York: Academy Press, 1970
- 28 Lupis, C. H. P. Chemical Thermodynamics of Materials. Amsterdam: Elsevier Publishing Co. Inc., 1983
- 29 Oonk, H. A. J. Phase Theory—Thermodynamic of Heterogeneous Equilibria, Amsterdam: Elsevier Publishing Co. Inc., 1980; [荷] H. A. J. 翁克. 施印华译. 黄述诚校. 相理论——多相平衡热力学, 北京: 科学出版社, 1987

- 30 Hillert, Mats Phase Equilibrium Phase Diagrams and Phase Transformation—Their Thermodynamic Basis. Cambridge University Press, 1998
- 31 Kononov, D. Ann. Phys. 1881, 14: 34
- 32 Van Ryn Alkemade, A. C. Z. Phys. Chem. 1983, 11: 289

第 17 章 固体中的扩散、化学反应与烧结

撰稿人 曾燕伟

17.1 固体中的扩散

正如前面所述，晶体材料的主要结构特征是其原子或离子的周期性规则排列。然而，实际晶体中原子或离子的排列总会或多或少地偏离这种严格的周期性。由于热起伏的存在，晶体中的部分原子或离子由于剧烈的热振动而脱离正常格点进入附近的间隙位置或晶体的表面，而且这些脱离正常格点的原子或离子可以从热起伏的过程中重新获得能量，在晶体结构中不断地改变位置而出现由一处向另一处的无规则迁移运动，这就是晶体中原子或离子的扩散。原子或离子的这种扩散迁移运动不仅可以出现在晶体材料中，同样可以发生在结构无序的非晶态材料中（如玻璃、高分子有机材料等）。

原子或离子的扩散过程是一种不可逆过程。它与热传导、导电、粘滞等不可逆过程一样，都是由于物质内部存在某些物性的不均匀性而发生的输运过程。如物体中温度梯度引起的热量输运——热传导现象；流体运动的速度差引起的动量迁移——内摩擦或粘滞现象；导体中电位差引起的电量输运——导电现象。扩散现象则是由于物质中存在浓度梯度或化学势梯度、温度梯度、势能梯度（如电场）所引起的质量输运过程。原子或离子的扩散是众多工程材料如金属材料、无机非金属材料、有机高分子等材料的制备、使用中很多重要的物理、化学以及物理化学过程得以实现的基础。例如半导体掺杂、固溶体的形成、离子晶体的导电、固相反应、相变、烧结、材料表面处理等等都与材料内部物质的迁移扩散有着密切的关系。因此研究并掌握扩散的基本规律对材料性质的认识与开发、材料的制备、生产和使用均有十分重要的意义。

17.1.1 扩散的基本特点与唯象理论

（1）扩散的基本特点 发生在气体或液体中的传质过程是一个早为人们所熟悉的现象。在流体中，质点间相互作用比较弱，且无一定结构，质点的迁移可以像图 17-1 中所描述的那样，完全随机地在三维空间的任意方向上发生，每一步的迁移行程也随机地决定于该方向上最邻近质点的距离。流体的密度越小，质点迁移的平均行程（也称自由程）越大。因此，发生在流体中的物质迁移过程往往总是各向同性和具有较大的速率。

与流体中的情况不同,质点在固体介质中的扩散远不如在流体中那样显著。首先构成固体的质点均束缚在三维结构的势阱中,质点之间相互作用强,故质点的每一步迁移必须从热涨落或外场中获取足够的能量以跃出势阱。实验表明,固体中质点的明显扩散往往在低于其熔点或软化点的较高温度下发生。此外,固体中原子或离子的扩散迁移方向和自由程还受到结构

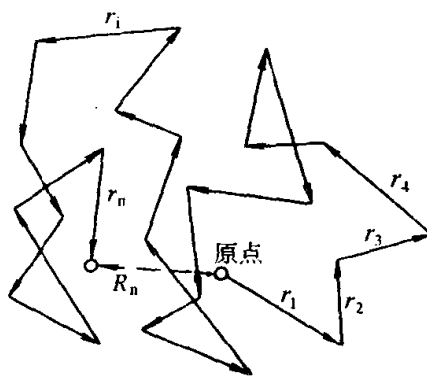


图 17-1 扩散质点的无规行走轨迹

中质点排列方式的限制。如图 17-2 所示,处于平面点阵内间隙位的原子只

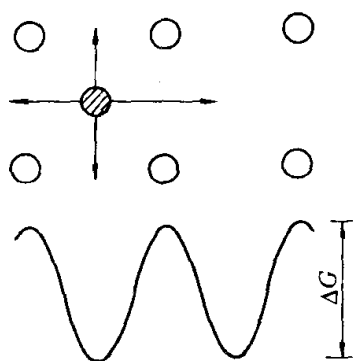


图 17-2 平面点阵中间隙原子扩散方向与势场结构示意图

存在四个等价的迁移方向,每一步迁移均需获取高于能垒 ΔG 的能量,迁移的自由程则相当于晶格常数大小。因此,固体中的扩散具有各向异性和扩散速率低的特点。

(2) 菲克定律与扩散方程 流体和固体中质点的迁移行为在微观上彼此存在较大的差异,但从宏观连续介质的角度看,它们则遵守着相同的统计规律。德国物理学家菲克(Fick)于 1855 年用类似傅里叶(Fourier)研究热传导的方法首先得出了著名的菲克定律,建立了定量描述浓度场下物质扩散所遵循的数学方程。

菲克定律认为,在连续介质构成的扩散体系中扩散质的浓度 c 一般是空间 r 和时间 t 的函数,单位时间内扩散质流过某一垂直截面元的扩散流量密度 J 与该处扩散质的浓度梯度 dc/dr 成正比,即有如下扩散第一方程:

$$J = -D \nabla c = -D \left(i \frac{\partial c}{\partial x} + j \frac{\partial c}{\partial y} + k \frac{\partial c}{\partial z} \right) \quad (17-1)$$

式中, D 为扩散系数,其量纲为 $L^2 T^{-1}$,在 SI 和 CGS 单位制中分别为 m^2/s 和 cm^2/s ;负号则表示质点的扩散流与浓度梯度方向相反。

应该指出,扩散第一方程是一个矢量式,故对于各向同性的固体材料如金属、陶瓷等多晶材料,扩散系数 D 常为与方向无关的标量。但在一些存在各向异性的单晶材料中,扩散系数 D 的变化取决于晶体结构的对称性。对于非立方对称的各向异性材料,扩散系数 D 为二阶张量,此时式(17-1)有如下分量形式:

$$J_i = -D_{ii} \frac{\partial c}{\partial x} - D_{ij} \frac{\partial c}{\partial y} - D_{ik} \frac{\partial c}{\partial z} \quad (i, j, k = x, y, z) \quad (17-2)$$

扩散第一方程是质点扩散定量描述的基本方程。它可以直接用于求解浓度分布不随时间变化的稳定扩散问题，同时又是浓度分布随时间变化的不稳定扩散动力学方程建立的基础。现考察不稳定扩散体系中任一由界面 S 围成的体积 V ，在单位时间内流入体积 V 中净扩散质的量应为：

$$Q = - \oint_S J_s dS \quad (17-3)$$

根据高斯定理则有：

$$Q = - \int_V \text{div} J_s dV \quad (17-4)$$

又根据质量守恒定律，单位时间流入体积 V 中净物质量与该体积内扩散质浓度变化应满足如下关系：

$$Q = \int_V \frac{dc}{dt} dV \quad (17-5)$$

比较式 (17-4) 和式 (17-5)，容易导出如下扩散第二方程：

$$\frac{\partial c}{\partial t} = - \text{div} J_s = \text{div} (D \nabla c) \quad (17-6)$$

当扩散具有各向同性且扩散系数 D 不随位置变化，则扩散第二方程有如下直角坐标和球坐标系中的表达式：

$$\frac{\partial c}{\partial t} = D \left(\frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2} + \frac{\partial^2 c}{\partial z^2} \right) \quad (17-7)$$

$$\frac{\partial c}{\partial t} = D \left(\frac{\partial^2 c}{\partial r^2} + 2 \frac{\partial c}{r \partial r} \right) \quad (17-8)$$

(3) 扩散方程的解与应用 在一定的边界和初始条件下，扩散方程可求得其解析解或数值解。许多不同条件下扩散方程的解可从一些关于扩散和热传导的专著中找到，本节仅介绍扩散系数为常数情况下的几种简单解及其应用。

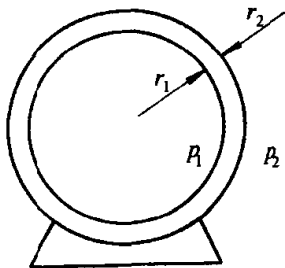


图 17-3 氧气分子
通过球罐壁扩散
泄漏示意图

① 稳定扩散。扩散第一方程可直接用于描述稳定扩散过程。现以一个高压氧气球罐的氧气泄漏问题作为例子。如图 17-3 所示，氧气储罐的内外径分别为 r_1 和 r_2 ，罐中气压为 p_1 ，罐外气压即为大气中的氧分压为 p_2 。由于氧气泄漏过程极为微弱缓慢，故认为 p_1 不随时间变化，达到稳态后，氧气将以一恒定速率泄漏。由扩散第一方程可知，氧气在钢罐内壁的扩散引起单位时间内的泄漏量为：

$$Q = -4\pi r^2 D \frac{dc}{dr} \quad (17-9)$$

对上式积分，便有：

$$Q = -4\pi D r_1 r_2 \frac{c_2 - c_1}{r_2 - r_1} \quad (17-10)$$

式中， c_1 和 c_2 分别为氧分子在球罐内、外壁的溶解度。根据 Sievert 定律，双原子分子气体在固体中的溶解度通常与压力的平方根成正比 $c = Kp^{1/2}$ ，于是可得单位时间内氧气泄漏量：

$$Q = -4\pi D r_1 r_2 K \frac{p_o^{1/2} - p_i^{1/2}}{r_2 - r_1} \quad (17-11)$$

② 不稳定扩散。不稳定扩散中典型的边界条件可分成两种情况，它们对应于不同扩散特征的体系。一种情况是扩散长度远小于扩散体系的尺度，故可引入无限大或半无限大边界条件而使方程得到简单的解析解；另一种情况是扩散长度与体系尺度相当，此时方程的解往往具有级数的形式。下面讨论前一种情况中的两个例子。

如图 17-4 所示的扩散体系为一长棒 A，其一端面暴露于扩散质 B 的恒压蒸气中，因而扩散质将由端面不断扩散至棒 A 的内部。不难理解，该扩散过程可由如下方程及其初始条件和边界条件得到描述：

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2}; \quad c(x > 0, 0) = 0; \quad c(0, t > 0) = c_0. \quad (17-12)$$

通过引入新的变量 $u = xt^{-1/2}$ ，并注意到在任意时刻 $c(\infty, t) = 0$ 和 $c(0, t) = c_0$ 的边界条件，容易解得长棒 A 中扩散质浓度分布为：

$$c(x, t) = c_0 \left[1 - \operatorname{erf} \left(\frac{x}{2\sqrt{Dt}} \right) \right] \quad (17-13)$$

式中 $\operatorname{erf}(z)$ 为高斯误差函数：

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \exp(-\xi^2) d\xi \quad (17-14)$$

其函数图线如图 17-4(b) 所示。

由式 (17-13) 容易看出，对于一定值的 $c(x, t)/c_0$ ，所对应的扩散深度 x 与时间 t 有着确定的关系。如假定 $c/c_0 = 0.5$ ，由图 17-5 可知 $x/2$

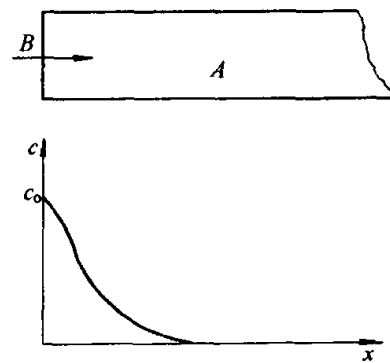


图 17-4 端面处于恒定蒸气压下的半无限大固体
一维扩散示意图

$(Dt)^{1/2} = 0.52$, 即在任何时刻 t , 对于半浓度的扩散距离 $x = 1.04(Dt)^{1/2}$, 并有更一般的关系:

$$x^2 = Kt \quad (17-15)$$

式中 K 为比例系数。这个关系式常称为抛物线时间定则。它被广泛地应用于如钢铁渗碳、晶体管或集成电路生产等工艺环节中控制扩散质浓度分布和扩散时间以及温度的关系。

长棒扩散的另一个典型例子是所谓的扩散薄膜解。如图 17-6 所示, 在一半无限长棒的一个端面上沉积 Q 量的扩散质薄膜, 此时扩散过程的初始和边界条件可描述为:

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2}; c(x > 0, 0) = 0; \int_0^{\infty} c(x) dx = Q (t > 0) \quad (17-16)$$

其相应的解有如下形式:

$$c(x, t) = \frac{Q}{\sqrt{\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right) \quad (17-17)$$

扩散薄膜解的一个重要应用是测定固体材料中有关的扩散系数。将一定量的放射性示踪原子涂于长棒的一个端面上, 测量经历一定时间后放射性示踪原子离端面不同深度处的浓度, 然后利用式(17-17)对所获实验数据作 $\ln c(x, t) \sim x^2$ 直线, 由其斜率和截距求得相应的扩散系数 D 。

对于扩散长度与扩散体系尺度相当的一些情况, 扩散第二方程的解往往借助于分离变量法, 即令 $c(x, t) = A(x)B(t)$, 进而得到用傅里叶级数形式表达的解:

$$c(x, t) = \sum_n A_n \sin(\lambda_n x) \exp[-\lambda_n^2 Dt] \quad (17-18)$$

可以看出, 此时扩散体系的浓度分布由一系列浓度波叠加而成, 它们的振幅随时间按指数关系衰减。

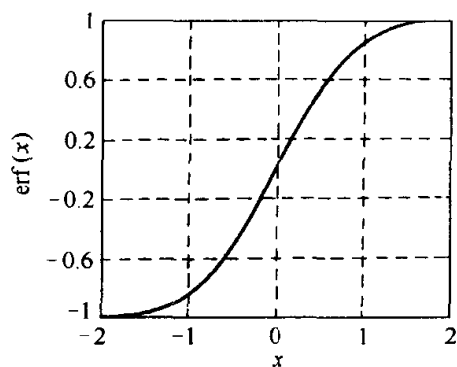


图 17-5 高斯误差函数曲线

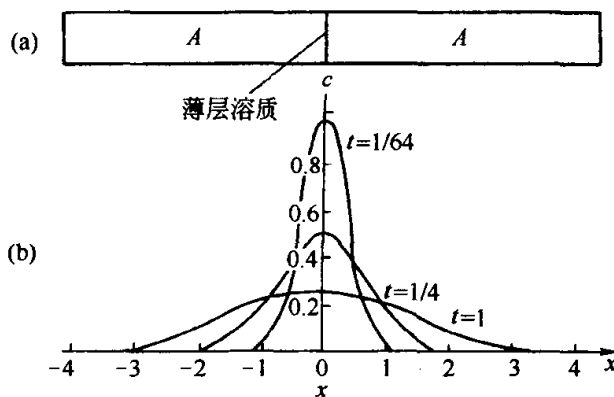


图 17-6 定量扩散质在无限长棒中扩散的薄膜解

(4) 扩散系数随浓度的变化 在上一节中, 为求解的简单起见, 扩散系数被看作为一个不随浓度变化的常数。然而, 实验表明, 大量的实际材料中扩散质浓度分布并非像数学解所描述的那样规则和理想, 即使在一些简单的合金系统中也是如此。例如在 Cu-Ni 合金系统中, 观测到试样中富 Cu 的一端扩散远比在富 Ni 的一端来得快。这种实验与理论的偏差实际上源于体系中扩散系数随体系成分变化而变化的事实。例如在 Si、Al、Sn、Zn、Cd 和 Be 溶于 Cu 的一次固溶体中, 当浓度从零增加到溶解极限时, 扩散系数可变化 10 倍之多。现以一对半无限大扩散偶为例, 用扩散第二方程讨论扩散系数随浓度变化的关系。显然, 此时扩散方程和对应的边界条件应为:

$$\frac{\partial c}{\partial t} = \frac{\partial}{\partial x} \left(D \frac{\partial c}{\partial x} \right); \quad c(x < 0, t = 0) = c_0; \quad c(x > 0, t = 0) = 0 \quad (17-19)$$

在上式中引入变换 $u = xt^{-1/2}$, 于是有:

$$-\frac{u}{2} dc = d \left[D \frac{\partial c}{\partial u} \right] \quad (17-20)$$

考虑 $[dc/dx]_{c=0} = 0$, 将上式作 $0 \sim \xi$ 的积分得:

$$-\frac{1}{2t} \int_0^\xi x dc = D \left(\frac{dc}{dx} \right)_\xi \quad (17-21)$$

因而有扩散系数 D 与浓度 ξ 的依赖关系:

$$D = -\frac{1}{2t} \left(\frac{dx}{dc} \right)_\xi \int_\xi^0 x dc \quad (17-22)$$

应该指出, 变换 $u = xt^{-1/2}$ 的引入使原偏微分方程变成了一个常微分方程。然而这一变换是有局限性的, 并非所有扩散偶的浓度分布均具有 $c = f(xt^{-1/2})$ 的形式。1933 年侯野 (Matano) 首先使用上述变换考察了扩散系数随浓度的变化。实验中利用上式考察扩散系数随浓度变化的一般方法是首先测出扩散质在 t 时刻的浓度分布曲线 $c(x)$, 然后如图 17-7 所示, 用作图法分别求出在浓度 ξ 点处的导数项与积分项, 最后依式(17-22) 得到扩散系数随浓度的变化关系。

(5) 互扩散中的克肯达尔效应 当扩散体系中存在两种或两种以上扩散质时, 扩散将完全不同于单一扩散质的情况。如图 17-8 所示, 将两种材料焊在一起而构成一个扩散偶, 并用高熔点的情性金属细丝埋在焊接界面处作为位置性标记。可以想像, 若两个组元的扩散速率不相同, 则退火后在焊接面处组元扩散较慢的那一边将积聚较多的物质, 出现体积膨胀; 而在焊接面的另一面, 则由于丧失了扩散较快的组元而出现空洞, 从而导致材料收缩。假如将试样扩散偶的一端固定, 则随扩散的进行, 焊接面标记物将出现如图 17-8(c) 所示的反扩散方向移动, 这

便是 1947 年克肯达尔 (Kirkendall) 等人在实验研究 α -黄铜中 Zn 和 Cu 的扩散时首先提出的所谓克肯达尔效应。由扩散方程可定量地考察这一效应对互扩散中组元扩散系数间关系的影响。

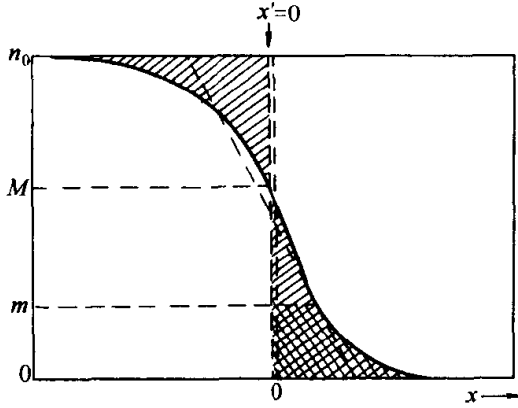


图 17-7 扩散随浓度变化的 Matano 分析

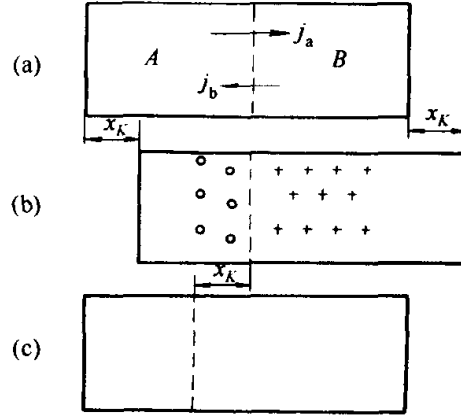


图 17-8 扩散偶中的 Kirkendall 效应

首先设 x 坐标原点位于焊接面, x' 坐标以扩散偶左端面为原点。于是在 x 坐标系中, 通过焊接面组元 B 的流量是:

$$J_b(x=0, t) = -D_b \left(\frac{\partial c_b}{\partial x} \right)_{x=0} = -D_b \left(\frac{\partial c_b}{\partial x} \right)_{x'=L(t)} \quad (17-23)$$

在 x' 坐标系中应包含一附加迁移项, 故有:

$$J_b(x'=L, t) = -D_b \left(\frac{\partial c_b}{\partial x'} \right)_{x'=L} + (vc_b)_{x'=L(t)} \quad (17-24)$$

式中, v 为焊接面 [$x'=L(t)$] 相对于试样左端面的迁移速度。显然, 对于组元 A 也有类似的方程。因此, 原子总浓度的连续性方程应为:

$$\begin{aligned} \frac{\partial c}{\partial t} &= \frac{\partial c_b}{\partial t} + \frac{\partial c_a}{\partial t} = - \left[\frac{\partial J_b}{\partial x'} + \frac{\partial J_a}{\partial x'} \right] \\ &= \frac{\partial}{\partial x'} \left(D_a \frac{\partial c_a}{\partial x'} + D_b \frac{\partial c_b}{\partial x'} - vc \right) \end{aligned} \quad (17-25)$$

当总浓度恒定时, 则 $\partial c / \partial t = 0$, 同时考虑到试样左端面 ($x'=0$) 处浓度梯度和 v 均为零, 则上式括号内表达式应为零, 故得焊接面移动速度为:

$$v = (D_a - D_b) \frac{dc_a}{c dx'} \quad (17-26)$$

将上式代入式 (17-25), 得:

$$\frac{\partial c_b}{\partial t} = \frac{\partial}{\partial x'} \left(\frac{D_b c_a + D_a c_b}{c} \left[\frac{\partial c_b}{\partial x'} \right] \right) \quad (17-27)$$

比较扩散方程标准形式, 可导出互扩散过程中的互扩散系数:

$$D = D_a \nu_b + D_b \nu_a \quad (17-28)$$

此处 D_a 、 D_b 、 ν_a 、 ν_b 分别为组元 A、B 的本征扩散系数及其浓度分数。根据上式关系和实验测得的 D 和 c ，可计算出组元的本征扩散系数 D_a 或 D_b 。例如在 Cu-22% Zn 合金中，于温度为 785℃ 时有： $D_{Zn}/D_{Cu} = 2.3$ ； $D_{Zn}(\nu_{Zn} = 0.22)/D_{Zn}(\nu_{Zn} = 0) = 17$ 。

17.1.2 质点迁移的微观机构与扩散系数

(1) 无规布朗运动理论 菲克定律定量地描述了质点扩散的宏观行为，在人们认识和掌握扩散规律过程中起了重要的作用。然而，菲克定律给出的仅仅是对现象的描述，它将除浓度以外的一切影响扩散的因素都包括在扩散系数之中，且难以赋予其明确的物理意义。1905 年爱因斯坦在研究大量质点作无规布朗运动的过程中，首先用统计的方法得到扩散方程，它使宏观扩散系数与扩散质点的微观运动得到联系。

考虑一个一维长棒中的扩散情况。设在扩散过程的某一时刻 t 参与扩散的质点浓度分布为 $c(x, t)$ ，扩散质点在 t 时刻位于 x 和 $x + dx$ 之间的几率为 $f(x, t)dx$ ，函数 $f(x, t)$ 称为质点位移的分布函。扩散经 τ 时间后，质点的浓度分布将由 $c(x, t)$ 变成 $c(x, t + \tau)$ ，并由下面方程得到联系：

$$c(x, t + \tau) = \int_{-\infty}^{+\infty} f(x - x', \tau) c(x', t) dx' \quad (17-29)$$

令 $\xi = x - x'$ ，并考虑当时间 τ 很短时，质点位移量 ξ 是一小量，故可将上式左右两边依 τ 和 ξ 的幂级数展开：

$$c(x, t + \tau) = c(x, t) + \tau \frac{\partial c}{\partial t} + \tau^2 \frac{\partial^2 c}{\partial t^2} + \dots \quad (17-30)$$

$$c(x - \xi, t) = c(x, t) - \xi \frac{\partial c}{\partial x} + \xi^2 \frac{\partial^2 c}{\partial x^2} + \dots \quad (17-31)$$

代入式 (17-29) 的右边，求积分、略去高次项并与式 (17-30) 比较便得爱因斯坦扩散方程：

$$\frac{\partial c}{\partial t} = \frac{\bar{\xi}^2}{2\tau} \left(\frac{\partial^2 c}{\partial x^2} \right); \bar{\xi}^2 = \int_{-\infty}^{+\infty} \xi^2 f(\xi, t) d\xi \quad (17-32)$$

不难理解，若质点可各向同性地沿空间三维方向迁移，式 (17-32) 可相应地推广至三维情况：

$$\frac{\partial c}{\partial t} = \frac{\bar{\xi}^2}{6\tau} \left(\frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2} + \frac{\partial^2 c}{\partial z^2} \right) \quad (17-33)$$

因此，与扩散第二方程比较，可得质点随机迁移统计意义上的扩散系数：

$$D = \frac{\bar{\xi}^2}{6\tau} \quad (17-34)$$

式中, $\bar{\xi}^2$ 为扩散质点在时间 τ 内位移平方的平均值。对于固态扩散介质, 设质点平均迁移自由程为 a , 有效跃迁频率为 ν , 则有 $\bar{\xi}^2 = \nu\tau a^2$ 。将此关系代入式 (17-34) 得:

$$D = \frac{\bar{\xi}^2}{6\tau} = \frac{\nu a^2}{6} \quad (17-35)$$

由此可见, 扩散的布朗运动理论诠释了菲克定律中扩散系数的物理意义, 为从微观角度研究扩散系数奠定了物理基础。它是建立扩散微观机制与宏观扩散系数间关系的桥梁。

(2) 质点迁移的微观机制 本章一开始已经指出, 构成晶体的每一个质点均束缚在三维周期性势阱中, 故质点的迁移方式或扩散的微观机制将受到固体结构对称性和周期性的限制。迄今已为人们所认识的晶体中原子或离子的迁移机构主要分为两种: 空位机构和间隙机构。

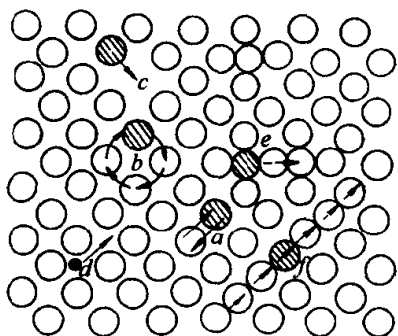


图 17-9 晶体中质点扩散
的微观机制

a —直接交换; b —环形交换; c —空位;
 d —间隙; e —填隙; f —挤列

所谓空位机构的原子或离子迁移过程如图 17-9 中 c 所示, 晶体中由于本征热缺陷或杂质离子的不等价取代而存在空位。于是, 空位周围格点上的原子或离子就可能跳入空位, 而原来的空位则相对跳入空位的原子作相反方向的迁移。因此在固体结构中, 空位的移动意味着结构中原子或离子的相反方向移动。这种以空位作为媒介的质点扩散方式就称为空位机构。实验表明, 无论金属体系或离子化合物系统, 空位机制是固体材料中质点扩散的主要机构。在一般情况下, 离子晶体是由离子半径不同的阴离子和阳离子构成,

而较大尺寸阴离子扩散多半是通过空位机构进行的。

图 17-9 中 d 给出了质点通过间隙机构进行扩散的物理图像。处于间隙位置的质点从一间隙位移入另一间隙位的过程必然引起周围晶格的变形, 且间隙原子相对晶格位上原子或离子的尺寸越大, 间隙机构引起的晶格变形越大, 间隙机构也就越难发生。

除以上两种扩散机构外, 还存在如图 17-9 中所示的 a 、 b 、 e 等几种扩散方式。 e 称之为亚间隙机构。位于间隙位的原子 A 通过热振动将格点上的原子 B 弹入间隙为 C , 而原子 A 进入晶格位 B 。这种扩散机构所造成的晶格畸变程度约处于空位机构和间隙机构之间。已有文献报道, AgBr 晶体中的 Ag 和具有萤石结构的 UO_{2+x} 晶体中的 O^{2-} 的扩散属于这种机构。此外, a 、 b 分别称为直接易位和环易位机构。在这些机构中, 处于对等位置

上的两个或两个以上的格点原子同时跳动进行位置交换而发生迁移。其中环易位机制被认为所需的激活能比直接易位机制要小的多，是一种无点缺陷晶体结构中可能发生的扩散机构。但出现这种机构的几率并不大，因为它需要扩散质点作集体有规则的协同运动。有文献报道，在 $\text{CaO-Al}_2\text{O}_3\text{-SiO}_2$ 三元系统熔体中的 O^{2-} 离子扩散近似依这种机构进行。

(3) 质点跃迁与扩散系数 虽然晶体中以不同微观机制进行的质点扩散有不同的扩散系数，但通过爱因斯坦扩散方程所赋予扩散系数的物理涵义[式(17-35)]，则有可能建立不同扩散机构与相应扩散系数的关系。在空位机构中，结点原子成功跃迁到空位中的频率 ν 应为原子成功跃过能垒 ΔG 的次数和该原子周围出现空位的几率的乘积所决定：

$$\nu = A\nu_0 N_v \exp\left(-\frac{\Delta G_m}{RT}\right) \quad (17-36)$$

式中， ν_0 为格点原子振动频率（约 $10^{13}/\text{s}$ ）； N_v 为空位浓度； A 为比例系数。若考虑空位来源于晶体结构中本征热缺陷，例如肖特基缺陷，则式(17-36)中 $N_v = \exp[-\Delta G_f/2RT]$ ，此处 ΔG_f 为空位形成能。将该关系和式(17-36)代入式(17-35)，便可得到空位机构的扩散系数：

$$D = \gamma a^2 \nu_0 N_v \exp\left(-\frac{2\Delta G_m + \Delta G_f}{2RT}\right) \quad (17-37)$$

式中， $\gamma = A/6$ 。因空位来源于本征热缺陷，故该扩散系数称为本征扩散系数或自扩散系数。

对于以间隙机构进行的扩散，由于晶体中间隙原子浓度往往很小，间隙原子周围的邻近间隙位置几乎均是空的，故间隙机构扩散时可提供间隙原子跃迁的位置几率可近似地看成为 100%。基于与上述空位机构同样的考虑，间隙机构的扩散系数可表达为：

$$D = \gamma a^2 \nu_0 \exp\left(-\frac{\Delta G_m}{RT}\right) \quad (17-38)$$

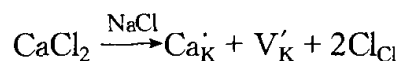
比较式(17-37)和式(17-38)，可以看出，无论扩散依空位机构或间隙机构进行，其扩散系数在数学上具有相同的形式，故可统一表达为如下阿累尼乌斯公式：

$$D = D_0 \exp\left(-\frac{Q}{RT}\right) \quad (17-39)$$

此处 D_0 为非温度显函数项，常称为频率因子； Q 为扩散活化能。对于空位扩散机构，扩散活化能由空位形成能和空位迁移能构成，而间隙机构扩散的活化能只包含间隙原子的迁移能。

应该指出，在实际晶体材料中空位的来源除本征热缺陷以外，往往还包

括杂质离子固溶所引进的空位。例如在 NaCl 晶体中引入 CaCl_2 则 will 发生如下取代反应:



因此, 空位机构扩散系数中应该考虑结构的总空位浓度 $N_v' = N_v + N_i$, 其中 N_v 和 N_i 分别为本征空位浓度和杂质空位浓度。此时, 扩散系数应为:

$$D = \gamma a^2 \nu_o (N_v + N_i) \exp\left(-\frac{2\Delta G_m + \Delta G_f}{2RT}\right) \quad (17-40)$$

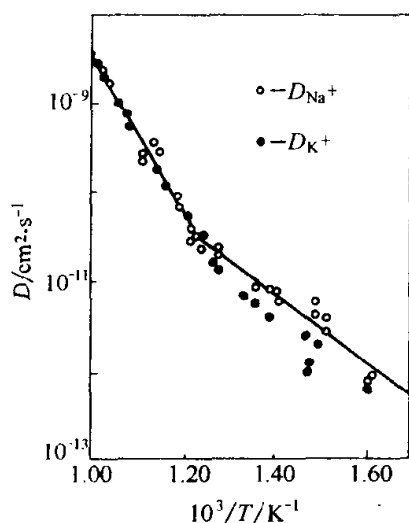


图 17-10 NaCl 单晶中 Na^+ 的自扩散系数

在温度足够高的情况下, 结构中来自本征缺陷的空位浓度 N_v 可远大于 N_i , 此时扩散常为本征缺陷所控制; 当温度足够低时, 结构中本征缺陷提供的空位浓度 N_v 远小于 N_i , 且因扩散主要受固溶引入杂质离子的电价和浓度等因素所控制, 故常称之为非本征扩散。此时, 相应的 D 则称为非本征扩散系。

图 17-10 给出了固溶微量 CaCl_2 的 NaCl 的晶体中, Na^+ 的自扩散系数 D 与温度 T 的关系。在高温区活化能较大的应为本征扩散, 而活化能较小的低温区所对应的应为本征扩散。Patterson 等人测量了单晶 NaCl 中 Na^+ 和 Cl^- 离子的本征扩散系数并得到如表 17-1 所示的活化能数据。

(4) 质点跃迁的结构相关效应 在上述原子跃迁机构和扩散系数关系讨论中默认了原子具有跃迁各向同性的条件, 但事实上原子跃迁的方向性会受到跃迁环境结构对称性的影响。

下面将讨论在化学上等同的同位素 A^* 在 A 中依空位机构进行的扩散来阐述这种影响。如图 17-11 所示, 密排面上同位素 A^* 从位 6 跃到 7, 相应于原位 7

表 17-1 NaCl 单晶中 Na^+ 和 Cl^- 自扩散活化能

离子	活化能 $Q/\text{kJ}\cdot\text{mol}^{-1}$		
	$H_m + H_{1/2}$	H_m	H_i
Na^+	174	74	199
Cl^-	261	162	199

上的空位移到 6。此时, 值得注意的是空位的下一次跳跃将具有各向同性, 而 A^* 周围的环境则完全不同于跃迁之前。几率最大的跃迁将是跳回位 6。显然, 这种择优回跳是直接与前一步跳跃相关的, 因为曾使前一步跳跃得以发生的空位仍在 A^* 的邻近。因此它与爱因斯坦扩散方程中各向同性的统计考虑存在偏差, 并导致 A 的自扩散系数不等于同位素 A^* 在 A 中实测

的扩散系数, 其偏差大小定义为:

$$\theta = \lim_{m \rightarrow \infty} \left(\frac{\bar{r}_m^2(A^*)}{\bar{r}_m^2(A)} \right) = \frac{D_{A^*}}{D_A} \quad (17-41)$$

如图 17-11 所示, 位 6 与位 7 间交换位置的几率正好是空位跃迁几率的 $1/6$ (若有 N 个最邻近的位置, 则是 $1/N$)。假如发生这种情况, 则 A^* 有两次跳跃对于 $\bar{r}_m^2(A^*)$ 没有贡献。因此, 从式 (17-41) 可以近似地得到 $\theta = 1 - 2/N$, 对于立方结构可精确地计算得到 $\theta = (1 + \cos\phi) / (1 - \cos\phi)$, 其中 ϕ 是 A^* 两

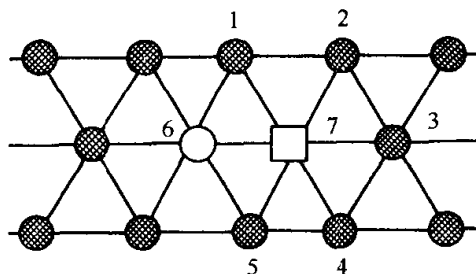
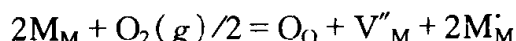


图 17-11 密排面上空位和同位素扩散示意图

次连续跳跃矢量间夹角按其频率对所有可能跳跃的平均值。对于不同晶体结构的典型 θ 值是: fcc 和 hcp 结构为 0.78, bcc 结构为 0.73, 金刚石结构为 0.5。

(5) 非化学计量氧化物中的扩散 除掺杂点缺陷引起非本征扩散外, 非本征扩散也发生在一些非化学计量氧化物晶体材料中, 特别是一些过渡金属元素氧化物如 FeO、NiO、CoO 或 MnO 等材料中。在这些氧化物晶体中, 金属离子的价态常因环境中的气氛变化而改变, 从而引起结构中出现阳离子空位或阴离子空位并导致扩散系数明显地依赖于环境中的气氛。在这类氧化物中典型的非化学计量空位形成可有如下两类情况。

① 金属离子空位型。造成这类非化学计量空位的原因往往是环境中氧分压升高迫使部分诸如 Fe^{2+} 、 Ni^{2+} 、 Mn^{2+} 等二价过渡金属离子变成三价金属离子:



当缺陷反应平衡时, 平衡常数 K_p 由反应自由能 ΔG_o 控制:

$$K_p = [\text{V}''_\text{M}] [\text{M}'_\text{M}]^2 / p_{\text{O}_2}^{1/2} = \exp \{ -\Delta G_o / RT \} \quad (17-42)$$

并有 $[\text{M}'_\text{M}] = 2 [\text{V}''_\text{M}]$ 关系, 因此非化学计量空位浓度 $[\text{V}''_\text{M}]$:

$$[\text{V}''_\text{M}] = (1/4)^{1/3} p_{\text{O}_2}^{1/6} \exp \{ -\Delta G_o / 3RT \} \quad (17-43)$$

将式(17-43)代入式(17-40)中空位浓度项, 可得非化学计量空位浓度对金属离子空位扩散系数的贡献:

$$D = \gamma a^2 \nu_o \left(\frac{1}{4} \right)^{1/3} p_{\text{O}_2}^{1/6} \exp \left(-\frac{3\Delta G_m + \Delta G_o}{3RT} \right) \quad (17-44)$$

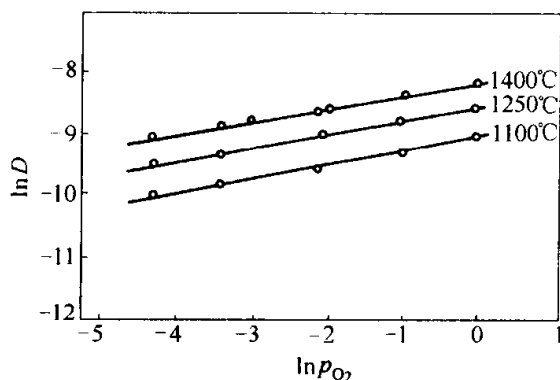
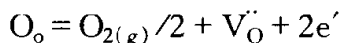


图 17-12 氧分压对 CoO 中 Co^{2+} 离子扩散系数的影响

显然, 在等温条件下, 由上式决定的 $\ln D - \ln p_{\text{O}_2}$ 图线是一斜率为 $1/6$ 的直线, 图 17-12 为实验测得氧分压对 CoO 中 Co 离子空位扩散系数的影响关系, 其直线斜率为 $1/6$, 理论分析与实验结果完全一致。

② 氧离子空位型。以 ZrO_2 为例, 高温时氧分压的降低将导致如下缺陷反应发生: 氧离子离开晶格生成氧气, 并在晶格上留下带正电

的空位 V_{O}^{\bullet} :



其反应平衡常数:

$$K_p = p_{\text{O}_2}^{1/2} [V_{\text{O}}^{\bullet}] [e']^2 = \exp \{ -\Delta G_{\text{O}}/RT \}$$

考虑平衡时 $[e'] = 2[V_{\text{O}}^{\bullet}]$, 故有:

$$[V_{\text{O}}^{\bullet}] = (1/4)^{-1/3} p_{\text{O}_2}^{-1/6} \exp \{ -\Delta G_{\text{O}}/3RT \} \quad (17-45)$$

于是, 非化学计量空位对氧离子的空位扩散系数贡献为:

$$D = \gamma a^2 \nu_{\text{O}} \left(\frac{1}{4} \right)^{-1/3} p_{\text{O}_2}^{-1/6} \exp \left(-\frac{3\Delta G_{\text{m}} + \Delta G_{\text{O}}}{3RT} \right) \quad (17-46)$$

比较式(17-46)和式(17-44), 可以看出, 氧分压 p_{O_2} 的增加将有利于金属离子的扩散而不利氧离子的扩散。考虑式(17-36)和式(17-37), 还可看出, 非化学计量氧化物扩散系数的温度依赖关系表现为在 $\ln D - 1/T$ 图中可能出现含两个转折点的线性区间。高温区和低温区分别为本征空位扩散和杂质空位扩散行为的表现, 而在中温区则可归属于非化学计量空位扩散机构所致。图 17-13 示意地给出了这一关系的图像。

17.1.3 扩散的热力学理论

(1) 扩散方程的推广 众所周知, 在实际材料系统中, 化学组分以及参与扩散的原子或离子往往不限于一种。因此, 在有多个组元构成的扩散体系中不同扩散组元之间不可避免地会发生交互作用而使得简单的扩散方程不能有效地描述这种交互作用引起的复杂扩散现象, 为此有必要将扩散方程作一定的推广。

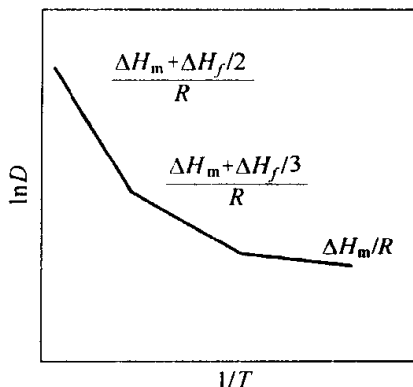


图 17-13 缺氧型氧化物中扩散系数与温度的关系

一种适应于偏离平衡态不太远的线性考虑是将多元扩散体系中组元 i 的扩散流密度 J_i 用如下线性方程表达:

$$J_i = - \sum D_{ij} \left(\frac{dc_j}{dx} \right); i, j = 1, 2, \dots, k \quad (17-47)$$

即认为第 i 种组元的扩散流 J_i 不仅受其本身浓度梯度的控制, 同时也受到其他组元浓度梯度的影响 ($D_{ij} \neq 0$)。

应该指出, 方程(17-47)的形式仍限于经典扩散理论的框架之内, 它表明在扩散体系中出现宏观物质流的驱动力是浓度梯度, 且扩散总是朝负梯度方向进行。显然, 这一扩散概念不能很好地解释温度场或其他外场存在下质点的定向迁移运动以及发生诸如在失稳分解过程 (spinodal decomposition) 中扩散质从低浓度区向高浓度区的所谓爬坡扩散 (如钢中奥氏体分解时, 过饱和奥氏体不断向 Fe_3C 输送碳原子)。根据普适的热力学理论, 可以认为扩散过程与其他物理化学过程一样, 其发生的根本驱动力应该是化学势梯度。基于这个基本概念, 式 (17-47) 可进一步推广为:

$$J_i = - \sum L_{ij} \left(\frac{d\mu_j}{dx} \right); i, j = 1, 2, \dots, k \quad (17-48)$$

这里 μ_j 为组元 j 的化学势; L_{ij} 为扩散的唯象系数, 它是组分、温度以及化学势等的函数。 L_{ii} 是组元 i 自身活动性大小的量度, 而 L_{ij} ($i \neq j$) 则是衡量第 i 组元与第 j 组元之间交互作用大小的参数。根据不可逆热力学 Onsager 倒易关系 $L_{ij} = L_{ji}$, 一个 N 元扩散体系的 N^2 个唯象系数中仅有 $(N+1)N/2$ 个是独立的。这一倒易关系的存在为多元扩散体系的实验研究提供了很大的方便。

(2) 扩散系数中的热力学因子 由式(17-48)可以看到, 浓度梯度为零并非扩散流消失的普遍有效条件, 而化学势恒定才是扩散达到平衡的必要条件。下面将运用这一普遍的扩散热力学理论讨论恒温无外场条件下浓度梯度与化学势梯度的关系。为简单起见, 考虑一二元系统并认为组元间相互作用甚微可忽略不计。因此, 根据式 (17-48) 得:

$$J_a = - L_{aa} \left(\frac{d\mu_a}{dx} \right); J_b = - L_{bb} \left(\frac{d\mu_b}{dx} \right) \quad (17-49)$$

将化学位热力学关系式 $\mu_i = \mu_{i0} + kT \ln c_i \gamma_i$ 代入上式, 即有:

$$\begin{aligned} J_a &= - L_{aa} kT \left(1 + \frac{d \ln \gamma_a}{d c_a} \right) \frac{dc_a}{dx} \\ J_b &= - L_{bb} kT \left(1 + \frac{d \ln \gamma_b}{d c_b} \right) \frac{dc_b}{dx} \end{aligned} \quad (17-50)$$

这里 μ_{i0} 为标准状态下纯组元 i 的化学势; c_i , γ_i 为组元 i 的分数浓度和活度系数。比较式(17-49)和式(17-50), 容易看出浓度梯度与化学势梯度对扩

散流作用的本质差别。更有意义的是将上式与扩散方程比较, 则可得出本征扩散系数的热力学表达式:

$$D_a = L_{aa} kT \left(1 + \frac{d \ln \gamma_a}{d c_a} \right) \quad D_b = L_{bb} kT \left(1 + \frac{d \ln \gamma_b}{d c_b} \right) \quad (17-51)$$

上式表明: 质点的本征扩散系数与其浓度、活动性、活度系数以及温度均密切相关。根据溶液热力学理论, 在无限稀的情况下, 体系近似为理想固溶体并有 $\gamma = 1$ 。将此条件代入上式并令 $D_i^* = L_{ii} kT$, 便推得著名的 Darken 方程:

$$D_a = D_a^* \left(1 + \frac{d \ln \gamma_a}{d c_a} \right) \quad D_b = D_b^* \left(1 + \frac{d \ln \gamma_b}{d c_b} \right) \quad (17-52)$$

式中, D_i^* 称为组元 i 的自扩散系数。Darken 方程说明了本征扩散系数与自扩散系数间的关系。其中 $[1 + d \ln \gamma_i / d c_i]$ 称之为扩散系数的热力学因子。容易看出, 通常发生的扩散系数 $D > 0$ 的扩散正是热力学上相应于 $[1 + d \ln \gamma_i / d c_i] > 0$ 的情况, 而当 $[1 + d \ln \gamma_i / d c_i] < 0$ 时, 扩散系数变为负数, 从而导致迁移质点的爬坡扩散。此时质点的扩散不是有利于体系组元浓度趋于均匀, 而是使其发生分聚。考虑普适的 Gibbs-Duhem 热力学关系式: $[d \ln \gamma_a / d \ln c_a] = [d \ln \gamma_b / d \ln c_b]$, 并引用式(17-51)和式(17-52), 可得:

$$\frac{D_a}{D_b} = \frac{L_{aa}}{L_{bb}} = \frac{D_a^*}{D_b^*} \quad (17-53)$$

该结果表明, 对于二元系统来说, 只有组元的活动性或自扩散系数相等, 它们的本征扩散系数 D_a 和 D_b 才能相等。而前面所述的克肯达尔效应正是由于扩散组元的活动性或自扩散系数不相等所造成的。

最后值得指出, 在上述热力学因子导出和分析过程中, 扩散质之间交互作用的忽略并非总是合理地近似。在很多情况下, 空位常常容易达到热力学的平衡状态, 但交互作用则不一定等于零。例如, 在有空位扩散流与各种质点扩散流交互作用的情况中, 根据克肯达尔效应, 空位的净扩散流不等于零, 而为:

$$J_v = - \sum L_{vj} \left(\frac{d \mu_j}{d x} \right) \quad (17-54)$$

即 J_v 由交叉项 $L_{vj} (j < n - 1)$ 所决定。故若 J_v 不为零, 交叉项必定不全为零。

(3) 热力学因子在扩散中的作用 在 α -黄铜(铜锌合金)中, 当锌的含量从零增加到 22% 时, 锌的扩散系数 D_{Zn} 增大 17 倍。在这个范围里, 热力学因子从 1 增加到 1.8。这种扩散系数 D_{Zn} 随空位 v_{Zn} 增加而迅速增大的部分原因是热空位的浓度随着与固相线差距的缩小而增加以及锌原子与空位间的相互吸引作用, 从而促进锌的扩散。

在具有失稳分相的合金体系中, 热力学因子会扮演极其重要的作用。在这种情况下, 两相区的亚稳分相线决定于自由能二阶导数 $[d^2 G / d c^2] = 0$

的条件并与有效热力学因子有如下关系：

$$c \left(\frac{\partial^2 G}{\partial c^2} \right) = \frac{d\mu}{dc} = \frac{kT}{c} \left(1 + \frac{d \ln \gamma}{d \ln c} \right) \quad (17-55)$$

在亚稳分相线以下，自由能二阶导数 $[d^2 G / dc^2] < 0$ ，因而导致扩散系数

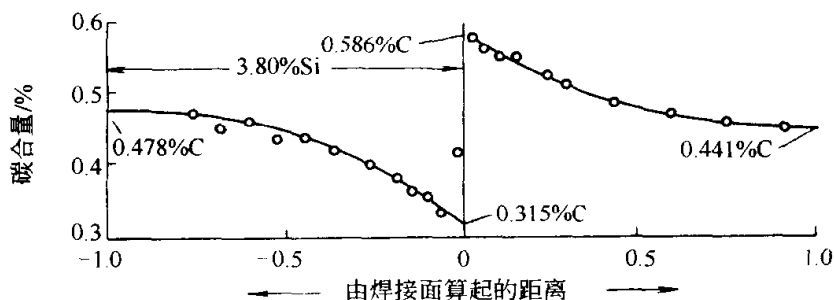


图 17-14 Fe-Si/Fe 扩散偶在 1050℃ 退化 13 天后的碳分布

出现负值，产生负扩散或爬坡扩散并使原来均匀的浓度场出现起伏。Dehlinger 和 Becker 首先描述了由负扩散所导致的失稳分相现象。

Darken 用 γ -Fe-0.4%C-0.4%Si 扩散偶在 1050℃ 扩散的数据，说明了热力学因子在三元扩散体系中的重要意义。如图 17-14 所示，在焊接面两边建立的碳浓度梯度，似乎违背了菲克定律，但却降低了由于 Si 浓度差所引起的化学势梯度。扩散进行一段时间后，当焊接面两边的硅浓度达到平衡时，碳浓度呈现相等。在图 17-15 所示的三元相图中，S 图 17-15 扩散偶两边 A、B 点组学成分随扩散时间的变化

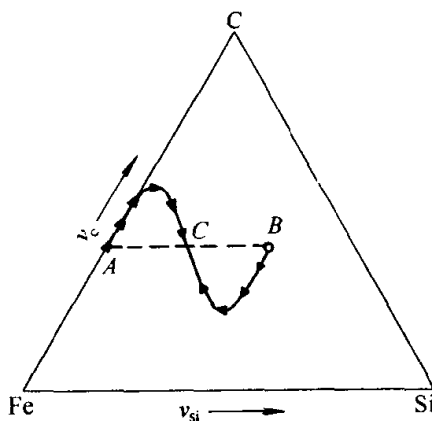


图 17-15 扩散偶两边 A、B 点组学成分随扩散时间的变化

17.1.4 短路扩散与非平衡态下点缺陷的扩散

在前面的各节中，有关的讨论均限于扩散发生在结构均匀的介质中。然而，实际的晶体材料内部往往存在许多部分或完全无序的结构区域，如晶界、相界、表面以及各种位错区。实验表明，在这些区域中，原子或离子的扩散比在结构完整的体相中要快得多。换言之，这些有着大量结构缺陷的区域实际上是扩散的快速通道，故而称之为短路扩散。此外，在一些材料加工处理过程中（如金属材料的冷加工，表面辐照等），非平衡态下的过饱和空位与间隙缺陷的稳定存在也可促使扩散过程的加快，并有可能造成合金中出现非平衡相沉淀。

(1) 晶界扩散 在晶界扩散的理论处理中，可考虑一个如图 17-16 所示

的垂直于晶体表面的薄夹层作为晶界。在晶界露头的表面上蒸发沉淀放射性同位素，它将沿晶界方向快速扩散，进而晶界中高浓度的原子向晶界的两边体相内扩散，从而形成如图 17-16 中所示的浓度分布曲线。费歇

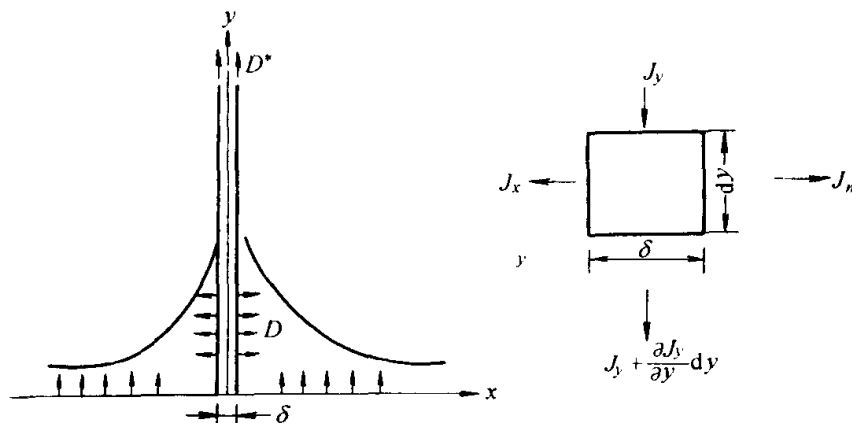


图 17-16 晶界扩散与扩散浓度分布示意图

(J.C.Fisher)对这种扩散作了近似数学处理：当 $t > 0$ 及 $y = 0$ 时， $c = c_0$ ；令同位素在厚度为 $\delta = 10^{-7} \text{ cm}$ 的晶界层中的扩散系数为 D_G^* ，在点阵中为 $D_L^* \ll D_G^*$ ，故在晶界区域的微元内，同位素浓度随时间按下式变化：

$$\frac{\partial c}{\partial t} = -\frac{\partial J_y}{\partial y} - 2J_x/\delta = D_G^* \frac{\partial^2 c}{\partial y^2} + 2D_L^* \frac{1}{\delta} \left(\frac{\partial c}{\partial x} \right)_{x=\delta/2} \quad (17-56)$$

晶界内的浓度 $c(x=0, y) = c_G$ 很快地达到准稳态的分布，然后扩散质沿 x 方向扩散。为简单起见，将含有初始未知数 $c_G(y)$ 的薄膜解 $c(x)$ 代入式(17-56)中，然后计算 c_G 得到完整的解为：

$$c(x, y, t) = c_0 \exp \left[\frac{-y}{(\pi D_L^* t)^{1/4}} \left(\frac{\delta D_G^*}{2 D_L^*} \right)^{1/2} \right] \left[1 - \operatorname{erf} \frac{x}{2 \sqrt{D_L^* t}} \right] \quad (17-57)$$

将上式对 x 全区间积分，便可得到浓度沿 y 方向的平均分布 $\bar{c}(y, t)$ 。绘制 $\ln \bar{c}$ 对 y 的关系图，可得一斜率倒数为 $(D_L^* t \eta)^{1/2} T^{1/4}$ 的直线，式中 $\eta = \delta D_G^* / 2 D_L^* (D_L^* t)^{1/2}$ ，称为晶界因子。因此，知道 D_L 及 η ，就可得到 $D_G \delta$ 。图 17-17 中给出了不同 η 值的浓度分布曲线，费歇的解则对应于 $\eta \gg 1$ 的条件。若 $t = 10^5 \text{ s}$ ， $D_L = 10^{-11} \text{ cm}^2/\text{s}$ ， $\delta = 4 \times 10^{-8} \text{ cm}$ ，可计算得 D_G/D_L 应大于 5×10^4 。

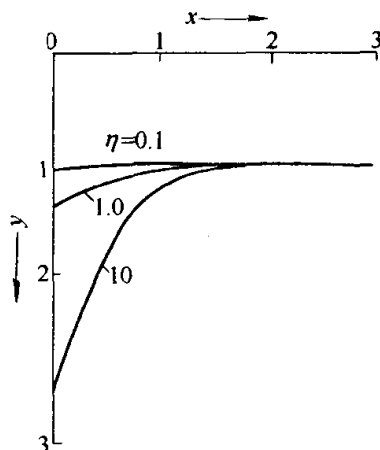


图 17-17 不同 η 值下
浓度分布曲线

对晶界扩散的研究除费歇的数学模型外，还有一些唯象的或微观的理论研究。但实际材

料中原子或离子沿晶界扩散的状况远比理论模型描述要复杂得多。这是因为材料内部晶界浓度高，它们的取向相对于扩散方向往往呈无规分布，内部的晶界可连通成一个极为复杂的三维扩散网络而使准确的定量描述变得非常困难。目前，众多晶界扩散研究达成的共识是晶界自扩散活化能约为体扩散活化能的 0.4~0.6 倍。

(2) 表面扩散 晶体材料的自由表面并不是一个完美的几何面。如图 17-18 所示，它们常常由一些低指数晶面的台阶组成，并由原子尺寸高度带有弯折的凸缘分开。此外，表面上还存在着点缺陷，包括吸附到表面上的原子和空位。它们的形成能和迁移能在数值上差别不大，但对表面取向高度敏感，故而预计这些缺陷对表面扩散起着重要作用。金属材料研究的一些结果表明，吸附原子的迁移能随表面指数的不同而不同，与吸附原子的迁移相比，吸附空位是较慢的扩散缺陷。而且，在非完整的表面上，扩散是高度各向异性的，沿密排方向的迁移能较小。

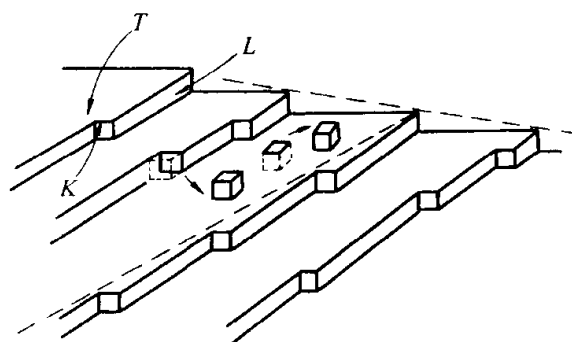


图 17-18 低指数表面的台阶(T)、凸缘(L)和弯折(K)

金属材料研究的一些结果表明，吸附原子的迁移能随表面指数的不同而不同，与吸附原子的迁移相比，吸附空位是较慢的扩散缺陷。而且，在非完整的表面上，扩散是高度各向异性的，沿密排方向的迁移能较小。

表面扩散的定量研究可采用晶界研究的数学模型。如图 17-16 所示，将构成晶界的其中一个晶粒移走，便是表面扩散的情况了。鉴于晶界模型的对称性，故可用式(17-57)解的 $1/2$ 来描述同位素从样品边缘开始，沿 y 方向在表面上展开的扩散动力学过程。当然，此时 δ 为表面扩散层厚度，并应用表面扩散系数 D_S 取代 D_G 。实验表明，通过类似蒸发与沉积的过程附着在表面上的原子，其迁移活化能 Q_{SD} 甚小，且与升华能 W_S 的比值 Q_{SD}/W_S 约在 0.1~0.6 之间。

(3) 非平衡态点缺陷对扩散的影响 前面讨论了许多在热平衡状态下有关点缺陷的扩散性质。然而，在材料加工和处理过程中所涉及的某些重要的技术，如淬火、冷加工、辐照等往往在材料结构中产生过饱和的空位与间隙缺陷。这些过饱和缺陷出现无疑将对相关的扩散作出附加贡献，而且过饱和和间隙缺陷由于形成能和迁移能高，除促进扩散加快以外，还可能导致平衡时出现的如组分偏聚、相变等现象。

晶态材料的自由表面、位错和晶界都可看成空位源，空位又可通过聚集而形成位错环、层错四面体或空洞而消失。但在有些情况下，如位错密度很低时，这些空位源将失效并造成大量过饱和空位从而促进扩散。金属材料淬

火而出现的大量空位就是一个重要的例子。金属材料冷加工的范性变形过程中, 位错的交互作用也可产生以空位为主的点缺陷, 而且这些点缺陷的过饱和度随着位错密度的升高而有所降低。因此这种过饱和点缺陷常表现为在较低温度下明显地促进扩散的进行。此外, 材料表面受到高能电子、离子或中子的辐照会产生大量缺陷。这些缺陷有时呈费仑克尔 (Frenkel) 对, 有时形成复杂的团簇, 并对扩散产生着不可忽视的影响。

17.1.5 带电质点的扩散行为

在离子晶体中, 参与扩散的质点是带有一定电荷的正离子或负离子。这些质点的移动无疑会引起结构内部电场的变化, 而电场的变化又反过来影响离子的扩散行为。考虑到结构内部存在电场和一定的空位浓度 (如发生不等价置换), 扩散方程应具有如下形式:

$$J_i = -D_i \left(\frac{dc_i}{dx} - \frac{c_i}{p} \frac{dp}{dx} + \frac{c_i}{RT} \frac{dh_i}{dx} + \frac{c_i Z_i e}{RT} \frac{d\phi}{dx} \right) \quad (17-58)$$

此处 p 为结构中空位浓度; D_i 、 Z_i 、 e 分别为离子 i 的迁移率、电价和单位电量。在热力学上, 把化学势和电势相结合的形式称为电势化学势。

首先考虑一种简单的不等价置换情况。在 AX_n 晶体中阴离子 X 的扩散系数可认为比阳离子 A 的扩散系数要小的多, 因此在无杂质离子存在时, 结构中仅发生 A 离子的自扩散过程。然而, 若让比 A 离子电价大而扩散系数小的 B 离子取代 A 离子生成固溶体, 则在置换区域内将出现较大的空位浓度和空位向外流动的趋势。显然, 空位的这种迁移流动必然受到电荷平衡的制约。当达到平衡时, 必有 $J_i = 0$ 。此时忽略式 (17-58) 右边的第一、三项, 由积分便可得到结构中电场 ϕ 与空位浓度 p 分布的关系:

$$\ln \left(\frac{p}{p_0} \right) = \frac{Z\phi e}{RT} \quad (17-59)$$

带电质点互扩散可用 AX_n 和 BX_m 晶体扩散偶来进行研究。阴离子 X 仍可认为其扩散系数非常小, 故参与扩散的质点是 A 、 B 离子。当 A 、 B 离子电价相同, 且 AX_n 和 BX_m 晶体结构相同, 则 AX_n 和 BX_m 可形成固溶体, A 、 B 离子的互扩散将以同一数目进行, 而不产生空位浓度梯度。此时应用式 (17-58), 可得:

$$J_i = -D_i \left[\frac{dc_i}{dx} + \frac{c_i Z_i e}{RT} \left(\frac{d\phi}{dx} \right) \right]; \quad i = A, B \quad (17-60)$$

为保持电荷平衡, $J_a + J_b = 0$, 且 $c_a + c_b = c_0$ 。因此, $dc_a/dx + dc_b/dx = 0$ 。代入该条件, 消去电位梯度项, 则得:

$$J_i = \frac{-D_a D_b (C_a + C_b)}{C_a D_a + C_b D_b} \left(\frac{dc_i}{dx} \right); \quad i = A, B \quad (17-61)$$

这一结果表明,在该类型的离子互扩散中,尽管 A、B 的自扩散系数和自扩散量均不尽相同,但其受电位控制的互扩散量决定于相同的互扩散系数和各自的浓度梯度。扩散平衡时,扩散电位为:

$$\frac{d\phi}{dc_a} = \frac{D_b - D_a}{D_a c_a + D_b c_b} \left(\frac{RT}{Ze} \right) \quad (17-62)$$

作 $0 \sim c_a$ 的积分,得出:

$$\phi = \frac{RT}{Ze} \ln (\tilde{D}/D_a) \quad (17-63)$$

17.1.6 金属中的电迁移和热迁移

当在合金材料中通过大电流时,电位梯度的出现往往可能导致像发生在离子晶体中的电解现象,荷电较大的组元离子实在材料中发生漂移并在阴极上聚集,如 α -Fe 中碳离子由间隙机制向阴极的迁移。在电场中,这种漂移运动的方向更多地取决于材料中载流子的类型——电子或空穴。同时,载流子的运动倾向于将迁移中的离子实拖向相反的电极。根据式 (17-58),对于 A 中溶有微量 B 的合金系统,当 B 处于间隙位时,其扩散流密度为:

$$J_B = -D_B \left[\frac{dc_B}{dx} + \frac{c_B Z_B^* e}{RT} \left(\frac{d\phi}{dx} \right) \right] \quad (17-64)$$

式中的第二项便是电迁移对总扩散的附加贡献, $d\phi/dx$ 为材料中电位梯度, Z_B^* 为 B 离子考虑了屏蔽效应后的有效荷电数。根据 Huntington-Fok 的模型, Z_B^* 与标称离子价 Z_B 存在如下关系:

$$Z_B^* = Z_B - \frac{Z_A}{2} \left(\frac{\Delta\sigma_{BS}}{\sigma} \right) \left(\frac{c_B m^*}{c_{BS} |m^*|} \right) \quad (17-65)$$

式中, m^* 为传导电子的有效质量, c_{BS}/c_B 为 B 离子位于鞍点的分数。这些离子可对传导电子产生强烈的散射作用,使合金的导电率相对降低 $\Delta\sigma_{BS}/\sigma$ 。故 $(\Delta\sigma_{BS}/\sigma)(c_B/c_{BS})$ 的乘积表示带电粒子位于点阵位置及鞍点位置时有效散射截面的相对差。为证实电迁移效应,可将两根铜棒之间涂焊一层金属锑,并于恒温下通过每平方厘米几万安培的直流电,然后测量铜-锑-铜结部位 Sb 的浓度分布并与上述的式子进行比较。图 17-19 给出了通电后,铜-锑-铜结部 Sb 的浓度的分布曲线。其拟合结果满足式(17-64)和式(17-65),且指出作用在溶质离子上的直流电场力似乎只取决于标称离子价而无需考虑屏蔽作用。

在存在温度梯度的材料中,热流将对粒子的扩散迁移产生附加贡献,这便是一种热迁移现象。沃兹 (K. Wirtz) 提出了一个单组元 A 的空位扩散模型,如图 17-20 所示,在温度梯度方向考虑质点在三个点阵面 1、2 和 3 上的迁移行为,得到原子 A 迁移的净扩散流密度:

$$\Delta J_{R-L} \propto c_B D_B (E_{VF} - E_{VM}) \frac{\Delta T}{akT^2} \quad (17-66)$$

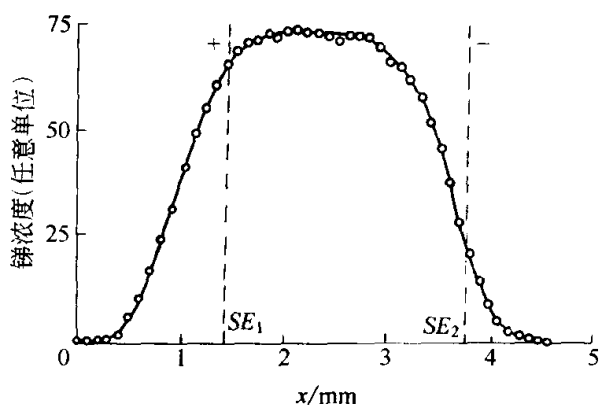


图 17-19 Sb 在铜中电迁移后浓度分布曲线
SE 间为 0.8% Sb-Cu 合金, SE 外为纯铜

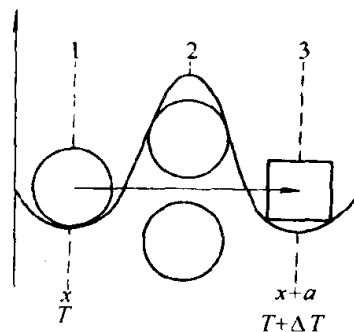


图 17-20 晶格中质点热
迁移示意图

由此可见, 当材料中存在温度梯度时, 在空位形成能和空位迁移能不相等的情况下将发生热迁移现象。

17.2 固相化学反应

17.2.1 固相反应动力学的基本特征

固相反应在固体材料制备的高温过程中是一个普遍的物理化学现象, 它是一系列金属合金材料、传统硅酸盐材料以及各种新型无机非金属材料制备所涉及的基本过程之一。广义地说, 凡是有固相参与的化学反应都可称为固相反应, 例如固体的分解、氧化以及固体与固体、固体与液体之间的化学反应等都属于固相反应研究范畴之内。但在狭义上, 固相反应常指固相与固相间发生化学反应生成新的固相产物的过程。

固相反应与一般气、液反应相比, 在反应机构和反应速度等方面均有自己的特点。①固相反应属非均相反应。因此参与反应的固相相互接触是反应物间发生化学作用和物资输送的先决条件。②固相反应开始温度常远低于反应物的熔点或系统的低共熔点。这一温度与反应物内部开始明显扩散作用的温度是相一致的, 常称为泰曼温度或烧结温度。不同物质的泰曼温度与其熔点 T_m 间存在一定的关系。例如对于金属为 $0.3 \sim 0.5 T_m$; 硅酸盐类的泰曼温度则分别为 $0.8 \sim 0.9 T_m$ 。此外, 当反应物之一存在有多晶转变时, 此转变温度也往往是反应开始变得显著的温度, 这一规律常称为海得华定律。

图 17-21 描述了固相 A 与 B 进行化学反应生成 C 的一种反应历程: 反应一开始是反应物颗粒间的混合接触, 并在表面发生化学反应生成细薄且含大量结构缺陷的新相, 随后发生产物新相的结构调整和晶体生长。当在两反应颗粒间所形成的产物层达到一定厚度后, 进一步的反应将由一种或几种反应物通过产物层的扩散而得以进行。这种物质的运输过程可能通过晶体晶格

内部、表面、晶界、位错或晶体裂缝进行。显然，对于广义的固相反应，由于反应体系存在气相或液相，进一步反应所需要的传质过程往往可以在气相或液相中进行。在这种情况下，气相或液相可能对固相反应起到重要的作用。由此可以认为，固相反应是固体参与化学作用并发生化学变化，同时至少在固体内部或外部存在某一使反应得以持续的传质过程的动力学过程。显然，在此过程中，决定反应速度的不仅限于化学反应本身，反应新相晶格缺陷调整速率、晶体生长速率以及反应体系中物质和能量的输送速率都将影响着反应的速度。

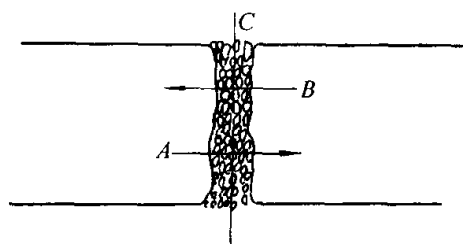


图 17-21 固相物质 A、B 化学反应过程模型

17.2.2 固相反应动力学方程

(1) 固相反应一般动力学关系 上节业已指出，固相反应的基本特点在于反应通常是由几个简单的物理化学过程，如化学反应、扩散、结晶、熔融、升华等步骤构成的。因此，整个反应的速度将受到诸有关动力学阶段速度的影响。显然，所有环节中速度最慢的一环，将对反应速度有着决定性的影响。现以金属氧化为例，建立反应总速度与各阶段反应速度间的关系。

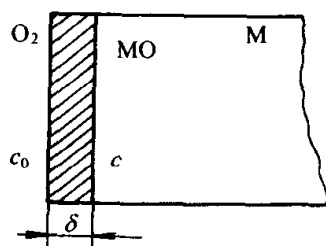
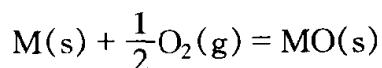


图 17-22 金属表面氧化反应模型

设反应依图 17-22 所示模式进行，其反应方程式为：



反应经 t 时间后，金属 M 表面已形成厚度为 δ 的产物层 MO。进一步的反应将由氧气 O_2 通过产物层 MO 扩散到 M-MO 界面上和金属 M 发生氧化反应两个过程所构成。根据化学反应一般原理和扩散定律，单位面积上金属的氧化速度 V_r 和氧气扩散速度 V_d 分别有如下关系：

$$V_r = Kc; \quad V_d = D \left(\frac{dc}{dx} \right)_{x=\delta} \quad (17-67)$$

式中， K 为化学反应速率常数； c 为界面处氧气浓度； D 为氧气在产物层中的扩散系数。显然，当整个反应达到稳定时，反应总速率 $V = V_r = V_d$ 。由 $Kc = Ddc/dx = D(c_0 - c)/\delta$ ，得界面氧浓度： $c = c_0/[1 + K\delta/D]$ ，故有：

$$\frac{1}{V} = \frac{1}{Kc_0} + \frac{1}{D \left(\frac{c_0}{\delta} \right)} \quad (17-68)$$

由此可见, 由扩散和化学反应构成的固相反应过程总速率的倒数为扩散最大速率的倒数和化学反应最大速率的倒数之和。若将反应速率的倒数理解成反应的阻力, 则式(17-68) 将具有熟悉的串联电路欧姆定律完全相似的形式: 反应的总阻力等于各环节分阻力之和。反应过程与电路的这一类同对于研究复杂反应过程有着很大的方便。例如, 当反应不仅包括化学反应、物质扩散, 还包括结晶、熔融、升华等物理化学过程, 且当这些单元过程间又以串联模式依次进行时, 那么固相反应的总速度应为:

$$V = \frac{1}{\left(\frac{1}{V_{1\max}} + \frac{1}{V_{2\max}} + \frac{1}{V_{3\max}} + \cdots + \frac{1}{V_{n\max}} \right)} \quad (17-69)$$

式中, $V_{1\max}$ 、 $V_{2\max}$ 、…… $V_{n\max}$ 等分别代表构成反应各环节的最大可能速率。

应该指出, 对实际的固相反应过程, 掌握所有反应环节的具体动力学关系有时往往十分困难, 故需抓住问题的主要矛盾才能使问题比较容易地得到解决。例如, 若在固相反应环节中, 物质扩散速率较其他各环节都慢得多, 则由式(17-69)可知反应阻力主要来源于扩散过程。此时, 若其他各项反应阻力较扩散项是一个小量并可忽略不计, 则总反应速率将几乎完全受控于扩散速率。

(2) 化学控制反应动力学 化学反应是固相反应过程的基本环节。根据物理化学原理, 对于二元均相反应系统, 若化学反应依式 $mA + nB = pC$ 进行, 则化学反应速率的一般表达式为:

$$V_r = \frac{dc_c}{dt} = K c_a^m c_b^n \quad (17-70)$$

式中, c_a 、 c_b 、 c_c 分别代表反应物 A、B 和产物 C 的浓度; K 为反应速率常数, 它与温度间存在阿累尼乌斯关系:

$$K = K_0 \exp\left(-\frac{\Delta G_r}{RT}\right) \quad (17-71)$$

此处, K_0 为常数; ΔG_r 为反应活化能。

然而, 对于非均相的固相反应, 式(17-70)不能直接用于描述化学反应动力学关系。这是因为对大多数固相反应, 浓度概念已失去应有的意义。其次, 多数固相反应是以反应物间的机械接触为基本条件。因此, 在固相反应中, 将用反应物的转化率 G 取代式(17-70)中的浓度项, 并同时考虑反应过程中反应物间的接触面积。

反应物转化率常定义为反应物在反应过程中被反应了的体积百分数。设反应物颗粒呈球状, 反应前半径为 R_0 , 经时间 t 后反应物颗粒外层 x 厚度已被反应, 则定义转化率 G :

$$G = \left[\frac{R_o^3 - (R_o - x)^3}{R_o^3} \right] = 1 - \left(\frac{1 - x}{R_o} \right)^3 \quad (17-72)$$

根据式(17-70)的含义, 有固相反应中动力学一般方程:

$$\frac{dG}{dt} = KF(1 - G)^n \quad (17-73)$$

式中, n 为反应级数; K 为反应速率常数; F 为反应截面。当反应物颗粒为球形时, $F = 4\pi R_o^2 (1 - G)^{2/3}$ 。考虑一个一级反应, 由式(17-73)得动力学方程:

$$\frac{dG}{dt} = 4\pi K R_o^2 (1 - G)^{5/3} = K_l (1 - G)^{5/3} \quad (17-74)$$

若反应截面在反应过程中不变(如金属平板的氧化过程), 则有:

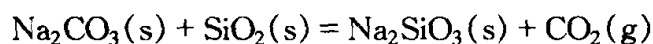
$$\frac{dG}{dt} = K_l' (1 - G) \quad (17-75)$$

积分式(17-74)和式(17-75), 并考虑初始条件 $t = 0, G = 0$, 得反应截面分别依球形和平板模型变化时, 固相反应转化率或反应度与时间的函数关系:

$$F_l(G) = [(1 - G)^{-2/3} - 1] = K_l t \quad (17-76)$$

$$F_l'(G) = \ln(1 - G) = -K_l' t \quad (17-77)$$

碳酸钠 Na_2CO_3 和二氧化硅 SiO_2 在 740°C 下进行固相反应:



当颗粒半径 $R_o = 36\mu\text{m}$, 并加入少许 NaCl 作溶剂, 整个反应动力学过程完全符合式(17-76)关系, 如图 17-23 所示。这说明上述反应于该条件下, 反应总速率为化学反应动力学过程所控制, 而扩散的阻力已小到可忽略不计, 且反应属于一级化学反应。

(3) 扩散控制反应动力学 固相反应一般都伴随着物质的迁移。在固相结构内部扩散速度常常较为缓慢, 因而在多数情况下扩散速度控制着整个反应的总速度。由于反应截面变化的复杂性, 扩散控制的反应动力学方程将彼此不同。在众多的反应动力学方程中, 基于平行板模型和球体模型所导出的所谓扬德尔(Jander)方程和金斯特林格(Ginstleringer)方程具有相当的代表性。

① 扬德尔方程。如图 17-24 (a) 所示, 设反应物 A 和 B 以平板模式相互接触反应与扩散, 经时间 t 后形成厚度为

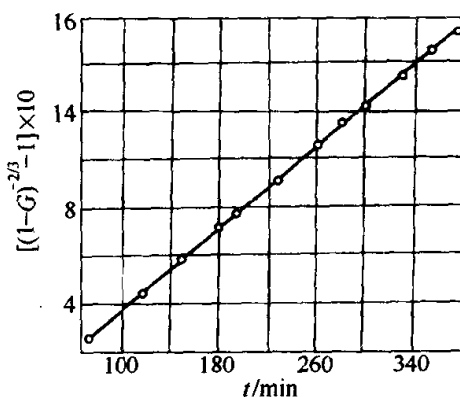


图 17-23 在 NaCl 参与下碳酸钠与二氧化硅反应动力学曲线

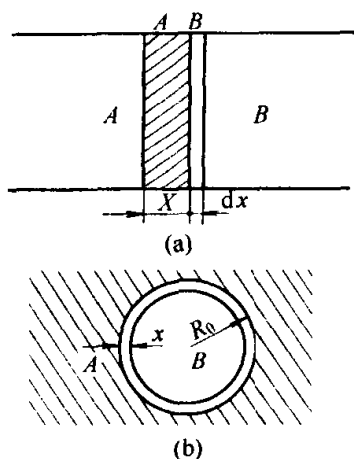


图 17-24 固相反应动力学 Jander 模型

x 的产物 AB 层。随后, 物质 A 通过 AB 层扩散到 B - AB 界面继续与 B 反应。若界面化学反应速度远大于扩散速率, 则可认为固相反应总速率由扩散过程控制。

设 t 到 $t + dt$ 时间内通过 AB 层单位截面的 A 物质量为 dm 。显然, 在反应过程中的任一时刻, 反应界面 B - AB 处物质 A 的浓度为零, 而界面 A - AB 处 A 的浓度为 c_0 。由菲克扩散定律可得:

$$\frac{dm}{dt} = D \left(\frac{dc}{dx} \right)_{x=\xi} \quad (17-78)$$

又设反应产物 AB 密度为 ρ , 分子量为 μ , 则

$dm = \rho dx / \mu$; 当扩散到达稳定时, 有:

$$\left(\frac{dc}{dx} \right)_{x=\xi} = \frac{c_0}{x}; \quad \frac{dx}{dt} = \frac{\mu D c_0}{\rho x} \quad (17-79)$$

积分上式并考虑边界条件 $t = 0$ 时 $x = 0$, 得:

$$x^2 = \frac{2\mu D c_0 t}{\rho} = Kt \quad (17-80)$$

上式说明, 反应物以平行板模式接触时, 反应产物层厚度与时间的平方根成正比, 故常称上式为抛物线速度方程式。当反应物为球状颗粒时, 利用式 (17-72) 可对上式进行修正, 得所谓的扬德尔方程积分式:

$$x^2 = R_0^2 [1 - (1 - G)^{1/3}]^2 = Kt \quad (17-81)$$

或

$$F_J(G) = [1 - (1 - G)^{1/3}]^2 = \frac{Kt}{R_0^2} = K_J t \quad (17-82)$$

较长时间以来, 扬德尔方程一直作为一个较经典的固相反应动力学方程被广泛地接受。但认真分析扬德尔方程推导的过程, 不难发现, 将圆球模型的转化率式 (17-72) 代入平板模型的抛物线速度方程中, 就限制了扬德尔

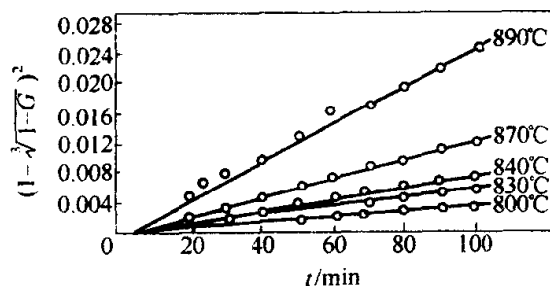


图 17-25 不同温度下碳酸钡与二氧化硅反应动力学 Jander 模型

方程只能用于反应转化率较小和反应截面 F 可以近似地看成常数的反应初期。图 17-25 和图 17-26 分别表示了反应 $\text{BaCO}_3 + \text{SiO}_2 = \text{BaSiO}_3 + \text{CO}_2$ 和 $\text{ZnO} + \text{Fe}_2\text{O}_3 = \text{ZnFe}_2\text{O}_3$ 在不同温度下 $F_J(G)-t$ 关系, 它们证实了扬德尔方程在反应初期的适用性。图中所示温度变化所引起直线斜率的改变则完全是由反应速率常数 K_J 变化所至。

② 金斯特林格方程。金斯特林格针对扬德尔方程只能用于反应转化率较小的情况, 考虑了反应截面随反应进程而变化这一事实, 认为实际反应一经开始, 反应产物层是一个厚度逐渐增加的球壳而不是一个平面。为此, 金斯特林格提出了如图 17-27 所示的反应扩散模型。假设反应物之一 A 的熔点低于 B 的熔点, 当反应物 A 和 B 混合均匀后, A 可通过表面扩散或气相扩散而布满整个 B 的表面。反应开始并生成产物层 AB 后, 反应物 A 在产物层中的扩散速率远大于 B 的扩散速率, 且 $AB-B$ 界面

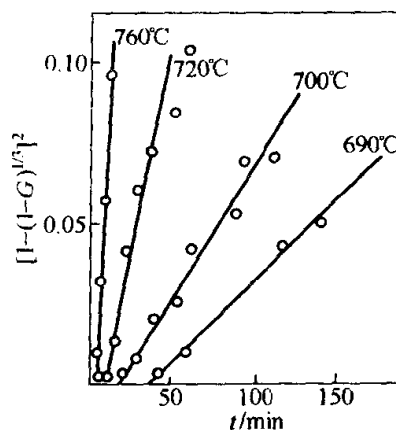


图 17-26 铁酸锌生成动力学的 Jander 模型

上, 由于化学反应速率远大于扩散速率, 扩散到该处的反应物 A 可迅速与 B 反应生成 AB , 因而 $AB-B$ 界面上 A 的浓度可恒为零。但在整个反应过程中, 反应物 A 与生成物 AB 的 $A-AB$ 界面上, 扩散相 A 的浓度则恒为 c_0 。根据该反应模型, 金斯特林格推出反应层厚度随时间变化关系为:

$$x^2 \left(1 - \frac{2x}{3R} \right) = 2K_0 t \quad (17-83)$$

将球形颗粒转化率关系式(17-72)代入上式, 经整理即可得出以转化率 G 表示的金斯特林格动力学方程:

$$F_K(G) = 1 - \frac{2G}{3} - (1 - G)^{2/3} = 2D\mu \frac{tC_0}{R_0^2 n\rho} = K_K t \quad (17-84)$$

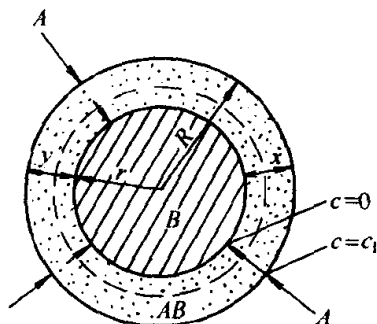


图 17-27 金斯特林格固相反应模型

式中, K_K 称为金斯特林格动力学方程速率常数。

大量实验研究表明, 金斯特林格方程比扬德尔方程能适应于更大的反应程度。例如, 碳酸钠与二氧化硅在 820°C 下的固相反应, 测定不同反应时间的二氧化硅的转化率 G 得如表 17-2 所示的数据。根据金斯特林格方程拟合实验结果, 在较宽的实验测定范围内 ($G = 0.246 \sim 0.616$), $F_K(G)$ 关于 t 有相当好的线性关系, 其速率常数 K_K 均为 1.83。但若以扬德尔方程处

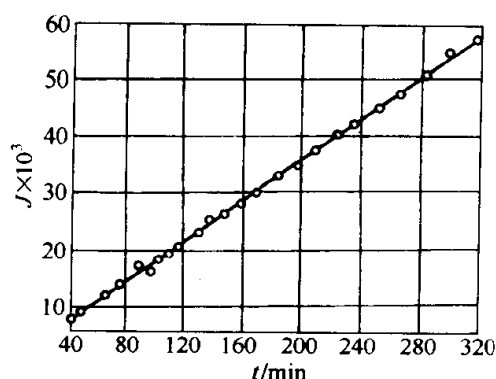


图 17-28 碳酸钠与二氧化硅反应动力学

制作用时，其反应动力学方程式为：

$$F_C(G) = (1 - G)\ln(1 - G) + G = Kt \quad (17-85)$$

然而，应该指出，无论是金斯特林格方程还是扬德尔方程以及圆柱颗粒模型所导出的动力学方程，它们在推导过程中尽管考虑了各自模型中不同的几何细节，但它们均建立在一个稳定扩散过程控制的基本假设上，而实际固相反应中的扩散过程可能会比这一假设更为复杂。因此，这些反应动力学方程并非对所有扩散控制的固相反应都能适用。此外，在上述的动力学方程中没有考虑反应物与生成物密度不同所带来的体积效应。由于实际反应物与生成物密度的差异，扩散相 A 在产物层中扩散的路程并非 $R_0 - r$ ，而是 $r_0 - r$ （此处 $r_0 \neq R_0$ ，为未反应的 B 加上产物层厚度的临时半径），并且 $r_0 - r$ 随着反应的进行而增大。为此，卡特（Carter）对金斯特林格方程进行了修正，得到如下卡特方程：

$$F_{CA}(G) = [1 + (Z - 1)G]^{2/3} + (Z - 1)(1 - G)^{2/3} = Z + 2(1 - Z)Kt \quad (17-86)$$

式中，Z 为消耗单位体积 B 组分所生成产物 C 组分的体积。

卡特将该方程用于镍球氧化过程的动力学数据处理，发现该方程与实验数据从 $G = 0 \sim 100\%$ 均符合得很好，如图 17-29 所示。H.O.Schmalryeel 也

理实验结果， $F_J(G)$ 与 t 的线性关系较差，速率常数 K_J 从 1.81 偏离到 2.25。图 17-28 给出了实验与拟合的图线。

由上述反应动力学分析可以看到，反应截面在反应过程中随时间变化的情况不同，反应所表现出的动力学行为也不同。因此，不同形状颗粒的反应物必然对应着不同的动力学方程。例如，对于半径为 R 的圆柱状颗粒，当反应物沿圆柱表面形成的产物层扩散的过程起控

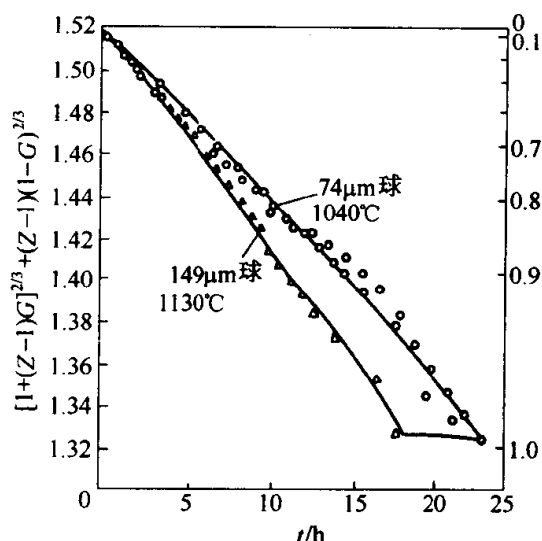


图 17-29 空气中镍球氧化动力学卡特模型

在 ZnO 与 Al_2O_3 反应生成 ZnAl_2O_4 实验中, 证实卡特动力学方程一直到反应度 100% 时仍然适用。

表 17-2 二氧化硅-碳酸钠反应动力学数据 ($R_0 = 36\mu\text{m}$, $T = 820^\circ\text{C}$)

时间/min	SiO_2 反应度 G	$K_K \times 10^4$	$K_J \times 10^4$
41.5	0.2458	1.83	1.81
49.0	0.2666	1.83	1.96
77.0	0.3280	1.83	2.00
99.5	0.3686	1.83	2.02
168.0	0.4640	1.83	2.10
193.0	0.4920	1.83	2.12
222.0	0.5196	1.83	2.14
263.5	0.5600	1.83	2.18
296.0	0.5876	1.83	2.20
312.0	0.6061	1.83	2.24
332.0	0.6156	1.83	2.25

17.2.3 影响固相反应的因素

固相反应过程涉及相界面的化学反应和相内部或外部的物质输运等若干环节, 因而除反应物的化学组成、特性和结构状态以及温度、压力等因素外, 其他可能影响晶格活化, 促进物质内外输运作用的所有因素均会对反应起影响作用。

(1) 化学组成与结构的影响 反应物化学组成与结构是影响固相反应的内因, 是决定反应方向和反应速率的重要因素。从热力学角度看, 在一定温度、压力条件, 反应可能进行的方向是自由能减小 ($\Delta G < 0$) 的方向。而且 ΔG 的负值越大, 反应的热力学推动力也越大。从结构的观点看, 反应物的结构状态-质点间的化学键性质以及各种缺陷的多寡与分布都将对反应速率产生影响。事实表明, 同组成反应物的结晶状态、晶型由于其热历史的不同会出现很大的差别, 从而影响到这种物质的反应活性。例如, 用氧化铝和氧化钴合成钴铝尖晶石 ($\text{Al}_2\text{O}_3 + \text{CoO} = \text{CoAl}_2\text{O}_4$) 的反应中, 若分别采用轻烧 Al_2O_3 和在较高温度下过烧的 Al_2O_3 为原料, 其反应速率可相差近 10 倍。研究表明, 轻烧 Al_2O_3 是由于 $\gamma\text{-Al}_2\text{O}_3 \rightarrow \alpha\text{-Al}_2\text{O}_3$ 转变, 而大大提高了 Al_2O_3 的反应活性。即利用了相转变温度附近, 质点可动性显著增加, 晶格松懈、结构内部缺陷增多等特点, 提高了反应和扩散的能力。因此, 在生产实践中为使反应达到高的生产效率, 如何有效地利用多晶转变、热分解和脱水反应等过程来诱导晶格活化效应是选择反应原料和设计反应工艺条件过程中应予以重视和考虑的。

此外, 在同一反应系统中, 固相反应速度还与各反应物的数量比例有关。

如果颗粒尺寸相同的 A 和 B 反应形成产物 AB, 则改变 A 与 B 的比例在一定范围内会影响到反应物表面积和反应截面积的大小, 从而影响到反应速率。增加反应混合物中“遮盖”物的含量, 有利于反应物间的接触机会和增加反应截面。

(2) 颗粒尺寸分布的影响 反应物颗粒尺寸对反应速率的影响, 首先在扬德尔方程和金斯特林格方程中明确地得到反映。反应速率常数 K 值是反

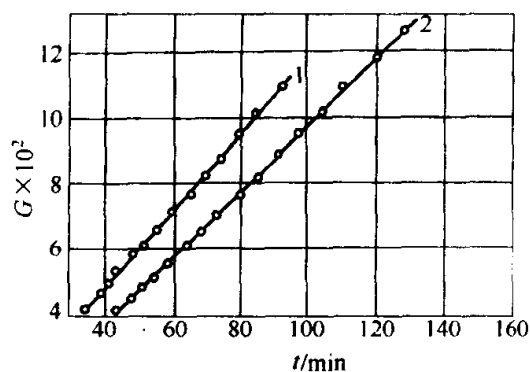


图 17-30 碳酸钠与氧化钼
固相反应动力学

$\text{MoO}_3 : \text{CaCO}_3 = 1:1, r(\text{MoO}_3) = 36\text{nm}$

1— $r(\text{CaCO}_3) = 130\mu\text{m}, T = 600^\circ\text{C}$;

2— $r(\text{CaCO}_3) = 135\mu\text{m}, T = 600^\circ\text{C}$

比于颗粒半径的平方。因此, 在其他条件不变的情况下, 反应速率受到颗粒尺寸大小的强烈影响。图 17-30 表示出不同颗粒尺寸对 CaCO_3 和 MoO_3 在 600°C 温度下反应生成 CaMoO_4 的影响, 比较图中曲线 1 和 2 可以看出颗粒尺寸的微小差别对反应速率的显著影响。

颗粒尺寸大小对反应速率影响的另一方面是通过改变反应界面和扩散截面以及改变颗粒表面结构来达到的。颗粒尺寸越小, 反应体系比表面积越大, 反应界面和扩散截面也相应增大, 因而反应速率增加。同时根据

威尔表面学说, 随着颗粒尺寸的减小, 反应物键强分布曲线宽化并变得平坦, 故而弱键比例增加有助于促进扩散和反应的加快。

应该指出, 由于反应物颗粒尺寸的不同, 同一反应体系的反应机理也可能发生变化而属于不同动力学范围控制。例如前面提及的 CaCO_3 和 MoO_3 的固相反应, 当取等分子比并在较高温度下反应时, 若 CaCO_3 颗粒尺寸大于 MoO_3 则反应表现为由扩散控制, 反应速率随 CaCO_3 颗粒尺寸的减小而加速; 倘若 CaCO_3 颗粒尺寸减小到小于 MoO_3 并且体系存在过量 CaCO_3 时, 反应转变为有 MoO_3 的升华过程所控制, 并随 MoO_3 颗粒尺寸的减小而加强。图 17-31 给出了 CaCO_3 和 MoO_3 反应受 MoO_3 升华过程所控制的动力学情况, 其动力学行为符合由布特尼可夫和金斯特林格推导的升华控制动力学方程:

$$F(G) = 1 - (1 - G)^{2/3} = Kt \quad (17-87)$$

反应物颗粒尺寸大小的分布是影响反应速率的又一因素。理论分析表明, 由于颗粒尺寸大小以平方关系影响着反应的速率, 在同一最可几尺寸条件下, 颗粒尺寸分布越集中对反应速率越是有利。因此, 在涉及固相反应的生产科研中, 设法缩小颗粒尺寸分布, 以避免少量较大尺寸的颗粒存在而显著延缓

反应进程是十分有意义的。

(3) 温度、压力与气氛的影响 温度是影响固相反应速度的重要外部条件之一。一般可以认为升高温度均有利于反应进行。这是因为温度升高, 固体结构中质点热振动动能增加, 原子的扩散和反应能力均得到提高的原因所致。研究表明, 无论对于化学反应或扩散, 其速率常数 $K = A \exp [-\Delta G_r/RT]$ 或扩散系数 $D = D_0 \exp [-Q/RT]$ 均随温度的升高而提高。而且, 由于扩散活化能 Q 通常比反应活化能 ΔG_r 小, 而使温度的变化对化学反应的影响远大于对扩散的影响。

压力是影响固相反应的另一外部因素。对应纯固相反应, 压力的提高可显著地改善粉料颗粒之间的接触状态, 如缩短颗粒之间距离, 增加接触面积, 从而提高反应速率。对于有液相或气相参加的固相反应, 扩散主要不是通过固相颗粒直接接触进行的。因此, 提高压力并不表现出积极的作用, 有时还适得其反。如粘土矿物脱水反应和伴有气相产物的热分解反应以及某些有升华控制的固相反应等等, 增加压力会使反应速率下降。表 17-3 所列数据表明随着水蒸气压力的增高, 高岭土的脱水温度和活化能明显提高, 脱水速率降低。

表 17-3 不同水蒸气压力下高岭土的脱水活化能

水蒸气压力 p/Pa	温度 $T/^\circ\text{C}$	活化能 $\Delta G_r/\text{kJ}\cdot\text{mol}^{-1}$
<0.10	390~450	214
613	435~475	352
1867	450~480	377
6265	470~495	469

此外, 气相对固相反应也有重要影响。它可以通过改变固相颗粒表面吸附特性而影响表面反应活性。对于一系列能形成非化学计量的化合物如 ZnO 、 CuO 等, 气氛可直接影响晶体表面缺陷的浓度、扩散机构和扩散速度。

(4) 矿化剂及其他影响因素 在固相反应体系中加入少量非反应物物质或由于某些可能存在于原料中的杂质, 则常会对反应产生特殊的作用。这些物质常被称之为矿化剂, 它们在反应过程中不与反应物或反应产物直接起化

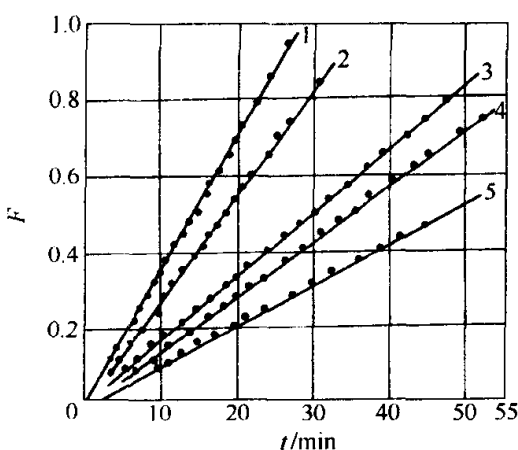


图 17-31 碳酸钠与氧化钼固

相反应 (升华控制)

$\text{CaCO}_3:\text{MoO}_3 = 15$; $r(\text{CaCO}_3) = 30\mu\text{m}$

1— $r(\text{MoO}_3) = 52\mu\text{m}$; 2— $r(\text{MoO}_3) = 64\mu\text{m}$;

3— $r(\text{MoO}_3) = 119\mu\text{m}$; 4— $r(\text{MoO}_3) = 130\mu\text{m}$;

5— $r(\text{MoO}_3) = 153\mu\text{m}$

学反应，但它们以不同的方式和程度影响着反应的某些环节。实验表明，矿化剂可以产生如下作用：①改变反应机构降低反应活化能；②影响晶核的生长速率；③降低体系的低共熔点，改善液相性质等。例如在 Na_2CO_3 和 Fe_2O_3 反应体系中加入微量的 NaCl ，可使反应转化率提高 1.5~1.6 倍之多。而且当颗粒尺寸越大，这种矿化效果越明显。又例如在硅砖中加入 1%~3% 的 $\text{Fe}_2\text{O}_3 + \text{Ca}(\text{OH})_2$ 作为矿化剂，能使其大部分 α -石英不断熔解析出 α -磷石英，从而促使 α -石英向磷石英的转化，关于矿化剂的一般矿化机理是复杂多样的，同种矿化剂可因反应体系的不同而表现出完全不同的矿化作用。但无论具体矿化细节的千差万别，矿化剂总是以某种方式参与到固相反应过程中去的。

以上从物理化学角度对影响固相反应诸因素进行了分析讨论，但必须指出，实际生产科研过程中遇到的各种影响因素可能会更多更复杂。对于工业性的固相反应，除上述物理化学方面的影响因素外，还存在许多工程方面的因素。例如水泥工业中的碳酸钙分解速率除受物理化学方面基本规律的影响，还与工程上换热传质速率有关。在相同的温度条件下，普通旋窑中的分解率要低于窑外分解炉中的。这是因为窑外分解炉中处于悬浮状态的碳酸钙颗粒在传质换热条件上比普通旋窑中好得多。因此，从反应工程的角度考虑传质传热效率对固相反应的影响是具有同样重要性的。尤其是对于硅酸盐材料，生产通常要求高温条件，此时传热速率对反应的进行将具有更强的影响。例如，把石英砂压成直径为 50mm 的小球，约以 $8^\circ\text{C}/\text{mm}$ 的速度进行加热使之进行 $\beta \rightarrow \alpha$ 相变，需要约 75min 完成。而在同样加热速度下，用同样的石英单晶球做实验，则完成相变所需时间为 13min。产生这种差异的原因除两者的传热系数不同以外[单晶约为 $5.23\text{W}/(\text{m}^2 \cdot \text{K})$ ，而石英砂球的导热系数为 $0.58\text{W}/(\text{m}^2 \cdot \text{K})$]，其两者传热方式也不一样。在石英单晶中，辐射传热使固相反应可以在相当的厚度范围内甚至在整个体积中同时进行，而石英砂球只能依靠传导机构由外及内地传热，相应的固相反应也只能由外及内地逐步深化。显然，前者的固相反应速度必然远大于后者。

17.3 固态烧结

烧结是一个古老的材料制备工艺技术。人类祖先很早就利用烧结来制备各种陶器和简单的耐火材料。现代的烧结技术是粉末冶金、陶瓷、耐火材料、超高温材料等多晶材料制备过程中的重要手段。烧结过程可概括为把由粉体经成型而得到的具有一定外形的坯体（通常含 25%~60% 的气孔）在低于熔点的温度条件下转变为具有一定显微结构的致密烧结体的物理过程（如图 17-32 所示）。烧结过程进行的条件直接影响着烧结体显微结构中晶体、玻璃体和

气孔等相组分的尺寸及其分布，从而也直接影响着烧结材料的有关物理或化学性能。因此，了解粉末坯体烧结过程的现象和机理，了解烧结动力学及其影响烧结的因素对控制和改进材料的性能有着十分重要的实际意义。

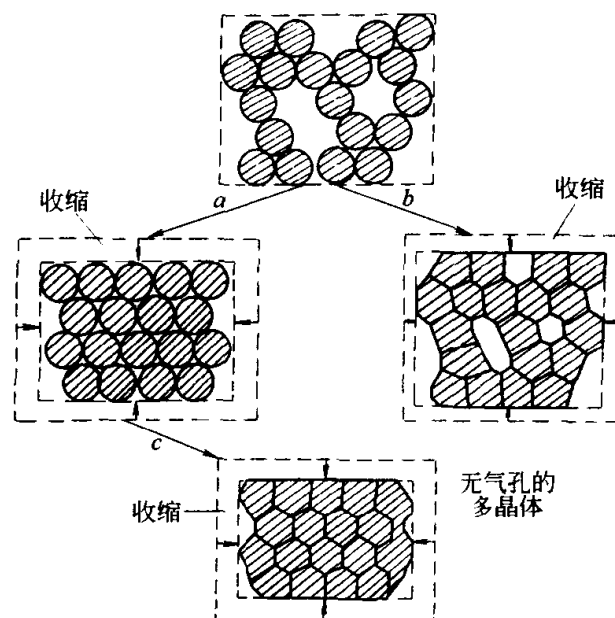


图 17-32 烧结现象示意图

a 颗粒聚积; *b* 颗粒中心拉近; *c* 致密化

17.3.1 固态烧结的初期动力学

固体粉料组成的素坯在开始烧结后，将出现颗粒间接触面的逐渐扩展、颗粒中心距不断减小以及素坯的明显收缩。当素坯的收缩率在 0~5% 范围内时，烧结进程被称之为初期阶段。在这一阶段，固态球形颗粒表面与它们间的颈部（图 17-33）区域之间所存在的化学位梯度将导致物质向颈部发生迁移。例如，若颗粒表面与其颈部区域间有较高的蒸气压差，经由气相而进行的蒸发-凝聚传质过程则很容易发生；若材料体系的蒸气压较低，则物质迁移易通过固相表面、晶格、晶界等方式的扩散而得以进行。

表 17-4 结合图 17-33 给出了各种扩散途径的描述。在素坯的烧结

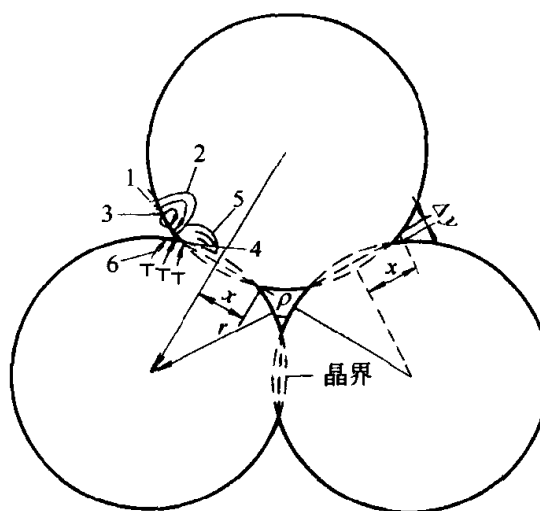


图 17-33 烧结初期物质迁移的可能路径

初期, 所列的传质过程一般均可能或多或少地发生。在这些传质过程中, 传质速率最大的一种将视为初期烧结的主导传质机理。下面以库津斯基 (G.C.Kuczynski) 首先提出的双球几何模型为基础, 对蒸发-凝聚传质机制的烧结初期动力学特点作一简单讨论。

表 17-4 图 17-33 中所示的传质途径

图示号	扩散途径	物质来源	物质抵达部位
1	表面扩散	表面	颈部
2	晶格扩散	表面	颈部
3	蒸发-凝聚	表面	颈部
4	晶界扩散	晶界	颈部
5	晶格扩散	晶界	颈部
6	晶格扩散	位错	颈部

如图 17-34 所示, 颗粒表面具有正的曲率半径, 而两颗粒相接触的颈部则具有负的曲率半径。根据 Kelvin 方程与图 17-34 模型的几何参数, 可得颗粒表面蒸气压 p_0 与颗粒间颈部处蒸气压 p_1 间的关系式:

$$\ln\left(\frac{p_1}{p_0}\right) = \frac{\gamma M}{dRT} \left(\frac{1}{\rho} + \frac{1}{x} \right) \quad (17-88)$$

式中, γ 为表面能; M 为蒸气相的分子量; d 为材料的密度; ρ 与 x 分别为颈部的曲率半径和高度。引入由气体分子运动论 Langmuir 公式推得的单位面积上分子的凝聚速率 $U_m = \alpha \Delta P [M / 2\pi RT]^{1/2}$, 并考虑图 17-34 所示颗粒间颈部的表面积和体积关系, 容易推得当凝聚速率等于颈部体积增长时颈部生长速率与颗粒尺寸、温度、时间以及蒸气压的关系:

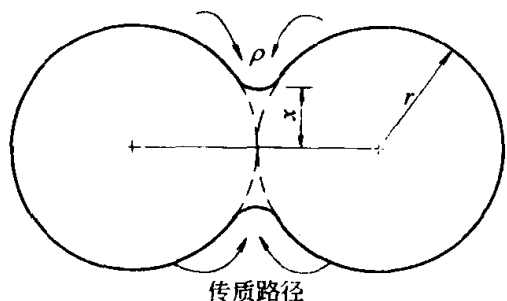


图 17-34 蒸发-凝聚传质机制

$$\frac{x}{r} = \left[\frac{3\sqrt{\pi} r M^{3/2} p_0}{\sqrt{2} (RT)^{2/3} d^2} \right]^{1/3} \left(\frac{t}{r^2} \right)^{1/3} \quad (17-89)$$

该方程得到了 Kingery 等人以具有较高蒸气压的氯化钠球进行烧结实验结果的证实。如图 17-35 所示, 烧结初期的颈部生长 x/r 随时间 t 的 $1/3$ 次方变化, 随着烧结的进一步进行, 颈部的增长很快就趋于停止。此外, 蒸发-凝聚传质烧结机制的另一个特点是烧结过程伴随球形颗粒的椭圆化、气孔形状改变, 而颗粒的中心距不变, 因此不导致素坯的宏观收缩。

然而, 在大多数固体材料中, 由于高温下材料蒸气压低而使烧结所依赖的传质趋于通过固体内质点的扩散过程来进行, 这样一种烧结机制将使颗粒

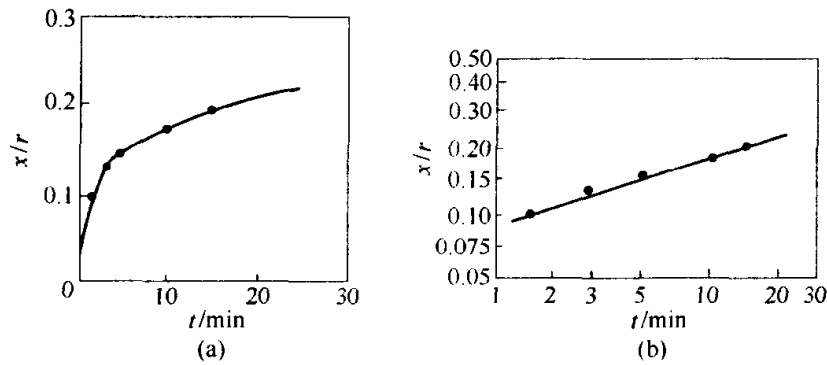


图 17-35 球形 NaCl 颗粒在 750℃ 下烧结颈部生长与时间关系

(a) 普通坐标; (b) 对数坐标

的中心距——即素坯尺寸收缩。考虑如图 17-36 所示双球几何模型，其颈部面积和体积分别为：

$$A = \pi^2 \frac{x^3}{2r}; \quad V = \pi \frac{x^4}{4r} \quad (17-90)$$

由于颗粒颈部具有负的曲率半径，而颗粒表面的曲率半径则为正值，因而在达到热力学平衡时，颗粒颈部区域相对颗粒的其他部位应存在更高的空位浓度。与式 (17-88) 类似，作一级近似可以推得空位浓度差： $\Delta c/c_0 = \gamma a^3/kT\rho$ ，并考虑颈部体积增长的速率由质点的晶格扩散所控制，则有：

$$\frac{dV}{dt} = -D_V \frac{\Delta CA}{\rho} \quad (17-91)$$

此处 a^3 为扩散空位的原子体积； c_0 为颗粒表面处的空位浓度； D_V 为空位的扩散系数，它与自扩散系数 D^* 的关系为： $D^* = D_V c_0$ 。将上述各式予以整理代入式可得：

$$\frac{x}{r} = \left(\frac{40\pi\gamma a^3 D^*}{kT} \right)^{1/5} \left(\frac{t}{r^3} \right)^{1/5} \quad (17-92)$$

随着初期烧结的进行，颗粒中心相互接近，素坯发生收缩而致密化。素坯的体收缩率与线收缩率由如下式表示：

$$\frac{\Delta V}{V} = 3 \left(\frac{\Delta L}{L} \right) = 3 \left(\frac{5\pi\gamma a^3 D^*}{4kT} \right)^{2/5} \left(\frac{t^2}{r^6} \right)^{1/5} \quad (17-93)$$

正如图 17-33 所示，固相材料烧结过程中还包括一些其他一些扩散烧结机制，如表面扩散机制、晶界扩散传质机制等，应用与上述相似的分析方法

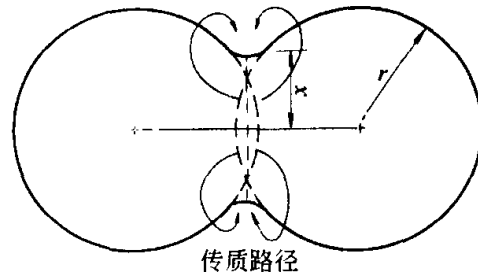


图 17-36 颗粒中心距收缩的双球模型

可分别得到相应的烧结初期动力学方程：

$$\frac{x}{r} = \left[\frac{\gamma a^3 \delta D_s t}{k T r^4} \right]^{1/n} \quad (17-94)$$

$$\frac{x}{r} = \left[\frac{\gamma a^3 W D_b t}{k T r^4} \right]^{1/6} \quad (17-95)$$

上式中 δ 为原子间距； W 为晶界宽度； D_s 为表面扩散系数； D_b 为晶界扩散系数；其余各项意义如同前述。式(17-94)中指数 n 的数值，不同的研究者提出了不同的看法，Kuczynski 和 Rockland 认为 $n = 7$ 。应该指出，在烧结的早期，表面扩散往往是重要和容易发生的。但与蒸发-凝聚类似，表面扩散机制不影响素坯的收缩率。

17.3.2 固态烧结中、后期动力学模型

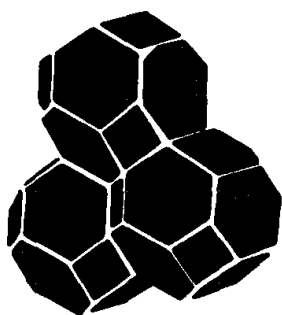


图 17-37 烧结中后期
晶粒的十四面体模型

在初期烧结阶段后，烧结过程使颗粒开始相互粘结，颈部扩大，气孔由不规则的形状逐渐变成由三个颗粒包围的、相互连通的圆柱形管道。同时，晶界开始移动，出现晶粒平均尺寸增大的晶粒生长现象。一般认为，进一步烧结所依赖的传质是以晶界和晶格扩散为主要的。此时，晶粒分布及空间堆积方式等均很复杂，这使烧结在该阶段的动力学描述变得十分困难。Coble 首先提出了一个简单的多面体模型，他认为此时的烧结体是由众多的十四面体堆积而成（如图 17-37 所示）。十四面体顶点是四个晶粒的交汇点，每个边是三个晶粒的交界线，它相当于圆柱形通道，成为进一步烧结的空位源。空位从圆柱形空隙向晶界扩散，而原子作反向扩散使坯体致密化。根据这样一个理想的晶粒空间堆积体，Coble 近似地导出了多晶材料中期烧结致密化或气孔率减小的动力学方程：

$$P = 10 D_v \gamma a^3 \frac{(t_f - t)}{l^3 k T} \quad (17-96)$$

式中， D_v 为原子的晶格扩散系数； γ 为表面能； a^3 为原子空位体积； t_f 为气孔全部消除所需的时间。

当多晶材料进入后期烧结时，气孔已由圆柱形孔道收缩成位于十四面体的二十四个顶点处的孤立气孔。根据相同的结构模型，Coble 导出了烧结体后期气孔率的关系式：

$$P = D_v \gamma a^3 \frac{6\pi(t_f - t)}{\sqrt{2} l^3 k T} \quad (17-97)$$

比较式(17-96)和式(17-97)，可以看出，固态材料在烧结中后期气孔收缩速

率关系并无质的差别，当温度和晶粒尺寸不变时，烧结体气孔率随时间而线性地减小。Coble 的这一推导结果基本上得到了 Al_2O_3 烧结实验结果的验证。

17.3.3 烧结过程中的晶粒生长与二次再结晶

晶粒生长与二次再结晶是伴随烧结中、后期传质过程同时发生的现象。晶粒生长是无应变的多晶材料在热处理时，晶粒尺寸在不改变其分布的情况下连续长大的过程，二次再结晶则是指在烧结过程中少数大晶粒在消耗小晶粒的同时出现异常长大的过程。在固态材料烧结过程中，合理地控制晶粒生长十分重要，因为晶粒生长是否合理密切地关系到最终产品的致密化程度及其性能。

晶粒生长不是小晶粒的简单粘结，它是晶界移动的结果。如图 17-38 所示，弯曲晶界两侧的原子具有不同的自由能。原子 A 的自由能高于原子 B 的自由能，因而原子 A 可以不断越过晶界而进入原子 B 所在的晶粒。此时，两晶粒间的晶界将表现为向原子 A 所在晶粒的曲率中心发生推移，直至晶界变得平直、晶界两侧的原子自由能相等为止。显然，这种表现为晶界移动的结果是原子 B 所在的晶粒长大，同时原子 A 所在的晶粒缩小，晶粒长大的速率决定于晶界移动的速率。

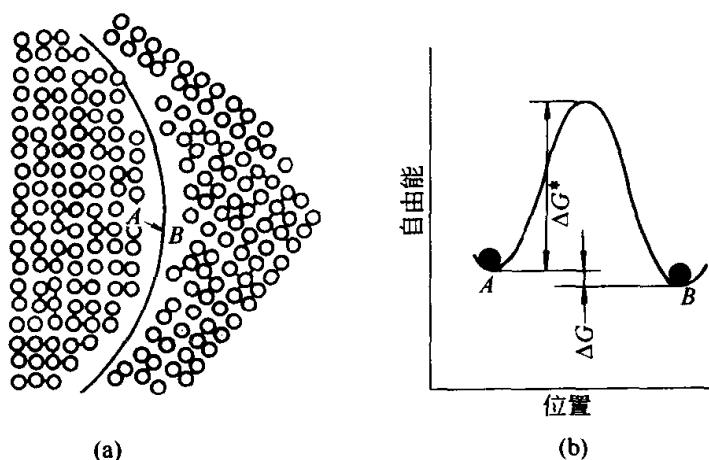


图 17-38 晶界结构与原子跃迁的能量变化

由简单的热力学概念可知，图 17-38 中晶粒 A、B 间晶界移动的推动力源于晶粒表面曲率不同而产生的自由能差：

$$\Delta G = \Delta p V = V \gamma \left(\frac{1}{\rho_1} + \frac{1}{\rho_2} \right) \quad (17-98)$$

考虑晶粒 A、B 上原子双向跃迁的有效频率分别为：

$$A \rightarrow B \quad f_{AB} = \frac{RT}{Nh} \exp \left(\frac{-\Delta G^*}{RT} \right) \quad (17-99)$$

$$B \rightarrow A \quad f_{BA} = \frac{RT}{Nh} \exp \left[-\frac{(\Delta G^* + \Delta G)}{RT} \right] \quad (17-100)$$

可推得晶界的移动速率近似为：

$$v = \lambda (f_{AB} - f_{BA}) = \lambda \gamma V \frac{1}{Nh} \left(\frac{1}{\rho_1} + \frac{1}{\rho_2} \right) \exp \left(\frac{\Delta G^*}{RT} \right) \quad (17-101)$$

式中， N 为阿佛加得罗常数； h 为普郎克常数； γ 为表面能； ρ_1, ρ_2 分别为曲面的主曲率半径； λ 为每次跃迁距离。

由上式可以看出，晶粒生长速率或晶界的移动速率随温度成指数规律增加，而与晶界面曲率半径成反比。温度越高、曲率半径越小，晶界向其曲率中心移动的速率越大，因而晶粒长大的平均速度与晶粒的直径成反比 $dD/dt = K/D$ ，即有：

$$D^2 - D_0^2 = Kt \quad (17-102)$$

式中， D_0 为晶粒的初始直径； D 为时间 t 时晶粒的直径； K 为比例常数。晶粒生长到后期时， $D \gg D_0$ ，上式可近似写成：

$$D = Kt^{1/2} \quad (17-103)$$

一些氧化物材料的晶粒生长实验表明，晶粒 D 与时间指数常在 $1/2 \sim 1/3$ 之间。这主要是晶界移动时遇到了第二相夹杂物如杂质原子或气孔等而受到阻碍，由此限制了晶粒的长大。图 17-39 描述了晶界移动时遇到夹杂物的三种情况。晶界遇到并企图通过夹杂物时，将损失正比于夹杂物横截面积大小的能量，因而使晶界继续推进的能力减小，甚至减弱至界面变得平直，使晶粒生长逐渐停止。但当晶界推进力较大时，损失能量后的晶界也可能越过杂质或带着气孔一起继续前进，如图 17-39 中的 (b)、(c) 两种情况。应该指出，在烧结后期，严格控制温度，使体系中保持图中 (b) 种状态，即 $V_b = V_p$ ，是实现迅速排除气孔达到烧结体致密化的有效措施。过高的温度将使 $V_b \gg V_p$ ，导致晶界越过气孔并使气孔包入晶粒的内部，这无疑对烧结体的致密化将产生十分不利的影响。

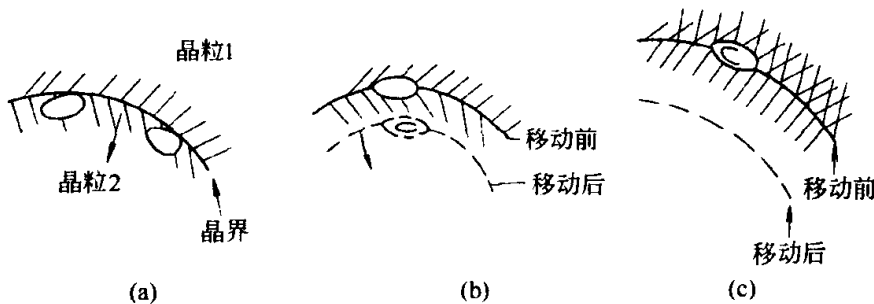


图 17-39 晶界移动碰到气孔的三种情况

(a) $V_b = 0$; (b) $V_b = V_p$; (c) $V_b > V_p$

在晶粒正常生长过程中, 由于夹杂物对晶界移动的牵制可使晶粒长大难以超过某一极限尺寸。曾讷 (Zener) 曾估计这一极限尺寸为:

$$D_l \propto \frac{d}{f} \quad (17-104)$$

式中, d 为夹杂物或气孔的平均直径; f 为夹杂物或气孔的体积分数。在烧结初期, 坯体内有许多小而多的气孔, 故 f 值相当大, 晶粒的初始尺寸 D_0 总是大于 D_l , 所以晶粒无法长大。随着烧结的进行, 小气孔不断沿晶界聚集或排除, d 由小增大, f 由大变小, 因而出现 $D_l > D_0$, 使晶粒生长得以开始。进入烧结后期, 气孔尺寸与体积分数都在进一步减小, 其趋于恒定的比值使得烧结体中晶粒的长大逐渐受到抑制直至最后停止。

与晶粒正常生长不同, 在烧结的中后期如在均匀基相中存在少数如图 17-40 所示的多边界的大晶粒, 便会出现以这些大晶粒为核心的异常迅速的所谓二次再结晶。二次再结晶的晶粒往往不仅有较多的边界, 同时晶界曲率也较大, 以至于晶界可以越过气孔或夹杂物快速向邻近小晶粒中心推进。如此不断吞并周围的小晶粒, 直至长大到与邻近大晶粒接触为止。在绝大多数情况下, 二次再结晶的出现首先很不利于

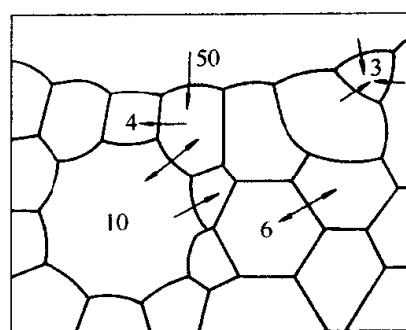


图 17-40 多晶坯体中晶粒生长示意图

烧结体的致密化, 因为晶界的快速移动使大量气孔包裹在晶粒内部; 同时大晶粒的异常快速长大而引入的大量结构缺陷对多晶材料的理化性能往往是有危害的。因而工艺上需采取适当措施防止其发生。仅在某些特殊情况下, 如要求材料内部晶粒分布具有择优取向 (硬磁铁氧体 $\text{BaFe}_{12}\text{O}_{14}$ 是一例), 才有意识地利用二次再结晶现象。从工艺控制角度考虑, 造成二次再结晶的原因主要是原始粉料粒度不均匀、烧结温度偏高和烧结速率太快, 此外还有坯体成型时压力不均匀, 局部有不均匀液相等。因此, 制备尽可能均匀的粉料、均匀成型压力、严格控制烧结温度与时间是避免出现二次再结晶的基本措施。此外, 有选择地引入适当的添加剂以抑制晶界的快速移动是防止二次再结晶的最好办法。如少量 MgO 加入 Al_2O_3 中可烧结出具有理论密度的制品, 还有 Y_2O_3 加入 ThO_2 中或 ThO_2 加入 CaO 中等等。

17.3.4 液相烧结和热压烧结

坯体在整个烧结过程中不出现任何液相的烧结是纯固相烧结。然而, 不少金属材料与无机非金属材料在烧结时常出现液相, 这类烧结过程称之为液相烧结。液相烧结有三个基本的过程: ①在颗粒间的液相可以产生毛细管

力, 从而引起颗粒间的压力并使颗粒易于滑动, 导致颗粒重排和改善颗粒的堆积结构; ②毛细管力将引起固态颗粒的溶解和再沉淀, 其结果是使颗粒在接触部位变得扁平、坯体发生收缩。溶解-沉淀过程是和固态颗粒的 Ostwald 生长密切相关的; ③由于液相的存在, 溶解-沉淀和流动传质将使烧结致密化速率比纯固相烧结大大提高。此外, 液相烧结的具体速率与液相数量、液相性质 (粘度、表面张力等)、液相与固相的润湿情况、固相在液相中的溶解度等等因素有着密切关系。因此, 定量地研究液相烧结较研究纯固相烧结往往更加复杂和困难。

热压烧结是在普通无压烧结的基础之上发展起来的一种特殊烧结技术。它是在烧结的同时在坯体上施加一定的压力, 以补偿烧结体内气孔中逐渐增大的气压抵消了作为烧结推动力的界面能, 使烧结得以继续进行从而达到制备高致密度材料的目的。

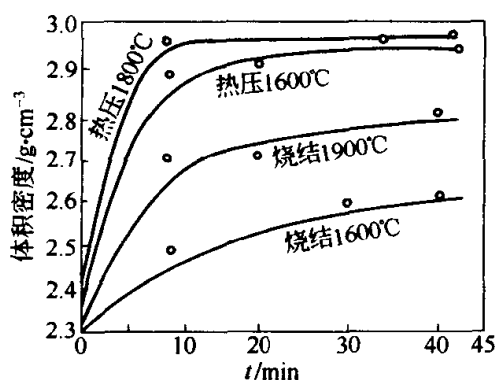


图 17-41 BeO 热压烧结与普通烧结致密化比较

BeO 的热压烧结与普通烧结对坯体密度的影响如图 17-41 所示。采用热压烧结技术可使制品达到理论密度的 99% 甚至 100%。热压烧结尤其适应于共价键材料, 如碳化物、硼化物、氮化物等。由于它们在烧结温度下有高的分解压和低的原子迁移率, 因此无压烧结难以使其致密化。例如 BN 粉粒在 200MPa 的等静压下成型后, 在 2500℃ 下无压烧结后相对密度为 0.66, 而采用 25MPa 压力在 1700℃ 下热压烧结可得

相对密度为 0.97 的 BN 材料。由此可见, 热压烧结对提高材料的致密度和降低烧结温度均具有显著的效果。

参 考 文 献

- 1 陆佩文等编. 无机材料科学基础. 武汉: 武汉工业大学出版社, 1996
- 2 浙江大学等. 硅酸盐物理化学. 北京: 中国建筑工业出版社, 1980
- 3 P. 哈森物理冶金学. 北京: 科学出版社, 1984
- 4 崔国文等. 缺陷扩散与烧结. 北京: 清华大学出版社, 1990
- 5 Mrowec S. Defects and Diffusion in Solids. An Introduction, Amsterdam: Elsevier, 1980
- 6 Mitchell J. W. et. al. Reactivity of Solids. New York: John Wiley & Sons, Inc., 1969
- 7 Kuczynski G. C. Sintering Processes. New York: Plenum, 1980

第 18 章 相变的基本原理

撰稿人 曾燕伟 冯 端 刘治国

物相作为物质系统中具有相同化学组成、聚集状态及相同物理、化学性质的均匀物质部分的概念最早是由吉布斯 (Gibbs) 于 19 世纪初首先提出的。在一定的条件下 (温度、压强等), 物质将以一种与外界条件相适应的聚集状态或结构形式存在着, 这种形式就是相。在某种意义上, 它和该物相的化学组成定义了其全部的物理和化学性质。相变则是指在外界条件发生变化的过程中物相于某一特定的条件下 (临界值) 时发生突变, 表现为: ①从一种结构变化为另一种结构, 如气相、液相和固相间的相互转变, 或固相中不同晶体结构或原子、离子聚集状态之间的转变; ②化学成分的不连续变化, 例如均匀溶液的脱溶沉淀或固溶体的脱溶分解等; ③更深层次序结构的变化并引起物理性质的突变, 例如顺磁体-铁磁体转变、顺电体-铁电体转变、正常导体-超导体转变等等。这些相变的发生往往伴随某种长程序结构的出现或消失。金属-非金属转变、液态-玻璃态间的转变等, 则对应于构成物相的某一种粒子 (原子或电子) 在两种明显不同状态 (如扩展态和局域态) 之间的转变。实际材料中所发生的相变形式可以是上述中的一种, 也可以是它们之间的复合, 如脱溶沉淀往往是结构和成分变化同时发生, 铁电相变则总是和结构相变耦合在一起。

相变现象的发现与研究是随着 19 世纪后半期采矿和冶金工业的出现和人们使用显微镜鉴定和研究矿石、钢铁等材料的细微结构而开始的。人们发现了相同化学成分的晶态物质在结构上的多形性及其相互转变的现象, 认识到了钢铁材料在热处理前后存在显微组织上具有差异的物相 (如奥氏体、马氏体、珠光体等) 及其它们之间在一定条件下的相互转变。随着对相变现象及其规律认识的逐步加深, 人们开始提出了相变的有关理论, 如经典的成核-生长理论、界面微观形貌控制的生长理论、固液相变界面失稳理论以及近代凝聚态物理学中的各种有关理论等等。纵观相变现象研究的发展可以看出, 相变现象以其在自然界的普遍存在性和多样性长期吸引着包括地质学、材料科学、物理学、化学等众多学科科学家的研究兴趣。相变现象研究与发展不仅使人们加深了对大量与相变有关的现象在理论上的认识和理解, 更重要的是, 它促进了构筑现代物质文明基础的现代科学技术、尤其是材料科学技

术的迅速发展。相变过程基本规律的研究、学习和掌握有助于人们合理、科学地优化材料制备的工艺过程，并对材料性能进行能动地设计和剪裁，其重要性和意义是不言自明的。

18.1 相变的基本结构特征

正如前面所述，物相作为一定化学组成和结构的物质的一种存在形式在发生变化时总是伴随着某一层次结构形式的改变。因此研究物质相变时所伴随的结构形式的改变，可以得到有关相变鲜明的物理图像。

18.1.1 重构型相变和位移型相变

M.J.Buerger 在研究大量晶态物质相变时结构变化特征的基础上，提出了涉及晶体结构变化的相变可分成重构型相变和位移型相变两种基本类型的观点。借助于示意图 18-1 可以看到，重构型相变表现为在相变过程中物相的结构单元间发生化学键的断裂和重组，并形成一种崭新的结构，其形式与母相在晶体学上没有明确的位向关系。位移型相变与此完全不同，在相变过程中不涉及到母相晶体结构中化学键的断裂和重建，而往往只涉及到原子或离子位置的微小位移，或其键角的微小转动。说明重构型相变和位移型相变的典型例子是碳的石墨-金刚石转变和石英变体的转变。石墨和金刚石同是由碳元素所组成，石墨具有层状结构，其特点为层内每个碳原子与周围三个碳原子形成共价键，而层间则由脆弱的分子键相连。但在高温高压下石墨可转变成结构完全不同的金刚石相，结构中每个碳原子均由共价键与其配位的四个碳原子相连，从而使金刚石具有完全不同于石墨的力学和电学性能。石英是自然界中广泛存在的矿物原料，其众多变体间的转变既有重构性相变又有位移型相变。如图 18-2 所示的横向相变过程，石英变体中 α -石英、 α -鳞石英和 α -方石英间的转化涉及到结构中化学键的断裂和重建，因为构成它们的硅氧四面体有着完全不同的连接方式，其转变过程具有势垒高、动力学速率低和相变潜热大等特点。相反，图中所示的纵向过程，石英、鳞石英和方石英本身 α 、 β 或 γ 变体间的转变在结构上仅表现为 Si—O—Si 键角的

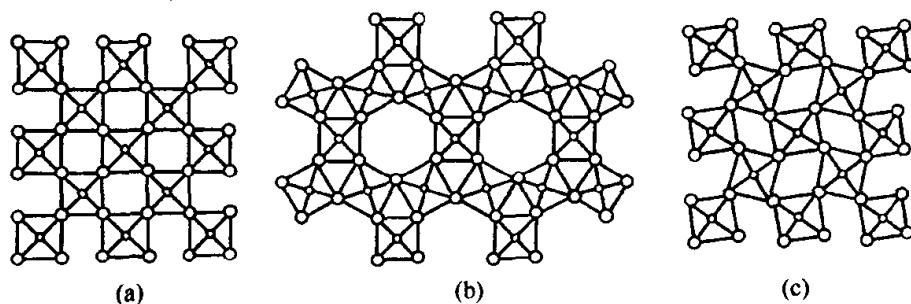


图 18-1 重构型和位移型相变结构变化示意图

(a) 高温相；(b) 重构机制相变；(c) 位移机制相变

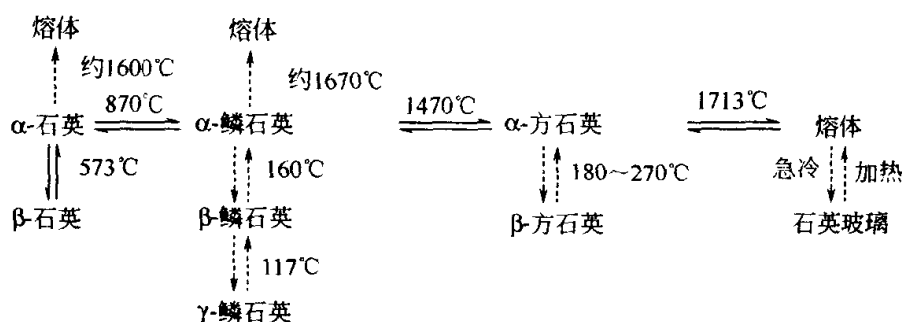


图 18-2 不同温度下石英变体间相变关系

微小变化，并在动力学上经历的势垒低、相变潜热小，因而有着较快的相变速度，以至于有时无法用淬火的方法将高温相保留到室温。

重构型相变不仅涉及于大量晶态材料不同晶相间的转化，实际上大量物质的气相-液相-固相间的相互转变也应归属于这一类型的相变。一般来说，一种物质的液相与固相在结构上均具有显著的差异。液相原子呈无规则排列，每个原子周围均可被其他原子以等几率的方式所配位，从而使液相的内部结构具有很高的对称性和不规则性。相反，晶相原子的排列具有严格的规则性和周期性。因而，液-固相间的转变总涉及到原子间键的断裂和重组，并伴随着较大的热效应。至于位移型相变虽不及重构型相变那样广泛地存在，但由于它的原子位移图像明确，而且又和一些重要的物理性质（如铁电性和反铁电性）的变化耦合在一起，已成为现代物理学和材料科学有关分支的研究热点，其中钙钛矿 ABO_3 结构的氧化物相变是最为令人感兴趣的。

图 18-3 中所示为一立方对称的钙钛矿结构。A 离子位于立方体的中心，B 离子位于立方体的 8 个顶角，而氧离子位于立方体棱边中位，若以 r_A 、 r_B 和 r_O 分别代表钙钛矿结构中 A、B 和 O 离子的离子半径，则它们之间满足下面关系时，钙钛矿结构将可以得到稳定存在：

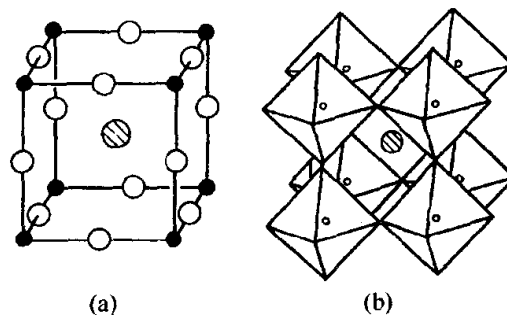


图 18-3 钙钛矿晶体结构示意图

$$r_A + r_O = t\sqrt{2} (r_B + r_O) \quad (18-1)$$

其中 t 为容差因子，其值为 $0.77 \sim 1.10$ 。由于钙钛矿结构中存在容差因子，且 A、B 离子的价数不局限于 $2+4$ ，所以具有钙钛矿结构的晶体种类十分丰富。除了大量的 $(2+4)$ 氧化物外，常见的还有 $(1+5)$ 氧化物 $NaNbO_3$ 、

KNbO_3 、 NaWO_3 ，(3+3) 氧化物 YAlO_3 、 LaAlO_3 、 LaCrO_3 以及 (1+2) 的氟化物 KNiF_3 、 KZnF_3 等等。

钙钛矿型结构在高温时具有立方对称。降低温度并在通过某一特定温度时， B 离子可沿一个轴向、两个轴向或沿对角线方向发生微小的位移，从而使原来的立方对称性遭到破坏而分别降至四方、正交和三方菱面体对称结构。正是由于这种离子位移相变的发生，使一些钙钛矿结构的离子晶体内部产生自发偶极矩，从而使其成为铁电体或反铁电体并具有介电和压电性能。

18.1.2 马氏体型相变

马氏体(Martensite)是钢高温淬火过程中通过相变而得到的一种高硬度产物。为纪念德国冶金学家 A. Martens 在钢铁显微结构研究方面的贡献，人们将该过程得到的高硬度产物相命名为马氏体，其相变过程称为马氏体相变。马氏体相变在许多金属、合金固溶体和化合物中都可观察到，本质上属于以晶格畸变为主、无成分变化、无扩散的位移型相变，其特征为发生于晶体中某一部分的极其迅速的剪切畸变。这种相变在热力学和动力学上都有相当显著的特点，如其相转变无特定的温度点、转变动力学速率可高达声速等，但其结晶学特点更是尤为鲜明。

如图 18-4 所示，(a) 为一四方结构的母相奥氏体块，(b) 为从母相中形成马氏体单晶片后的示意图。其中 $A'_1B'_1C'_1D'_1-A_2B_2C_2D_2$ 是母相奥氏体通过切变转变成的马氏体。相变使原来母相中的直线 $PQRS$ 变成了由线段 PQ 、 QR' 和 $R'S'$ 构成的折线。但应注意到，相变前后连接母相与马氏体的平面 $ABCD$ 既不发生扭曲也不发生旋转，通常称之为习性平面。同时， $A'_1B'_1$ 和 A_2B_2 两条棱的直线性表明了马氏体相变宏观上剪切的均匀整齐性。因此，马氏体相变可以概括为沿母相习性平面生长、形成与母相保持着确定的切变共格结晶学关系的新相的相变过程。这一点是马氏体相变最重要的结晶学特点。在金属中常见的马氏体相变包括以下几种：①面心立方-体心立方，如 Fe-C，Fe-Ni 中等合金的马氏体相变，其母相与马氏体的取向关系有 $\{111\}_P // \{011\}_M$ ， $\langle 011 \rangle_P // \langle 111 \rangle_M$ ；②面心立方-面心四方，

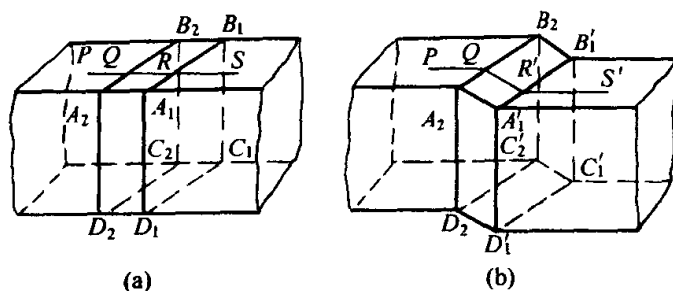


图 18-4 奥氏体-马氏体相变结晶学关系示意图

如 In-Tl 合金系统, 习性面为 $\{101\}_P$; ③体心立方-正交, 如 Au-Cd 合金系统, 取向关系近似为 $\{011\}_P // \{001\}_M$, $\langle 11\bar{1} \rangle_P // \langle 110 \rangle_M$, 习性面为 $\{133\}_P$; ④体心立方-密堆六方; ⑤面心立方-体心立方以及⑥面心立方-密堆六方等。

马氏体相变不仅发生在金属中, 在大量的陶瓷材料中也有出现, 例如钙钛矿结构的 BaTiO_3 、 PbTiO_3 等高温顺电立方相-低温铁电四方相以及 ZrO_2 中都存在这种相变, 利用 ZrO_2 的四方-单斜马氏体相变可有效地进行陶瓷高温结构材料的增韧。

18.1.3 有序-无序相变

有序-无序相变在结构上往往涉及到多组元固溶体中两种或多种原子在晶格点阵上排列的有序化。大量的多组元固溶体当温度降低时常会发生晶格中原子从统计随机分布的状态向不同原子分别占据不同亚点阵的有序化状态转变。例如, 组分为 AB 的合金, 在无序状态时, A、B 原子随机地占据在同一个点阵上。当温度降低到某一临界值时, 一种原子开始优先占据某一亚点阵, 而另一种原子则趋于占据另一个亚点阵, 从而形成一种部分有序的结构。随着温度的继续降低, 这种有序结构的有序化程度可能会进一步增加, 直至形成完全有序的固溶体。这类相变属于结构性相变, 它们发生于某一温度区间并涉及原子或离子的长程扩散和系统序参量的变化。

有序-无序相变的一个例子是连续固溶体铜-金合金中的 Cu_3Au 。由 X 射线衍射分析得知, 在高温无序状态下, 合金中 Au 和 Cu 原子近乎完全无规则地排列在面心立方 (FCC) 点阵上, 如图 18-5 (a) 所示。当温度降至其临界温度 ($T_c = 390^\circ\text{C}$) 以下, 合金中的 Au、Cu 原子开始发生偏聚, Au 原子择优占据立方体的顶点, Cu 原子则择优占据立方体的面心位置 [如图 18-5 (b) 所示], 并最后达到一种完全有序的结构。此时, 原面心立方晶胞可看成是由四个分别被一种原子占据的、相互穿插的简单立方亚点阵组成。

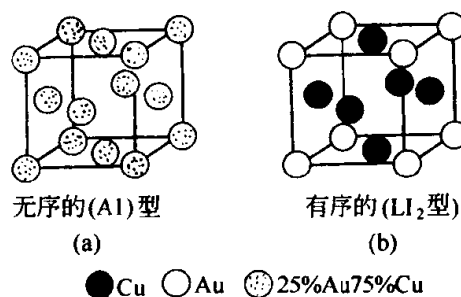


图 18-5 Cu_3Au 的有序-无序相变结构

有序-无序相变的发生常会伴随超结构现象的出现, 这是因为有序化过程使结构中出现富 A (或完全被 A 占据) 的晶面与富 B (或完全被 B 占据) 的晶面交替排列情况, 从而使布拉格衍射图上出现超结构衍射线。如图 18-6 所示, 铜-金合金中 CuAu 在高温时呈无序的面心立方结构, 在 385°C 以下退火则变为有序的四方结构。沿着结构的 c 轴, 出现交替排列的 Cu 原子层和

Au 原子层。有序-无序相变在金属材料中是普遍的，在无机非金属材料中也屡见不鲜。例如在几乎所有的尖晶石结构铁氧体中，高温时阳离子可同时无序地处在八面体或四面体位置，并呈无铁磁性状态。随着温度降低至某一临界值时，结构中开始出现离子在不同亚点阵上的择优占据有序化过程，并使材料出现铁磁性。相似的相变也出现在诸如 KDP (KH_2PO_4) 的铁电材料中，氢离子在其临界温度以下发生的靠近 PO_4^{-3} 基团的有序化排列导致了顺电-铁电-反铁电等相变，如图 18-7 所示。

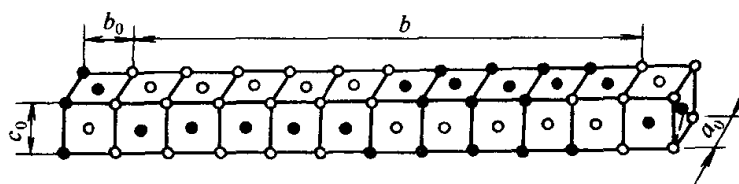


图 18-6 CuAu 的超结构示意图

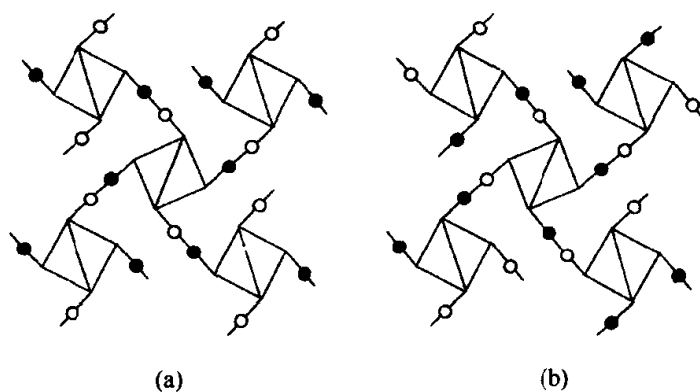


图 18-7 KDP 晶体结构中氢原子的有序化排列
(实心圆和空心圆分别为满和空的质子位)

18.1.4 其他形式相变

在丰富多彩的材料系统中，除前面所述的相变形式外，还存在一些有着特殊意义的相变形式，如无公度相变和液晶中所发生的相变。

所谓无公度相变是指晶态物质失去平移对称性的相变过程。人们发现，某些晶态材料在温度降低至某一温度 T_I 时，由于其长程关联作用使晶格不再具有严格的三维平移周期性，出现局域原子的性质受到一个周期性调制，调制波的波长与母相中晶体结构的周期之比为一无理数，故而称之为无公度调制，其相变产物称为无公度相。涉及的调制波，可以是结构上的调制、成分上的调制、乃至更细微层次如自旋结构上的调制。图 18-8 给出了两种无公度调制的示意结构。随着温度的继续降低并达到某一所谓锁定温度 T_L 时，材料的晶格平移性会重新出现而进入另一公度相。新相晶胞尺寸将是高

温相晶胞边长的整倍数。因此，无公度相就存在于 T_I 和 T_L 温度之间。无公度相变本质上属结构相变。在相变发生时，虽母相每个晶胞中的原子位移量互不相同，但其位置仍被某一周期函数所调制而保持着长程有序。目前已发现 100 多种材料具有无公度相变，例如 NaNO_3 ， $(\text{NH}_4)_2\text{BeF}_4$ ， BaMnF_4 ， KSeO_4 ， SiO_2 等等。无公度相存在的温区，窄的只有 $1\sim 2$ 度，宽的可达数十度甚至上百度。

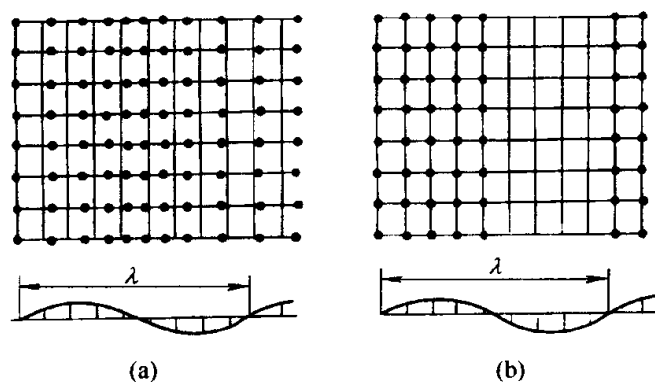


图 18-8 无公度相调制结构示意图

(a) 位移调制；(b) 成分调制

液晶是人们最为常见的现代电子产品数码、图像显示材料。然而，液晶的结构则具有人们不太熟悉的、介于液体与固体之间的中间相特征。它的力学性质相似于普通的液体，具有流动性；而其光学性质则呈各向异性，与晶体类似。大量研究表明，液晶相是具有棒状分子结构 ($1.5\sim 4.0\text{nm}$) 的晶态有机物在一定温度条件下部分或完全失去分子长程位置序，但在统计的意义上仍保留其长程方向序相变过程的产物。根据液晶结构有序性的差别，液晶可划分为向列型、胆甾型和近晶型，如第 3 章所述。图 18-9 显示了从各向同性相到向列相的转变。典型的向列型液晶和近晶型液晶材料分别是对氧化偶氮甲氧苯 (*p*-azoxyanisole) 和乙基对氧化偶氮苯酯 (ethyl *p*-azoxybenzoate)。

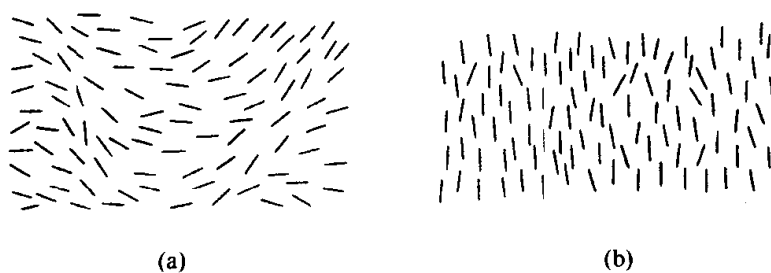


图 18-9 从各向同性相 (a) 到向列相 (b) 的转变

18.2 相变热力学

热力学是物理学的一个分支,是迄今发展得最为完善和普遍适用的理论。它在涉及热与能的众多学科领域中都有着广泛的价值和意义。借助于热力学普遍适用的理论来研究相变已经使人们在认识相变基本规律或机制等方面取得了很大的进展。相变热力学已是相变研究的重要部分。

18.2.1 重要的热力学函数

热力学研究的对象——系统与环境之间可以发生各种相互作用并使系统的状态发生变化。热力学状态函数对系统的状态及其所发生的现象给出宏观的描述,而无需考虑系统内部的结构细节。在经典热力学中,对系统的描述除温度、压力等强度变量外,还引入了广延性质的状态函数,它们分别为体积 V 、内能 U 、熵 S 、焓 H 、赫姆赫兹自由能 F 和吉布斯自由能 G 。内能描述了系统内部的能量总和;熵定义为系统在可逆过程中所吸收的热量与系统温度之比;焓为系统内能与外界对系统所作的功之和: $H = U + pV$;赫姆赫兹自由能 F 和吉布斯自由能 G 并非独立的热力学状态函数,它们分别被定义为:

$$F = U - TS; \quad G = H - TS \quad (18-2)$$

在讨论相变问题时,系统的吉布斯自由能 G 常更多地被用来描述系统的状态。根据上述定义,吉布斯自由能 G 可表示为:

$$G = G(T, p) = U - TS + pV \quad (18-3)$$

对于固体材料,考虑到其各向异性弹性应力应变能的存在,上式应改写成:

$$G = G(T, \sigma) = U - TS - \sum_{ij} \sigma_{ij} e_{ij} V_0. \quad (18-4)$$

此处, σ_{ij} 为应力张量的分量: e_{ij} 为应变张量的分量; V_0 为固体形变前的体积。如所研究的材料属于电介质或磁介质体,吉布斯自由能中外界对系统做功部分还应增加相应的电能和磁能项,于是有:

$$G = U - TS - \left(\sum_{ij} \sigma_{ij} e_{ij} + \sum_i E_i P_i \right) V_0. \quad (18-5)$$

$$G = U - TS - \left(\sum_{ij} \sigma_{ij} e_{ij} + \sum_i H_i M_i \right) V_0. \quad (18-6)$$

其中, E_i 为电场强度的分量: P_i 为电极化强度的分量: H_i 为磁场强度的分量; M_i 为磁化强度的分量。

根据可逆热力学过程的特点和数学上全微分原理,系统的热力学状态函数可分别有如下重要的微分形式:

$$dU = TdS - pdV; \quad dH = TdS + Vdp$$

$$dF = -SdT - pdV; \quad dG = -SdT + Vdp$$

其中,它们偏导数间数学上的恒等关系使许多在实际上难以测量的热力学量

得到了方便的实际应用。例如由内能微分式，可以得到：

$$\left(\frac{\partial U}{\partial S}\right)_V = T; \quad \left(\frac{\partial U}{\partial V}\right)_S = -p \quad (18-7)$$

由此再求其二阶导数得：

$$\left(\frac{\partial^2 U}{\partial V \partial S}\right) = \left(\frac{\partial T}{\partial V}\right)_S; \quad \left(\frac{\partial^2 U}{\partial S \partial V}\right) = -\left(\frac{\partial p}{\partial S}\right)_V \quad (18-8)$$

可得所谓的 Maxwell 关系式之一：

$$\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial p}{\partial S}\right)_V \quad (18-9)$$

将同样的方法用于 H 、 F 和 G 函数，可求出其余的 Maxwell 关系式：

$$\left(\frac{\partial T}{\partial V}\right)_p = -\left(\frac{\partial p}{\partial S}\right)_T; \quad \left(\frac{\partial T}{\partial p}\right)_S = -\left(\frac{\partial V}{\partial S}\right)_p; \quad \left(\frac{\partial T}{\partial p}\right)_V = \left(\frac{\partial V}{\partial S}\right)_T \quad (18-10)$$

常见易测的物理量与热力学偏导数的关系式有恒压热容、恒容热容、等温与绝热压缩率以及体热膨胀系数等：

$$c_V = T\left(\frac{\partial S}{\partial T}\right)_V = -T\left(\frac{\partial^2 F}{\partial T^2}\right)_V; \quad c_p = T\left(\frac{\partial S}{\partial T}\right)_p = -T\left(\frac{\partial^2 G}{\partial T^2}\right)_p \quad (18-11)$$

$$K_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T = -\frac{1}{V}\left(\frac{\partial^2 G}{\partial p^2}\right)_p; \quad K_S = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_S = -\frac{1}{V}\left(\frac{\partial^2 H}{\partial p^2}\right)_S \quad (18-12)$$

$$\alpha_V = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p = -\frac{1}{V}\left(\frac{\partial S}{\partial p}\right)_T \quad (18-13)$$

18.2.2 一级相变与高级相变

平衡态热力学理论指出，在等温等压的条件下，体系内各种自发进行过程的方向及其平衡状态的判据为： $(\Delta G)_{TP} \leq 0$ 。这一判据表明，在给定温度、压力条件下，若存在 $G_{II} - G_I < 0$ ，则 I 相可自发地转变为 II 相，反之则是不可能的。同时又表明，在任何相变点上，平衡共存两相的吉布斯自由能函数必须连续、相等。但作为自由能函数的各阶导数（对应于系统的熵、体积、比热容等），在相变点却可能发生不连续的跃迁。P. Ehrenfest（厄任费斯特）根据相变的这一热力学特征，首先提出按自由能函数导数连续情况来定义相变的级别：一个系统在相变点有直到 $(n-1)$ 阶连续的导数，但 n 阶导数不连续，则该相变定义为 n 级相变。

根据这一定义，一级相变的自由能函数的一阶导数在相变点上是不连续的，因而熵和体积的变化应该是跳跃的，意味着存在相变潜热和体积的变化。二级相变在相变点的一阶导数连续，所以熵和体积在相变过程中无跳跃变化，但其二阶导数（对应于系统的比热容、压缩率、膨胀系数等）将有不

连续的跳跃。至于三级或更高级相变，依此类推，将出现自由能函数的三阶或更高阶导数在相变点的跳跃变化。

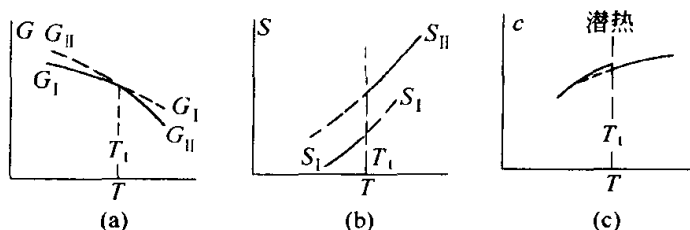


图 18-10 一级相变中有关热力学函数的变化

(a) 自由能; (b) 熵; (c) 比热容

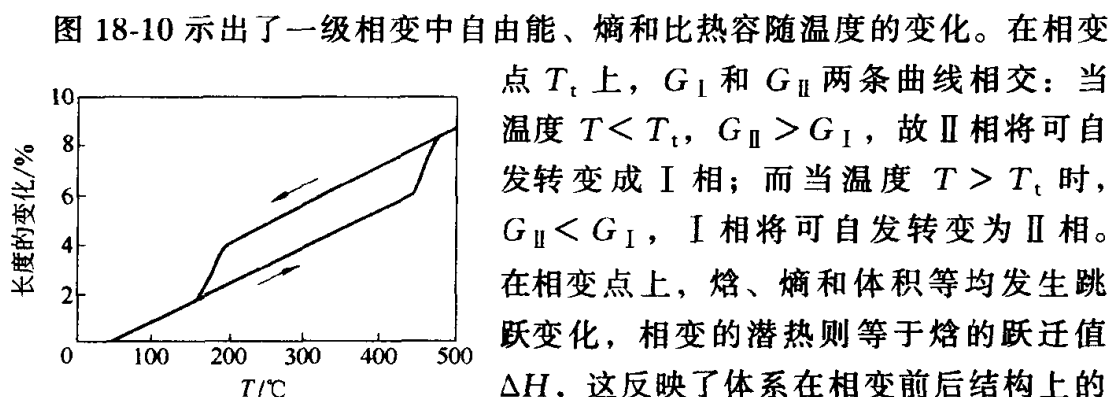


图 18-11 钴的一级相变中的
温度滞后现象

图 18-10 示出了一级相变中自由能、熵和比热容随温度的变化。在相变点 T_t 上, G_I 和 G_{II} 两条曲线相交: 当温度 $T < T_t$, $G_{II} > G_I$, 故 II 相将可自发转变成 I 相; 而当温度 $T > T_t$ 时, $G_{II} < G_I$, I 相将可自发转变为 II 相。在相变点上, 焓、熵和体积等均发生跳跃变化, 相变的潜热则等于焓的跃迁值 ΔH , 这反映了体系在相变前后结构上的明显差异。因此, 一级相变往往又属于结构上重构型相变, 在动力学上由于涉及到结构的重组, 常出现所谓的相变滞后现象, 如图 18-11 所示。

对于单组元系统, 两相的自由能函数可用变量 p 和 T 的曲面来描述。一级相变发生于两个相的交截线, 这个线的斜率就是著名的克拉佩龙-克劳修斯 (Clapayron-Clausius) 关系式:

$$\Delta G = G_{II} - G_I = (V_{II} - V_I)dp - (S_{II} - S_I)dT = \Delta V dp - \Delta S dT = 0 \quad (18-14)$$

因此有:

$$\frac{dp}{dT} = \frac{\Delta S}{\Delta V} = \frac{\Delta H}{T \Delta V} \quad (18-15)$$

图 18-12 给出了典型二级相变体系中热力学函数变化的情况。可以看出, 在低于相变点 T_c 时, II 相是稳定的: 在高于 T_c 时, I 相是稳定的。自由能曲线及其一阶导数在相变点连续, 故二级相变的自由能曲线实为一根连续曲线, 相变点为该曲线的奇点。因而, 在相变点两相合二为一, 不存在有明显的差异, 不会存在两相共存和相变滞后现象。在二级相变中, 熵和体

积呈连续变化, 因而相变不伴随潜热和体积突变发生。但熵和体积曲线在相变点并不光滑, 对应的一阶导数如比热容或热胀系数等将会发生跳跃变化。在有些情况下, T_c 点比热容会趋于无穷大而使比热容-温度曲线形似希腊字母 λ , 故常称为 λ 相变, 其 T_c 称为 λ 点。

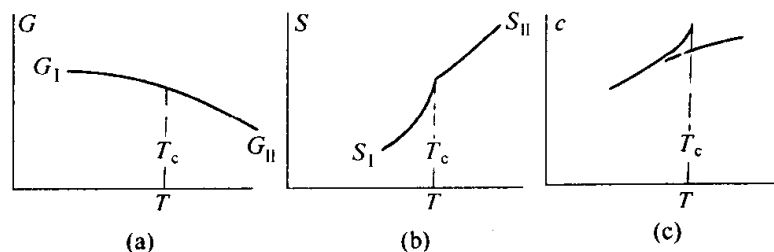


图 18-12 二级相变中有关热力学函数的变化

(a) 自由能; (b) 熵; (c) 比热容

在二级或二级以上高级相变中, 熵和体积变化的连续性使克拉佩龙关系式成为 $0/0$ 型的不定式。此时, 根据罗必达法则对不定式分子分母分别求导, 可以得到:

$$\frac{dp}{dT} = \frac{\partial \Delta S / \partial T}{\partial \Delta V / \partial T} = \frac{\Delta C_p}{TV \Delta \alpha_V}; \quad \frac{dp}{dT} = \frac{\partial \Delta S / \partial p}{\partial \Delta V / \partial p} = \frac{\Delta \alpha_V}{\Delta K_T} \quad (18-16)$$

这就是二级相变的 Ehrenfest 关系式。它适合于有限跳跃的二级相变。对具有无穷大跳跃 λ 型相变, Ehrenfest 关系式将不再适用, 而常采用 A. B. Pippard 相变点邻域近似分析结果:

$$c_p/T_c = \xi V_\alpha + C; \quad \alpha_V = \xi K_T + C \quad (18-17)$$

式中, C = 常数, $\xi = -(\partial^2 S / \partial T \partial p) / (\partial^2 S / \partial p^2)_T$ 。图 18-13 所示是 α - β 石英相变中有关参数与皮帕德关系式的拟合结果。

大量研究表明, 发生于自然界中的相变多半属于一级相变, 在大部分的金属或非金属材料中, 所涉及到的相变也多半为一级相变。二级相变的存在往往不及一级相变那样普遍, 但其丰富的物理内容一直吸引着物理学家和材料科学家们的研究兴趣。常见的二级相变包括在临界点的气-液相

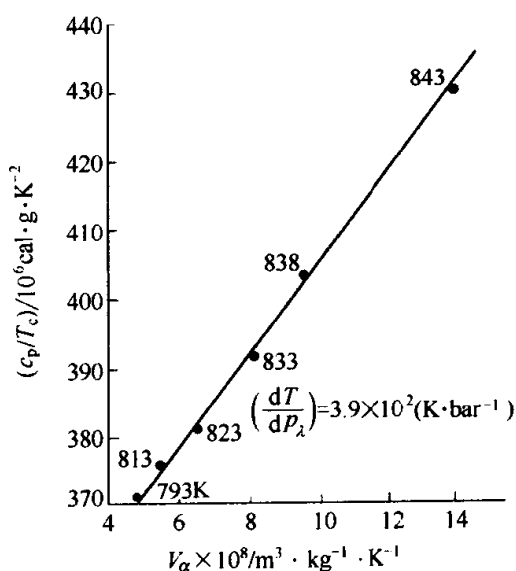


图 18-13 α - β 石英相变点邻域内 c_p/T_c 对 V_α 的实测关系

注: $1\text{cal} = 4.18\text{J}$, $1\text{bar} = 10^5\text{Pa}$

变、铁磁相变、超导相变、超流相变、部分固溶体的有序-无序相变、部分铁电相变等等。二级或更高级相变因其自由能函数导数的连续性常被称之为连续相变。与此对应，一级相变则被称之为不连续相变。

18.2.3 朗道相变理论简介

正如前面所述，二级相变遵循厄任费斯特关系式，在相变时熵和体积作连续变化。然而，二级相变的另一个重要的显著特征是其相变过程总伴随着高对称相的对称破缺和结构有序化的连续发展（或其逆过程）。物质结构的对称性取决于结构中存在的对称元素，对称性的改变意味着某些对称元素的消失或出现，因而总是突变式的，其结果是二者必居其一或非此即彼。在一级相变中，相变不仅改变了物相结构的对称性，同时反映结构有序程度的所谓序参量在相变前后也均会发生有限的突变。与此不同，在二级相变中伴随结构对称性突变的同时仍存在着某些作连续变化的序参量，它或它们在相变点只经历一个无限小量的变化，然后随着温度、压力或外场的变化而连续地变化。例如，在无序-有序相变中，结构中某一格点位置被原子 A 占有几率与原子 A 的浓度差是系统的序参量，在高温无序相中为零，而在低温有序相中不为零，序参量在相变点处开始从零连续增长；又如在 BaTiO_3 和 SiTiO_3 的立方-四方相变中，Ti 原子的位移和氧八面体的倾角是系统的序参量。应该指出，与某一相变相联系的序参量，在有些情况下比较容易确认，在有些情况下则需要认真分析方能得到。系统的序参量可以是标量，也可以是矢量。例如顺磁-铁磁相变，序参量为材料的宏观磁化强度 M 。

朗道的二级相变理论实际上是在平衡态热力学基础之上通过引入序参量来研究相变点邻域内二级相变所遵循的基本规律的理论，其形式简单，不仅是理解连续相变的基础，同时在众多其他类型的相变中可得到推广性的应用。

考虑一个简单的系统，其序参量为—标量 ζ_0 将系统吉布斯自由能函数 G 在相变点附近展开成序参量 ζ 的幂级数，即

$$G(T, p, \zeta) = G_0 + \alpha\zeta + A\zeta^2 + C\zeta^3 + B\zeta^4 + \cdots \quad (18-18)$$

此处系数 α 、 A 、 B 、 C 等均为 T 、 p 的函数。序参量为 ζ_0 物相的稳定条件为：

$$\left(\frac{\partial G}{\partial \zeta}\right)_{\zeta_0} = 0; \quad \left(\frac{\partial^2 G}{\partial \zeta^2}\right)_{\zeta_0} > 0 \quad (18-19)$$

在高对称相中， $\zeta = \zeta_0 = 0$ ，并在 T_c 以上是稳定相，而在 T_c 以下为不稳定相。因而要求一次项系数 $\alpha \equiv 0$ ；而二次项的系数 $A(T, p)$ 应满足下面的条件：

$$T > T_c, \quad A(T, p) > 0$$

$$T < T_c, \quad A(T, p) < 0$$

显然, 这一条件意味着当 $T = T_c$ 时, $A(T, p) = 0$ 。同时, 若相变点本身是稳定的, 那么将要求三次项系数为零, 四次项系数为正值, 即有

$$C(T, p)_{T_c} = 0; \quad B(T, p)_{T_c} > 0$$

由于系数 B 在相变点为正值, 在相变点附近仍应为正值。对于 C 则可能存在如下两种情况: ①由于系统对称性要求 $C(T, p) \equiv 0$, 即在式 (18-18) 中不出现三次项。于是, 决定相变条件的方程只有 $A(T, p) = 0$, 由此方程在 T - p 平面上可确定二级相变点所构成的一根曲线; ②假如 $C(T, p)$ 不恒为零, 那么决定相变的条件将由 $A(T, p) = 0$ 和 $C(T, p) = 0$ 两方程共同确定。此时, 连续相变只能发生在 T - p 平面上一个孤立的点上。

考虑 $C \equiv 0$ 的情况, 将吉布斯自由能函数 G 中的 $A(T, p)$ 在相变点附近展开并取其一次近似 $A(T, p) = a(T - T_c)$, $B(T, p)$ 可用 $B(T_c, p) > 0$ 代替。于是, 在相变点 T_c 附近的 G 函数写成:

$$G(T, p) = G_0 + a(T - T_c)\zeta^2 + B\zeta^4 \quad (18-20)$$

在 T_c 附近 G 函数取极小值的条件:

$$\left(\frac{\partial G}{\partial \zeta}\right) = 0 \Rightarrow \zeta(A + 2B\zeta^2) = 0 \quad (18-21)$$

可确定序参量 ζ 与温度的依赖关系。解出上述方程, 得:

$$\zeta_0 = 0$$

$$\zeta_0 = \pm \left(-\frac{A}{2B}\right)^{1/2} = \pm \left[\frac{a(T_c - T)}{2B}\right]^{1/2} \quad (18-22)$$

考虑当 $T > T_c$ 时, 只出现无序的高温相, 故其解为 $\zeta_0 = 0$ 。选择 $B = 0.1$, 作自由能函数 G 关于序参量 ζ 的变化图线, 如图 18-14 所示。从图中可以看到, 曲线关于 $\zeta = 0$ 是对称的; 当 $T > T_c$, 极小值在 $\zeta = 0$ 处, 无序相为平衡相; 当温度 T 趋近 T_c 时, 自由能曲线在 $\zeta = 0$ 处变得更加平坦。此时, 不仅 $\partial G / \partial \zeta = 0$, 而且 $\partial^2 G / \partial \zeta^2 = 0$; 当 $T < T_c$ 时, $\zeta = 0$ 为极大值, 而在 $\zeta = \pm [a(T_c - T)/2B]^{1/2}$ 处为极小值, 表明其有序相为稳定相。而且, 随着 T 从 T_c 下降, 序参量 ζ 从零按 $|\zeta| \propto (T_c - T)^{1/2}$ 增长。

在相变点附近熵的变化可由自由能函数关于温度 T 的偏导数给出。考虑自由能函数在相变点附近有展开式

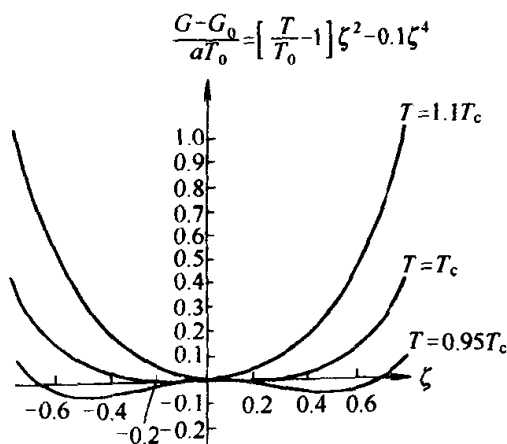


图 18-14 不同温度下自由能随序参量变化

(18-20), 故有:

$$\begin{aligned} S &= -\frac{\partial G}{\partial T} = S_0 - \frac{\partial}{\partial T} [a(T - T_c)\zeta^2 + B\zeta^4] \\ &= S_0 - a\zeta^2 - 2\zeta \frac{\partial \zeta}{\partial T} [a(T - T_c) + 2B\zeta^2] \end{aligned} \quad (18-23)$$

在高对称相中, $\zeta = 0$, 故 $S = S_0$; 在低对称相中, $\zeta^2 = a(T_c - T)/2B$, 所以有

$$S = S_0 + \frac{a^2}{2B}(T - T_c) \quad (18-24)$$

显然, 当 $T = T_c$ 时, $S = S_0$, 满足熵在 T_c 点连续的条件。对高、低对称相的熵函数差 $S - S_0$ 求关于 T 的导数, 可得比热容 c_p 在 T_c 点的 Ehrenfest 式的有限跃迁值:

$$\Delta c_p = T \left(\frac{\partial(S - S_0)}{\partial T} \right)_p = \frac{a^2 T_c}{2B} \quad (18-25)$$

至此, 朗道理论通过将序参量和平衡态热力学的结合对二级相变作出了定量的描述。遵循同样的方法, 不难求出定容比热容、压缩系数以及热膨胀系数在 T_c 处的跃迁值。

18.2.4 朗道理论的推广

朗道理论原来是针对二级相变提出的, 做适当修正可以推广应用到一级相变。

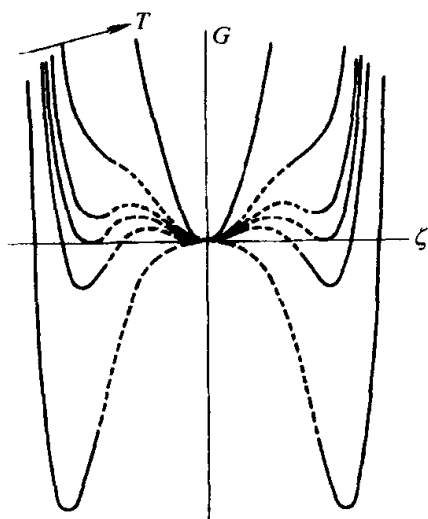


图 18-15 在不同温度下自由能作为序参量的函数 (朗道-德冯谢亚理论)

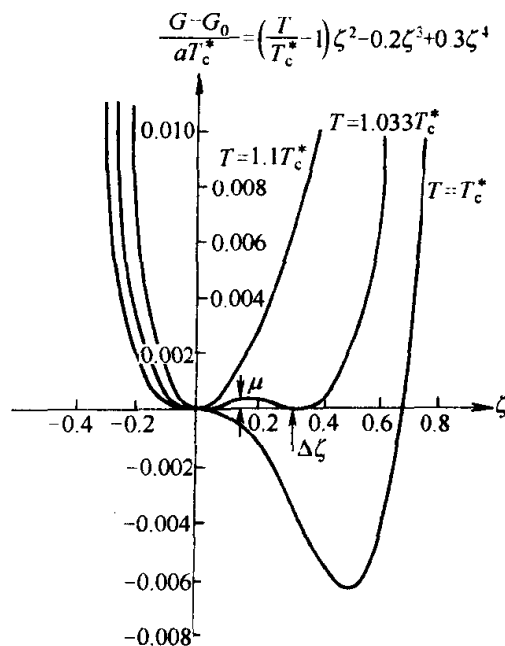


图 18-16 不同温度下自由能作为序参量的函数 (朗道-德让理论)

为了处理铁电相变中的弱一级相变，德冯谢亚提出了如下修正方案：即假设 $B(p) < 0$ ，为保持低温相的稳定性，展开式必须包含六次项，并假定其系数 $D > 0$ ，即

$$G = G_0 + a(T - T_c)\zeta^2 + B\zeta^4 + D\zeta^6 \quad (18-26)$$

图 18-15 给出用该修正得到的几个温度下的自由能曲线。

德让 (P.G.deGennes) 为了处理液晶中的一级相变，提出了朗道理论的另一修正方案，即在自由能中保留三次项，即 $C \neq 0$ ，因而，

$$G = G_0 + a(T - T_c)\zeta^2 + C\zeta^3 + D\zeta^6 \quad (18-27)$$

依然按朗道理论原样 $B > 0$ ，而令 $C < 0$ ，从而使低温平衡相中的 ζ 为正值。图 18-16 给出用该方法得到的几个温度下的自由能曲线。

磁场对铁磁相变的影响以及应力场对铁弹相变的影响与上面两种情况相似，也可对朗道理论的自由能进行类似修正。一般地，由于外场的存在，导致了在任何温度时序参量不为零，这样将使高对称相的对称性下降，无序相与有序相的差异减小。定性地看，外场附加项使自由能曲线呈现不对称性（见图 18-17）。

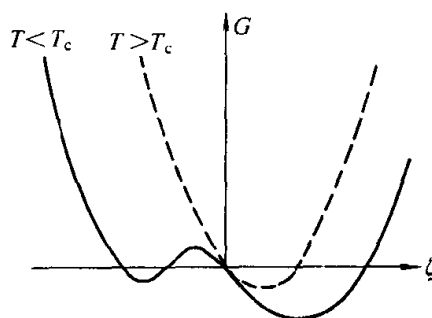


图 18-17 外场对朗道自由能曲线的影响

也有人将朗道理论推广到处理强一级相变的情况，如马氏体相变和重构型相变等。

18.3 固态相变动力学

与其他动力学过程研究类似，相变动力学旨在具体地描述相变过程的微观机制、过程进行的速率及其外界条件对它们产生的影响。经典的相变动力学研究主要是基于对相变中原子过程的具体分析，建立适当的物理模型，并利用物理学、尤其是热力学的一些基本概念对模型进行唯象或半唯象的处理来进行的。事实证明，这是一种非常有效的方法，它使相变动力学的研究取得了重要进展。随着近代物理学非线性理论和非平衡态热力学理论的发展与成熟，相变的动力学理论汲取了它们的一些精华，使原有的模型在适用范围或描述精度等方面得到了提高。当然，现代高精度的实验测试手段为相变动力学的研究与发展提供了必不可少、大量准确和可靠的资料。本节相变动力学讨论着重于其较为成熟的经典理论。

18.3.1 新相胚核形成过程

(1) 相变的热力学驱动力 从平衡态热力学观点看，当外界条件（如温

度、压力等)的变化使系统达到相转变点时,则会出现相变而形成新相。然而,事实上新相的出现往往需要母相经历一个“过冷”或“过热”的亚稳态才可能发生。这是因为欲使相变能自发进行,则必须使过程自由能变化 $\Delta G < 0$, 另一方面则是因为在非匀相转变过程中,由涨落而诱发产生的新相胚芽与母相间存在着界面。它的出现使系统的自由能升高,所以新相胚芽的出现所带来的系统体自由能项的下降必须足够大,才能补偿界面能的增加,于是必然出现“过冷”或“过热”等亚稳态。这种“过热”或“过冷”状态与平衡态所对应的自由能差便是相变的热力学驱动力。

以系统在恒压条件下进行相变为例。在相变平衡点 T_0 上,应有 $\Delta G = \Delta H - T_0 \Delta S = 0$ 。而在相变平衡点附近的某一温度 T 下, $\Delta G = \Delta H - T \Delta S \neq 0$ 。考虑在 T_0 的小邻域内 ΔH 和 ΔS 近似不随温度变化,比较上述两式便可得到:

$$\Delta G = \Delta H \left(\frac{T_0 - T}{T_0} \right) = \Delta H \left(\frac{\Delta T}{T_0} \right) \quad (18-28)$$

由此可见,自发相变要求 $\Delta G < 0$, 即应有 $\Delta H \Delta T / T_0 < 0$ 。若相变过程放热(如凝聚、结晶等过程),则 $\Delta H < 0$, 要使 $\Delta G < 0$, 必须有 $\Delta T > 0$ 。此时应有 $T_0 > T$ 而表明系统必须存在过冷的相变条件;若相变为吸热过程(如蒸发、熔融等过程),则 $\Delta H > 0$, 要满足自发相变的条件应有 $\Delta T < 0$, 即 $T_0 < T$, 这表明要发生相变的条件是过热。

在一些与浓度和蒸气压有关的相变过程中(如结晶沉淀、气相凝聚成液相或固相等过程),相变的热力学推动力则是系统浓度或蒸气压的过饱和度。由平衡态热力学理论可推得气相凝聚过程自由能变化 ΔG 与过饱和蒸气压 p 和平衡蒸气压 p_0 的关系式:

$$\Delta G = RT \ln \left(\frac{p_0}{p} \right) \approx RT \ln \left(1 + \frac{\Delta p}{p} \right) \approx RT \left(\frac{\Delta p}{p} \right) \quad (18-29)$$

式中, $\Delta p = p_0 - p$, 所以冷凝相变自发进行的推动力是蒸气压的过饱和 $p > p_0$ 。这一关系式在形式上完全适用于溶液中结晶或沉淀等与浓度相关的相变过程。考虑平衡时系统某一有关组分的饱和浓度为 c_0 , 而偏离该点的浓度为 c , 此时相变的自由能变化为:

$$\Delta G = RT \ln \left(\frac{c_0}{c} \right) \approx RT \ln \left(1 + \frac{\Delta c}{c} \right) \approx RT \left(\frac{\Delta c}{c} \right) \quad (18-30)$$

同理,欲使相变自发进行 ($\Delta G < 0$) 的条件是 $\Delta c < 0$, 即 $c > c_0$ 。

(2) 均匀成核过程 具备相变条件的系统一旦获取相变推动力,系统就具有发生相变的趋势。经典的成核-生长相变理论认为,新相的出现首先是通过系统中局域能量或浓度大幅度起伏涨落形成新相的胚芽而开始的,随后

由源于母相中的组成原子不断扩散至新相表面而使新相的胚芽得以长大。但是，在一定亚稳的条件下，并非任何尺寸的胚芽都可以稳定存在并得以长大而形成新相。尺寸过小的胚芽由于溶解度大很容易重新溶入母相而消失，只有尺寸足够大的胚芽才不会消失而成为可以继续长大形成新相的晶核。

现考虑在均匀亚稳态母相 α 中自发形成新相 β 胚芽时系统自由能的变化。显然，亚稳的母相转变为稳定的新相将使系统单位体积自由能降低，即 $\Delta G_{\alpha\beta} < 0$ ；但同时由于新相与母相存在结构上的差异而形成的新的界面可使系统的自由能增加，即 $\Delta G_s > 0$ 。因此，新相胚芽的出现使系统自由能发生的总变化应为此两项自由能变化的代数和：

$$\Delta G = V\Delta G_{\alpha\beta} + \Delta G_s = V\Delta G_{\alpha\beta} + A_{\alpha\beta}\sigma \quad (18-31)$$

式中， V 为新相胚芽的总体积， $A_{\alpha\beta}$ 为新相胚芽的总表面积， σ 为新相-母相间的界面能。

若假设所出现的新相胚芽为 n 个等大的球形颗粒，其半径为 r ，并考虑式 (18-31) 的关系，则有：

$$\Delta G = \frac{4}{3}\pi r^3 n \Delta H \frac{\Delta T}{T_0} + 4\pi r^2 n \sigma \quad (18-32)$$

由此可见，相变自由能变化 ΔG 为胚芽颗粒半径 r 和过冷度 ΔT 的函数。上式中的第一项常称为体积项，第二项为表面项。图 18-18 中给出了不同过冷度 ($T_3 > T_2 > T_1$) 条件下， ΔG 随胚芽颗粒半径 r 变化的关系。图中 r_k 为不同温度时所对应的胚芽临界半径，它是由式 (18-32) 关于 r 求极值所确定的：

$$r_k = -\frac{2\sigma}{\Delta G_{\alpha\beta}} = -\frac{2\sigma T_0}{\Delta H \Delta T} \quad (18-33)$$

所对应的自由能变化极值为：

$$\Delta G_k = \frac{16\pi\sigma^3}{3\Delta G_{\alpha\beta}^2} = \frac{4}{3}\pi\sigma r_k^2 \quad (18-34)$$

此自由能变化临界值实际上为形成一个临界晶核所必须越过的能垒，所以又常称为成核功。由图 18-18 和式 (18-33) 可以看到，一定的过冷条件对应于一定的临界晶核尺寸。当系统中由于涨落而出现的胚芽小于临界尺寸时，这些胚芽将自发地重新消溶回母相，只有当尺寸大于临界值的胚芽才可发育成晶核并得以继续长大成新相。相变系统的临界晶核

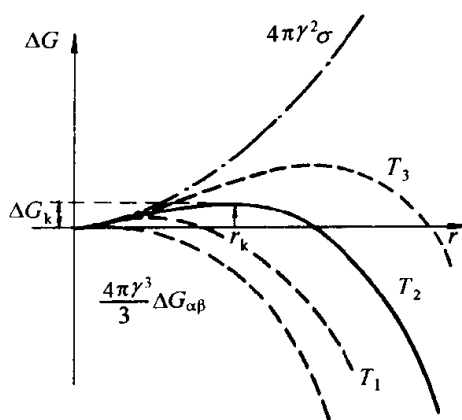


图 18-18 自由能变化与胚芽尺寸及过冷度的关系 ($T_3 > T_2 > T_1$)

尺寸决定于相变单位体积自由能变化和新相-母相界面能的相对大小。在液相结晶过程中 ($\Delta H < 0$)，过冷度越高，临界尺寸越小，成核功越低，说明可能存在的可长大的晶核数越多，相变也越易进行；当系统温度接近相变温度时， $\Delta T \rightarrow 0$ 而 $r_k \rightarrow \infty$ 。显然，此时析晶相变过程将变得不可能发生。根据玻尔兹曼统计，在平衡的条件下母相中具有临界尺寸的晶核密度应为：

$$n_k = n_0 \exp\left(\frac{-\Delta G_k}{k_B T}\right) \quad (18-35)$$

式中， n_0 为母相中可提供成核的位置的密度，对于均匀成核过程，它应等于母相单位体积中的原子或分子数， k_B 为玻耳兹曼常数。

然而，应该指出，临界晶核的平衡密度并非等于系统中实际长大的晶核密度。这是因为临界晶核与母相处于一种动态平衡，晶核的长大和消溶具有相同的几率。为了使一个临界晶核稳定并得以长大，至少要有一个原子从母相转入晶核。因此实际长大的晶核产生的速率应该是临界晶核密度与一个反映临界晶核附近母相中原子进入晶核多少和快慢的频率因子 β 的乘积：

$$\beta = S v_0 \exp\left(\frac{-\Delta G_a}{k_B T}\right) \quad (18-36)$$

式中， S 为在晶核附近母相的原子数， v_0 为这些原子的振动频率， ΔG_a 为这些原子进入晶核所必须越过的能垒高度。

影响晶核产生速率的另一个因素是系统中晶核的形成和长大也处于一个动态过程。在此过程中获得长大条件的晶核不断地离开临界晶核的群体，同时新的临界晶核又不断地补充进这个群体。当这个过程达到稳定态时，其临界晶核密度将低于平衡态下的密度。J.B.Zeldovich 从理论上对此进行了分析并提出在式 (18-36) 中引入因子：

$$Z = \frac{-1}{2\pi k_B T} \left(\frac{\partial^2 \Delta G(n)}{\partial n^2} \right)^{1/2} \quad (18-37)$$

在一般情况下，因子 $Z \approx 0.05$ 左右。基于上述两方面的影响并考虑在非匀相转变中，相变条件具备的瞬间并不能立即成核，而需要经历一段“孕育时间” τ ，可将系统在 t 时刻的稳定成核速率表达为：

$$J(t) = Z S v_0 n_0 \exp\left(-\frac{\Delta G_k + \Delta G_a}{k_B T}\right) \exp\left(-\frac{\tau}{t}\right) \quad (18-38)$$

应该强调指出，在有限的相变体系中稳定的成核过程不可能无限期地延续。随着相变的进行和母相量的减少，往往会出现相变驱动力的下降或成核势垒的升高，而最终使成核过程趋于停顿。在非匀相转变过程中，典型的新相粒子数随时间变化的关系如图 18-19 所示。

(3) 非均匀成核过程 在上述的均匀成核讨论中，我们曾假定相变系

统中各个位置上具有相同的成核几率。然而, 实际情况并非总是如此。当趋于冷凝的蒸气中悬浮着尘埃、趋于结晶的液相中含有杂质、或趋于发生晶型转变的固体中存在着界面、位错等缺陷时, 相变所需的成核过程往往会优先并容易地发生在这些特殊的区域上。在这种情况下, 成核过程将不再均匀地分布于整个系统, 故常称之为

非均匀成核。非均匀成核之所以比均匀成核更容易发生, 其主要原因是均匀成核中新相胚芽与母相间的高能量界面被非均匀成核中新相胚芽与杂质异相间的低能量界面所取代。显然, 这种界面的代换比界面的创生所需要的能量要少, 从而使成核过程所需越过的能垒降低, 进而使非均匀成核能在较小的相变驱动力下进行。

现考虑一个简单的情况。假设在过饱和的母相 α 中存在杂质 S , 新相

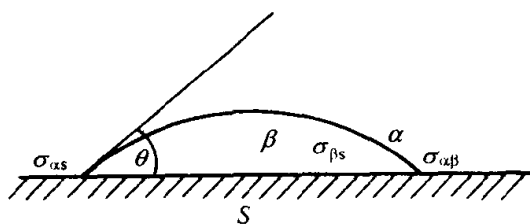


图 18-20 新相 β 胚核在母相与杂质 S 的界面上形成

β 的胚芽依附在 S 的表面形成, 其成核几何模型如图 18-20 所示。当新相胚芽形状为一曲率半径为 r 的球冠, 并考虑各界面能具有各向同性和彼此达成平衡, 则有如下关系式:

$$\sigma_{\alpha\beta}\cos\theta + \sigma_{\beta s} = \sigma_{\alpha s} \quad (18-39)$$

式中, θ 为新相-杂质界面接触角。由初等几何知识不难求出, 球冠状新相胚核体积、表面积以及与杂质 S 的界面面积:

$$\begin{aligned} V_{\beta} &= \frac{\pi}{3} r^3 (2 - 3\cos\theta + \cos^3\theta) \\ A_{\alpha\beta} &= 2\pi r^2 (1 - \cos\theta) \\ A_{\beta s} &= \pi r^2 \sin^2\theta \end{aligned} \quad (18-40)$$

并由下式可确定成核过程系统自由能的变化:

$$\Delta G_s(r) = \sigma_{\alpha\beta} A_{\alpha\beta} + (\sigma_{\beta s} - \sigma_{\alpha s}) A_{\beta s} + V_{\beta} \Delta G_{\alpha\beta} = \Delta G(r) f(\theta) \quad (18-41)$$

式中, $\Delta G(r)$ 是均匀成核中形成半径为 r 球形胚核时系统自由能的变化, $f(\theta)$ 仅为接触角 θ 的函数:

$$f(\theta) = \frac{1}{4} (2 + \cos\theta) (1 - \cos\theta)^2 \quad (18-42)$$

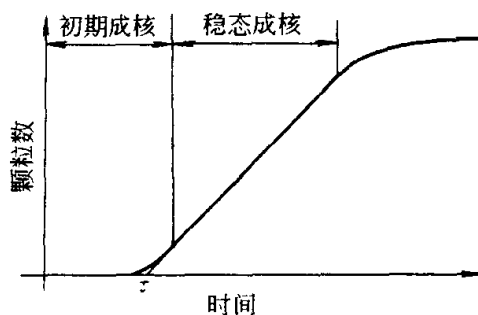


图 18-19 新相粒子随时间的变化

由式 (18-41) 容易求出临界胚核半径 r_{sk} 和相应的成核势垒 ΔG_{sk} , 它们与均匀成核过程的胚核临界半径和成核势垒具有如下关系:

$$r_{sk} = r_k, \quad \Delta G_{sk} = \Delta G_k f(\theta) \quad (18-43)$$

由此看出, 非均匀成核过程具有和均匀成核过程相同的胚核临界半径,

但成核势垒则随成核所依附界面的接触角 θ 的减小而下降。

当 $\theta = \pi$, $\Delta G_{sk} = \Delta G_k$; 若 $\theta \rightarrow 0$, 则 $\Delta G_{sk} \rightarrow 0$ 。即一个可完全被新相润湿的界面可使成核势垒完全消失。

在固态相变中更常见的情况是新相在母相的晶界面上成核。这种成核机制如图 18-21 所示。设 $\sigma_{\alpha\beta}$ 为各向同性, 于是在诸表面张力达成平衡时,

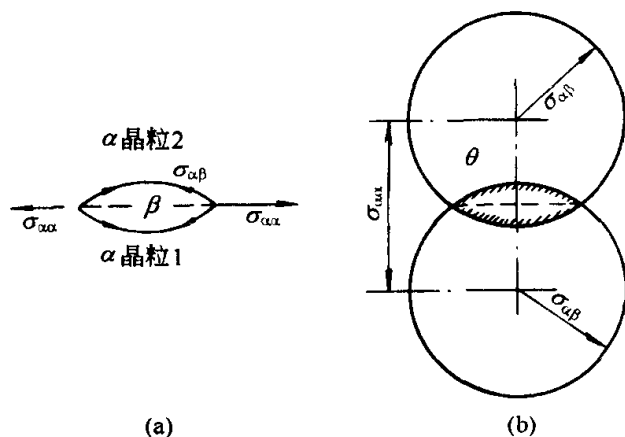


图 18-21 新相 β 胚核在母相 α 晶界上形成

β 相胚核将为球面凸透镜形。当其半径为 r , 半夹角为 θ , 在 $2\sigma_{\alpha\beta} > \sigma_{\alpha\alpha}$ 条件下, 可求得其成核势垒为:

$$\Delta G_{bk} = \Delta G_k g(\theta)$$

$$g(\theta) = (1 - \cos\theta)^2 (2 + \cos\theta) / 2 \quad (18-44)$$

式中, ΔG_k 为新相在均匀成核条件下的成核势垒; 接触角依赖关系 $g(\theta)$ 相同于式 (18-44)。因此, 当 $\sigma_{\alpha\alpha} \geq 2\sigma_{\alpha\beta}$ 时, 晶界成核势垒可完全消失, 新相的成核过程将会沿着母相的晶界作浸润式的扩展。

非均匀成核过程也常在晶粒的棱边 (三个晶粒的交线) 或角隅 (四个晶粒的交点) 上发生。有关计算表明, 成核势垒随接触角的余弦快速下降, 并且较低几何维数的成核位置有更低的成核势垒, 如图 18-22 所示。但应该注意, 实际材料中均匀成核位置、界面成核位置或更低维数的成核位置在数量上是依次急速递减的。这两种趋势竞争的结果可概括成一张“成核机制图”, 如图 18-23 所示。图中纵坐标与晶粒尺寸 L 、晶界有效厚度 δ 以及成核势垒 ΔG_k 有关, 横坐标为 $\sigma_{\alpha\alpha} / \sigma_{\alpha\beta}$ 。由图可见, 大的晶粒尺寸、小的成核势垒或小的 $\sigma_{\alpha\alpha}$ 有利于均匀成核过程。反之, 则有利于界面、晶棱或角隅位置上的非均匀成核过程。考虑到固相界面能的各向异性, 一般来说, 新相晶核与其相邻晶粒间的界面能并不相等。此时, 新相晶核将偏向并逐步长入与它匹配较好因而界面能较低的那个晶粒, 而不是对称地长在相邻晶粒的中间。

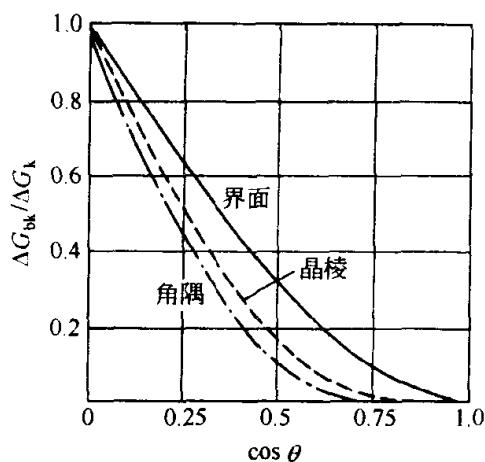


图 18-22 界面、晶棱和角隅成核势垒与均匀成核势垒比的 $\cos \theta$ 关系

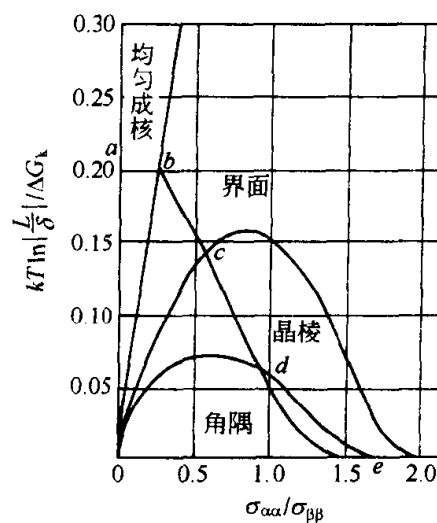


图 18-23 母相晶粒尺寸、成核势垒与 $\sigma_{\alpha\alpha}/\sigma_{\beta\beta}$ 的关系

发生非均匀相变成核过程的另一类位置是母相晶粒中存在的各种位错线和点缺陷。由于实际晶体中位错线和点缺陷的密度往往很大，所以它们对成核的贡献是不能忽视的。卡恩 (J.W.Cahn) 曾提出新相在母相位错线上成核的模型。如图 18-24 所示，假设以母相中的位错线为轴线，形成一个纺锤形新相胚核使位错线的弹性能完全消弛，此时单位长度胚核而导致的自由能变化为：

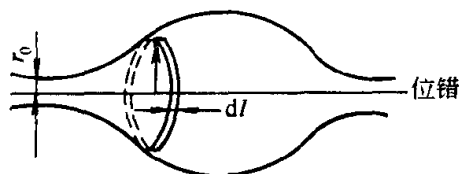


图 18-24 位错线上新相成核的几何模型

$$\Delta G_d = -A \ln r + 2\pi\sigma r + \pi r^2 \Delta G_{\alpha\beta} \quad (18-45)$$

式中第一项为位错弹性能，系数 A 对于刃位错和螺位错分别为： $\mu b^2/4\pi (1-\nu)$ 和 $\mu b^2/4\pi$ 。对式 (18-45) 关于 r 求导并令其为零，可得晶核的临界半径为：

$$r_{dk} = \frac{\sigma}{2\Delta G_{\alpha\beta}} [1 - (1 - \alpha_d)^{1/2}] \quad (18-46)$$

$$\alpha_d = 2A\Delta G_{\alpha\beta}/\pi\sigma^2$$

ΔG_d 与 r 的依赖关系如图 18-25 所示，当 $\alpha_d < 1$ 时， ΔG_d 有一极大和一极小值，存在成核势垒 ΔG_{dk} ；当 $\alpha_d > 1$ 时， ΔG_d 随 r 的增大而减小，成核势垒消失。在一般情况下， $\alpha_d < 1$ ，此时位错位置成核势垒与均匀成核势垒的比值与 α_d 的关系示于图 18-26 中。

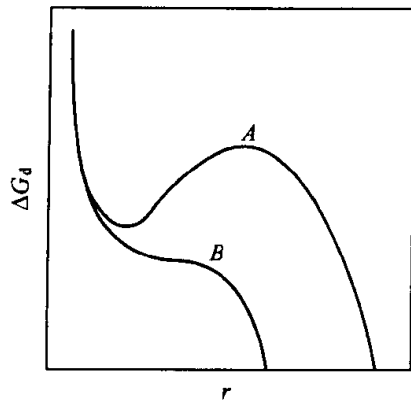


图 18-25 位错线上成核自由能变化与半径的关系
A— $\alpha_d < 1$; B— $\alpha_d > 1$

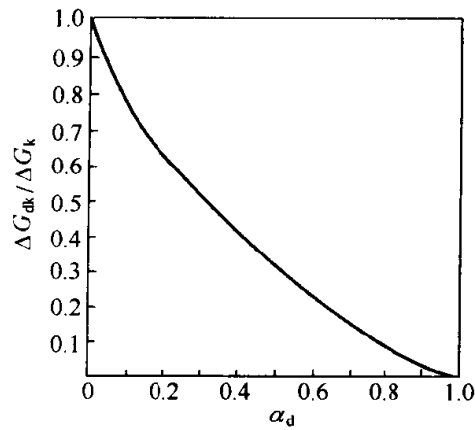


图 18-26 位错线上成核势垒与均匀成核势垒比值和 α_d 的关系

就非均匀成核过程中的成核速率而言，它与均匀成核情况下的式 (18-38) 应具有相同形式，所不同的是对于非均匀成核过程，应在式 (18-38) 中引入相应的非均匀成核势垒并将成核位置密度 n_0 理解成系统中的实际非均匀成核位置密度，而不是原来均匀成核情况下的单位体积中原子或分子的数目。

18.3.2 胚核生长和粗化过程

新相的稳定胚核一旦形成，随之便是通过相界面的移动而得到长大，即开始胚核的生长过程。一般来说，等温生长的速度取决于相变的驱动力和跃迁于新相上原子的具体迁移过程。如果新相和母相具有相同的化学组成，那么控制生长速度的过程将是原子由母相穿过界面跃迁于新相上这一短程扩散。如果新相和母相具有不同的化学组成，例如过饱和固溶体的脱溶分解，则新相生长不仅需要原子穿越相界面这一环节，同时还涉及有关组分在母相中的长程扩散。此时，新相的生长速度将取决于两者中较慢的环节，但多数情况将受控于长程扩散。

(1) 界面控制型生长 新相胚核的生长过程受控于界面区原子的短程扩散则属界面控制型生长。在这种情况下，新相界面的向前推移有两种可能的方式：一是在界面所在的各处，母相原子可独立、同时地穿过界面而成为新相的原子。这种生长模式的界面在微观上是模糊、粗糙的，可由多个原子层所构成，但界面的移动则是完全连续的，并在界面各处同时发生。另一种生长模式是台阶生长机制，如图 18-27 所示。其界面在微观上是光滑的，但存在原子尺度的台阶，母相中的原子只在这些台阶附近进入新相，界面的推移往往通过台阶的横向生长来进行。只有当这种台阶横向扫过后，界面才可能向前得到一个原子尺度的生长。应该指出，光滑界面生长和粗糙界面生长之

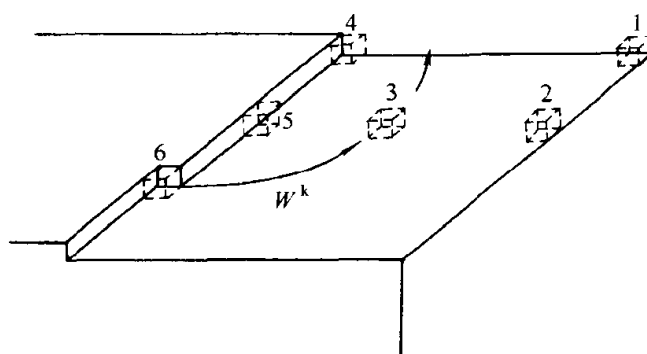


图 18-27 光滑界面的台阶生长机制示意图

间并无一条明确的界线，随着系统热力学量的变化可发生相互转化。通常相变驱动力的大小起着重要作用，驱动力大，趋向于连续生长；驱动力小，则趋向于台阶式生长。此外在固态相变中，新相母相间完全不共格，则其生长机制常类似于连续生长；如两相结构不同，但存在某些匹配良好的晶面，那么这些共格界面的生长就会出现类似光滑界面生长的情况，需要借助于台阶的机制。下面着重讨论第一种模式的生长速率。

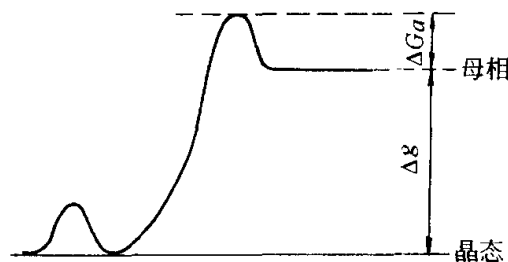


图 18-28 母相中原子穿越界面至新相表面自由能变化

图 18-28 所示为一个母相原子穿过粗糙界面并占据新相上一个格点位置所引起系统自由能的变化。考虑此时新相表面上可供占据的位置数足够的大，故新相生长的速率可按单位时间内由母相跃迁到新相表面上的原子数 n_{12} 和由新相表面跃迁回母相的原子数 n_{21} 之差来计算：

$$\begin{aligned} n_{12} &= n_0 \nu \exp\left(-\frac{\Delta G_a}{k_B T}\right) \\ n_{21} &= n_0 \nu \exp\left(-\frac{\Delta G_a + \Delta g}{k_B T}\right) \end{aligned} \quad (18-47)$$

其中 n_0 为新相表面单位面积上的位置数， ν 为界面附近原子的振动频率， ΔG_a 为母相原子到达新相表面所需越过的势垒， Δg 为相变驱动力。于是，可求得如下新相界面向前推进的线速率：

$$V = a \frac{n_{12} - n_{21}}{n_0} = a \nu \exp\left(-\frac{\Delta G_a}{k_B T}\right) \left[1 - \exp\left(-\frac{\Delta g}{k_B T}\right)\right] \quad (18-48)$$

式中， a 为新相表面的原子层间距。当相变驱动力 $\Delta g \ll k_B T$ ，如相变过程离开平衡态很小时，式(18-48)简化为：

$$V = av \frac{\Delta g}{k_B T} \exp\left(-\frac{\Delta G_a}{k_B T}\right) = M \Delta g \quad (18-49)$$

此处引入迁移率 $M = av \exp(-\Delta G_a/k_B T)$, 新相的生长速率与相变驱动力成正比。这便是威尔逊-夫仑克尔(Wilson-Frenkel)公式; 当相变过程远离平衡态时, 即 $\Delta g \gg k_B T$, 新相生长线速率式(18-48)可简化为:

$$V = av \exp\left(-\frac{\Delta G_a}{k_B T}\right) \quad (18-50)$$

它是新相可能的最大生长速率, 为新相生长的一种极限情况。

(2) 长程扩散控制型生长 在相变过程中, 新相与母相具有不同的化学组成是经常发生的。例如在过饱和固溶体中生成沉淀相。这时随着新相稳定胚核的形成和长大, 在新相附近的母相中便会出现溶质原子的贫化区, 并产生浓度梯度。远离新相的溶质原子在此浓度梯度的驱动下将朝着新相作长程扩散, 不断地提供新相长大所需的组分物质。于是, 长程扩散成了新相生长的重要环节, 并对生长速率常起着决定性的作用。

首先讨论简单的一维长程扩散型生长。为方便起见, 设新相中溶质浓度大于母相浓度: $c_\beta > c_\alpha$ (实际情况也可以是新相溶质浓度低于母相的浓度, 如图 18-29 所示), 同时假设扩散系数与浓度无关和忽略母相中平衡浓度与新相颗粒尺寸或新相界面曲率大小的依赖关系, 即考虑无限大平板一维增厚生长, 由连续性方程可得:

$$(c_\beta - c_\alpha) dx = D \left(\frac{\partial c}{\partial R} \right)_{R=x} dt \quad (18-51)$$

因而有生长速率 v 的表达式:

$$v = \frac{dx}{dt} = \frac{D}{(c_\beta - c_\alpha)} \left(\frac{\partial c}{\partial R} \right)_{R=x} \quad (18-52)$$

此时, 若由实验测出新相附近母相中溶质浓度随时间变化的具体分布, 则可由式(18-52)求得相应时刻新相的生长速率。然而, 实际上测量浓度分布常常是非常困难的。曾纳(C. Zener)曾假设新相附近的浓度分布具有线性特征, 如图 18-30 所示。显然, 溶质原子守恒要求图中的两阴影区面积必须相等, 由此可得溶质原子贫化区厚度 ξ 为:

$$\xi = 2 \frac{(c_\beta - c_0)}{(c_0 - c_\alpha)} x \quad (18-53)$$

并有

$$\left(\frac{\partial c}{\partial R} \right)_{R=x} = \frac{c_0 - c_\alpha}{\xi} \quad (18-54)$$

代入式(18-52), 并积分得:

$$x^2 - x_0^2 = \frac{(c_0 - c_a)^2}{(c_\beta - c_a)(c_\beta - c_0)} Dt \quad (18-55)$$

若 $x \gg x_0$, 可以得出生长速率为:

$$v = \frac{dx}{dt} \approx \frac{(c_0 - c_a)}{2(c_\beta - c_a)^{1/2}(c_\beta - c_0)^{1/2}} \left(\frac{D}{t} \right)^{1/2} \quad (18-56)$$

此式表明, 随着新相的长大, 其生长速率将逐渐降低。这是因为新相的长大使其外侧的溶质原子贫化区进一步扩大, 于是继续长大所需的溶质原子需从更远的地方扩散而来, 从而需要更长的扩散时间。

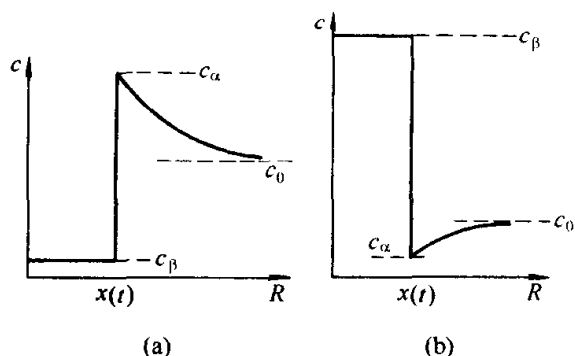


图 18-29 新相生长过程中系统

内溶质原子浓度分布

(a) 新相溶质浓度低于母相;

(b) 新相溶质浓度高于母相

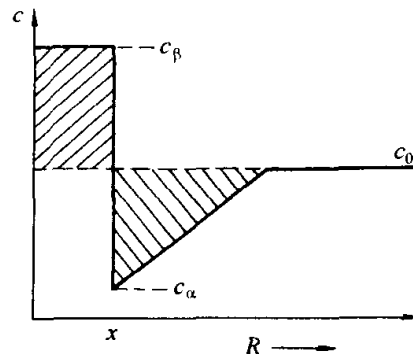


图 18-30 一维生长中溶质原

子浓度分布的线性近似

对于三维球形胚核的长大过程, 若生长过程具有球对称性, 同样可以利用一维的连续性微分方程:

$$4\pi r^2 (c_\beta - c_a) dr = 4\pi r^2 D \left(\frac{\partial c}{\partial R} \right)_{R=r} dt \quad (18-57)$$

考虑到在母相相对过饱和度不大的情况下, 球形胚核周围溶质原子贫化区尺寸要比胚核尺寸大得多, 故近似认为 r 为一常数。此时由菲克方程求出满足边界条件的稳定态解为:

$$c(R) = c_0 - (c_0 - c_a) \frac{r}{R} \quad (18-58)$$

代入式(18-57), 并积分得:

$$r^2 - r_0^2 = \frac{2(c_0 - c_a)}{(c_\beta - c_a)} Dt \quad (18-59)$$

对于 $r \gg r_0$, 有相似于式(18-56)的生长速率:

$$v = \frac{dr}{dt} \approx \frac{(c_0 - c_a)^{1/2}}{\sqrt{2}(c_\beta - c_a)^{1/2}} \left(\frac{D}{t} \right)^{1/2} \quad (18-60)$$

比较一维平板增厚生长和球形颗粒三维球面生长的解, 可以看出, 其生

长速率 v 和生长尺度 x 或 r 均具有相同的时间函数关系:

$$x \propto (Dt)^{1/2}, v \propto \left(\frac{D}{t}\right)^{1/2} \quad (18-61)$$

不难证明, 对于二维生长 (如无限长柱面生长), 也可得出相同的结果。事实上式 (18-61) 的关系是各种形状胚核均匀稳定生长所共同遵循的关系。根据这种关系, 生长过程中的新相颗粒形状将保持不变。然而, 实际生长过程中的情况要复杂得多, 界面能的各向异性以及存在的各种干扰都可以使这种均匀稳定的生长难以保持而出现非稳定的情况。

(3) 新相颗粒的粗化 粗化是指在相变过程中所形成的新相颗粒平均尺寸增大的动力学过程。它可以发生在相转变的后期, 也可能伴随着新相胚核形成和长大等过程而发生在相变的中期或甚至于前期。粗化过程的驱动力是蕴藏于系统内部的巨大界面能。一个平均粒径为 30nm 的新相颗粒系统粗化成平均粒径为 300nm 后, 系统中的总界面能可降低 100 倍。粗化过程的具体机制往往因系统不同而不同。其中, 新相颗粒的粗化可通过晶界移动实现大颗粒吞并小颗粒的过程来达到, 犹如烧结中、后期出现的二次再结晶过程, 其速度受界面处原子穿越界面的短程扩散控制。此外, 小颗粒的溶解、溶质原子的长程扩散以及大颗粒的长大是大多数多元系统中新相颗粒粗化的重要机制, 其中溶质原子的长程扩散往往控制着粗化的速度。下面讨论一个弥散分布的新相颗粒粗化过程。

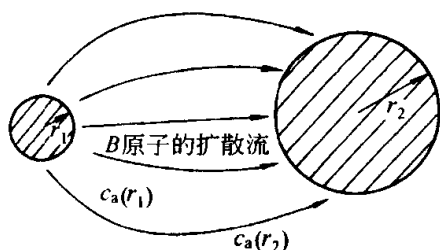


图 18-31 不同尺寸新相颗粒
周围溶质扩散流场

假设母相为 B 在 A 中的稀薄固溶体, 新相则几乎为纯 B 构成。同时假设新相所占的体积分数甚小于 1, 这意味着颗粒间平均距离远大于颗粒尺寸。如图 18-31 所示, 系统中的某两颗粒 1 和 2, 其尺寸具有 $r_2 > r_1$ 关系, 所以母相中与颗粒 1 平衡的溶质浓度将大于与颗粒 2 平衡的浓度, 即 $c_1 > c_2$ 。这种两颗粒间出现的浓度差将导致溶质原子由小颗粒向大颗粒的扩散,

进而使小颗粒周围的溶质浓度降低, 而大颗粒周围的溶质浓度升高。为保持与颗粒的平衡浓度, 小颗粒将逐渐溶解, 大颗粒则将吸收其附近母相中的过剩溶质而逐渐长大。现假设系统中颗粒的平均粒径为 r_a , 因而母相中溶质原子的平均浓度应为 $c(r_a)$, 它对应于远离颗粒处母相的溶质浓度。所以, 一般可以认为, 如果颗粒 $r > r_a$, 此颗粒将会长大, 而若 $r < r_a$, 此颗粒将会逐渐缩小。

选择某一颗粒为中心, 在母相中半径为 R 的球面上溶质原子扩散的总

流量应为:

$$J = -4\pi R^2 D \frac{dc}{dR} \quad (18-62)$$

若考虑扩散进球面的溶质原子全部用于其中心处球形颗粒的粗化, 则有:

$$4\pi r^2 \frac{dr}{dt} = -4\pi R^2 D \frac{dc}{dR} \quad (18-63)$$

亦即

$$-\frac{dR}{R^2} = \frac{D}{r^2 \left(\frac{dr}{dt} \right)} dc \quad (18-64)$$

取边界条件 $R = r, c = c(r); R \rightarrow \infty, c = c(r_a)$ 并积分得:

$$\frac{1}{r} = \frac{D}{r^2 \left(\frac{dr}{dt} \right)} [c(r_a) - c(r)] \quad (18-65)$$

利用吉布斯-汤姆逊母相中平衡浓度与颗粒表面曲率半径的近似关系式:
 $c(r) = c(\infty)[1 + \alpha/r]$ (对于球形颗粒式中 $\alpha = 2\sigma v/k_B T$), 可得:

$$\frac{dr}{dt} = \frac{2D\sigma v c(\infty)}{k_B T r} \left(\frac{1}{r_a} - \frac{1}{r} \right) \quad (18-66)$$

根据该式可绘出 dr/dt 与 r 的关系, 如图 18-32 所示。由图可以看出: ①半径小于 r_a 的新相颗粒都将逐渐变小, 其收缩速度随 r 的变小而增大, 并最终使这些颗粒消失; ②半径大于 r_a 的颗粒均可长大, $r = 2r_a$ 的颗粒长大的速度最大; ③随着粗化过程的进行, 系统中颗粒的平均粒径 r_a 也在不断地增大。因此, 一些虽在粗化过程的早期就开始长大但尺寸仍较小的颗粒, 会因为长大速度小而被 r_a 赶上或超过, 从而将出现由长大转为逐渐溶解的过程并最终归于消失; ④随着平均粒径的增大, 颗粒长大的平均速度逐渐下降。

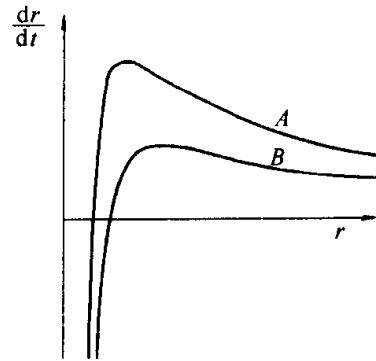


图 18-32 球形颗粒粗化速度与粒径的关系

平均粒径: $r_a(B)/r_a(A) = 1.5$

为了求出粗化过程中平均粒径随时间的变化关系, 格林伍德(G.W. Greenwood)曾考虑取生长速度最快的颗粒作为代表, 由式 (18-66) 得:

$$\frac{dr_a}{dt} = \frac{D\sigma v c(\infty)}{2k_B T r_a^2} \quad (18-67)$$

积分后得:

$$r_a^3 - r_{a0}^3 = \frac{3}{2} \frac{D\sigma v c(\infty)}{k_B T} t \quad (18-68)$$

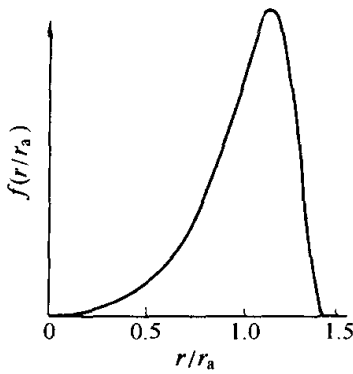


图 18-33 粗化过程中不同
粒径颗粒出现的几率

栗夫雪兹等人对新相颗粒粗化的问题进行了更为严密的理论处理, 除利用格林伍德考虑的条件外, 还假设新相初始颗粒尺寸为一狭窄的高斯分布。根据他们的计算结果, 新相颗粒系统在粗化后不同尺寸颗粒出现的几率如图 18-33 所示, 系统颗粒的平均粒径 r_a 、单位体积中颗粒的数量 N_v , 以及系统母相中残余过饱和度 Δc 随时间变化的关系分别为:

$$r_a^3 - r_{a0}^3 = \frac{8}{9} \frac{D\sigma v c(\infty)}{k_B T} t \quad (18-69)$$

$$N_v = \frac{k_B T}{4\sigma D v c(\infty)} t^{-1} \quad (18-70)$$

$$\Delta c = \left[\frac{D k_B^2 T^2}{9\sigma^2 v^2 c^2(\infty)} \right]^{-1/3} t^{-1/3} \quad (18-71)$$

比较式 (18-68) 和式 (18-69) 可以看出它们仅有在常数上的差异。上述栗夫雪兹等人导出的关系式已被大量实验结果所证实。

18.3.3 相变动力学速率形式理论

正如前面所述, 非匀相相变从动力学机制方面可涉及到成核、生长和粗化等重要环节, 它们各有其明确的物理模型。然而在相变的实际过程中, 所有这些环节往往是难以截然分开的, 它们可以以相互伴随的方式发生于整个相变过程。因为当第一批稳定胚核开始出现时, 生长也就随之开始了, 而生长的开始并不意味着成核的停止。此外, 当系统中一旦出现不同尺寸的新相区, 粗化的条件也就已经具备了。因此, 仅仅从物理模型给出各个环节的动力学描述仍是难以得出系统在相变过程中新相的宏观转化率与相变时间的关系, 即相变动力学速率。为了建立在相变动力学研究上具有十分实际意义的新相转化率与时间的明确关系, 人们作过许多努力并发展了在很大程度上与相变机制模型细节无必然联系的形式理论。本节将就这方面的内容作简略的介绍。

(1) 恒温转变的 JMA 方程 现考察一个均匀的母相 α 在恒温的条件下通过成核、长大等过程转变成新相 β 的情况。为简单起见, 假设新相成核均匀地发生于母相中, 且在相变过程中体积成核率 I_v 为一常数; 同时假设新相的生长各向同性, 其线速度 u 不随时间变化。由此, 系统中某一新相区在 τ 时刻成核并开始长大, 在 t 时刻其体积可写成为:

$$\Gamma = f u^3 (t - \tau)^3 \quad (18-72)$$

式中, f 为一与新相区形状有关的几何因子, 对椭球形有 $f = 4\pi/3$, 对立方

体则 $f=1$ 。此系统在 $\tau \rightarrow \tau + d\tau$ 时间间隔内出现的新核数量为: $dN_v = I_v V_a d\tau$, 因而新相长大的体积增量为:

$$dV_\beta = f I_v V_a u^3 (t - \tau)^3 d\tau \quad (18-73)$$

式中, V_a 和 V_β 分别为母相和新相在 τ 时刻的体积。现考虑在相转变初期阶段, 成核过程可随机均匀发生, 新相区的生长彼此不存在干扰, 并有 $V_a \approx V$ (系统的总体积)。此时对上式进行积分, 得时间 t 时系统中新相的转变体积:

$$V_\beta = \int_0^t f I_v V_a u^3 (t - \tau)^3 d\tau \approx f I_v V u^3 \frac{t^4}{4} \quad (18-74)$$

显然, 对于球形新相颗粒, 在 t 时刻的体积转化分数则为:

$$\zeta = \frac{V_\beta}{V} \approx \frac{\pi}{3} I_v u^3 t^4 \quad (18-75)$$

然而, 应该指出, 上述导出的转化率-时间关系式仅仅适用于相转变的初期。事实上随着相转变的进行, 不同胚核长大成的新相区会出现相互干扰和碰挤。当两个新相区发生相互碰挤时, 在固态相变的情况中应在两区间发展出一个公共的边界而不是相互贯穿, 其生长在公共界面处将停止, 而在其他地方仍可继续正常生长。约翰逊(W.A.Johnson), 迈尔(R.F.Mehl)和阿弗拉米(M.Avrami)首先从统计的角度成功地解决了这一相转变过程中几何学的问题。

在JMA的理论中, $dN_v = I_v V_a d\tau = I_v (V - V_\beta) d\tau$ 仍为基本的关系式, 但把式中 $I_v V_\beta d\tau$ 一项称作为“幻想核”数, 并将包括“幻想核”数在内和不考虑新相因生长互相碰挤的情况所得出的如下新相总体积称为“扩展体积”:

$$V_{ex} = \int_0^t \Gamma I_v d\tau = \int_0^t f I_v u^3 (t - \tau)^3 d\tau = \frac{f}{4} I_v u^3 t^4 \quad (18-76)$$

显然, 在扩展体积中包括了如图 18-34 所示区域交叠的重复部分。若将图中为斜线覆盖一次的区域体积记为 V_1 , 覆盖两次的记为 V_2 , 覆盖 n 次的记为 V_n , 则有:

$$V_{ex} = \sum_i i V_i, V_\beta = \sum_i V_i \quad (18-77)$$

在相变的过程中, 若 V_β 增大至 $V_\beta + dV_\beta$, 同时相应的 V_{ex} 增大至 $V_{ex} + dV_{ex}$ 。如果新相成核属完全随机, 则由统计的考虑它们之间应存在如下关系:

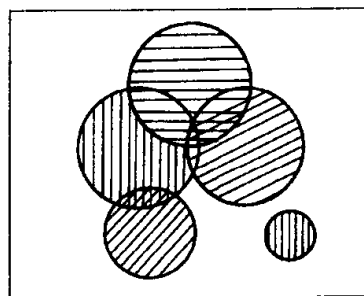


图 18-34 系统中相变区域相互碰撞和交叠的情况

$$\frac{d(V_\beta/V)}{dV_{ex}} = 1 - \frac{V_\beta}{V} \quad (18-78)$$

于是可以得到 JMA 精确处理的相转变率-时间关系式:

$$\zeta = \frac{V_\beta}{V} = 1 - \exp(-V_{ex}) = 1 - \exp\left(-\frac{fI_v u^3 t^4}{4}\right) \quad (18-79)$$

容易看出, 如新相区为球形, 则有

$$\zeta = 1 - \exp(-\pi I_v u^3 t^4/4) \quad (18-80)$$

更一般地称为约翰逊-迈尔-阿弗拉米方程 (或 JMA 方程) 为:

$$\zeta = 1 - \exp(-Kt^n) \quad (18-81)$$

应该指出, 一般而言, 假设成核率在整个相变过程中恒为常数以及假设新相区生长速率的各向同性并不总是正确的。若考虑到成核率随时间的变化以及生长速率的各向异性, 式 (18-79) 应改写成:

$$\zeta = 1 - \exp\left[-fu_x u_y u_z \int_0^t I_v(t-\tau)^3 d\tau\right] \quad (18-82)$$

式中, u_x 、 u_y 、 u_z 分别为新相区在 x 、 y 、 z 等方向的主生长速率。显然, 成核率随时间增加或减少将引起转化率-时间关系式中指数 $n \neq 4$ 。

现考察成核率变化的一种情况: 成核只发生在系统的某些有利位置, 在母相的单位体积中原有 N_0 个这种位置, t 时间后剩下 N 个, 在 $t+dt$ 时间段中消耗 $dN = -N\nu dt$ 。其中频率 ν 给出了单个有利位置变成核的速率。因此有: $N = N_0 \exp(-\nu t)$, 即单位体积的成核速率为:

$$I_v = -\frac{dN}{dt} = N_0 \nu \exp(-\nu t) \quad (18-83)$$

将此式代入式 (18-82), 并积分可得:

$$\zeta = 1 - \exp\left\{8\pi N_0 u^3 / \nu^3 \left[\exp(-\nu t) - 1 + \nu t - \frac{\nu^2 t^2}{2} + \frac{\nu^3 t^3}{6}\right]\right\} \quad (18-84)$$

讨论上式的两个极限是颇有意义的。当 νt 很小时, 意味着成核率几乎为一常数, 此时上式展开处理后得到的极限形式与式 (18-80) 完全相同, $n = 4$; 与此相反, 当 νt 很大时, N 可非常迅速地降为零, 即成核中心在相变的早期就已完全耗尽, 于是式 (18-84) 相应的极限形式为:

$$\zeta = 1 - \exp\left(-\frac{4\pi N_0 u^3 t^3}{3}\right) \quad (18-85)$$

由此可见, 当成核率为常数或为时间的有限减函数时, JMA 方程一般式中的指数应为: $3 \leq n \leq 4$ 。当然, 上述讨论的情况虽均属于三维成核和长大的过程, 但对于二维生长或一维生长过程的转化率时间关系也可容易地从式 (18-81) 中导出。此时 JMA 方程一般式中时间指数 n 分别为 $2 \leq n \leq 3$ (二

维生长) 和 $1 \leq n \leq 2$ (一维生长)。

对于在固态相变中晶界、晶棱和晶角处成核长大恒温相转变动力学速率, 卡恩于 1956 年在沿用了 JMA 理论中“扩展体积”概念的基础上进行了详尽的分析。根据他的分析结果, 在晶粒尺寸不太小的情况下, 晶界面或晶棱上的核心很快便会出现饱和状态, 此时新的核心已不再能够形成, 因而新相的生长只能是通过附在母相晶面上的新相向母相晶粒内部平行推进, 或是由位于晶棱上的新相沿其径向粗化的方式进行。根据卡恩的计算, 此时相应于 JMA 方程中的时间指数分别为 1 和 2。

(2) 长程扩散控制的恒温转变速率 现在考察长大速率由长程扩散所控制的相转变动力学速率。根据式(18-56)和式(18-70), 一个在 τ 时刻成核的球形新相区在 t 时刻的体积应为:

$$V_\tau = \frac{4\pi}{3} \alpha_3^3 D^{3/2} (t - \tau)^{3/2} \quad (t > \tau) \quad (18-86)$$

对于一个扩散控制的转变过程, 如连续脱溶过程, 转变率 ζ 的定义必须重新考虑。设 V_β 为整体体积 V 中新相的平衡体积, $V_\beta(t)$ 为新相的实际转变体积。因此, 为保持平衡时必须的溶质原子浓度, 并定义转变率 $\zeta = V_\beta(t) / V_\beta$, 则有:

$$V_\beta / V = (c_m - c_a) / (c_\beta - c_a) \quad (18-87)$$

$$\zeta = \frac{V_\beta(t)}{V} [(c_\beta - c_a) / (c_m - c_a)] \quad (18-88)$$

对于随机分布均匀成核早期饱和的情况, 即在 $t = 0$ 母相单位体积中已存在核心数 N_v 并不再有新核心形成, 其转变率在无碰撞情况下可由下式简单地决定:

$$\zeta = N_v V_c (c_\beta - c_a) / (c_m - c_a) = K t^{3/2} \quad (18-89)$$

对于更一般的情况并考虑新相区间存在相互碰挤, 引用 JMA 理论的原理, 新相在 t 时间的转化率应为:

$$\zeta = 1 - \exp \left[- \frac{(c_\beta - c_a)}{(c_m - c_a)} \frac{4\pi}{3} \alpha_3^3 D^{3/2} \int_0^t (t - \tau)^{3/2} I_v d\tau \right] \quad (18-90)$$

容易看出, 当成核速率 I_v 为常数时, 上式积分后 JMA 一般式中的时间指数 $n = 5/2$ 。对式(18-88)作碰撞处理, 可得与上式相似的结果。

然而, 应该指出, 式(18-88)实际上仅适用于长程扩散控制的相变初期。这是因为随着新相区的长大, 母相中溶质原子的浓度场会逐渐受到不同生长区所谓“软碰撞”的干扰。沃特 (C. Wert) 和曾纳首先给出了球形颗粒生长“软碰撞”的近似处理。他们考虑一个低度过饱和母液中的沉淀过程。假设所有的核胚一开始即已存在, 新相生长界面的推移是足够地缓慢以近似达

到一种浓度场的稳态过程 (即 $\nabla^2 c = 0$)。此时根据连续性方程, 有:

$$(c_\beta - c_\alpha) \frac{dr}{dt} = D \frac{[c_\infty(t) - c_\alpha]}{r} \quad (18-91)$$

式中, $c_\infty(t)$ 为时间 t 时远离新相生长界面的溶质原子浓度。现设最终生长区的尺寸为 R , 则转化率 $\zeta = (r/R)^3$, 并有关系式 $c_\infty(t) - c_\alpha = (c_m - c_\alpha)(1 - \zeta)$ 。将之代入式(18-91)得:

$$\frac{d\zeta}{dt} = \frac{3D(c_m - c_\alpha)}{R^2(c_\beta - c_\alpha)} \zeta^{1/3}(1 - \zeta) \quad (18-92)$$

沃特和曾纳将式(18-92)之解与式(18-88)在考虑碰撞处理后得到的结果

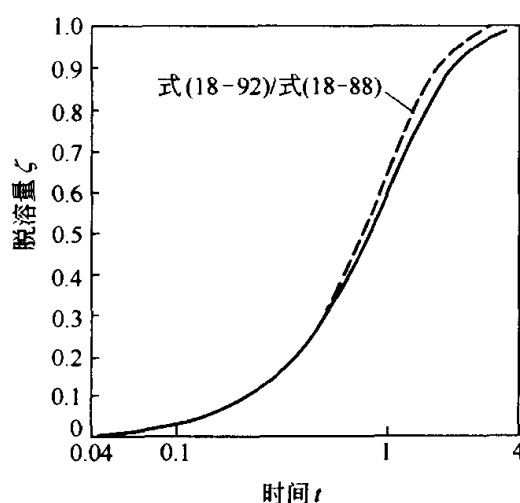


图 18-35 扩散长大中“软碰撞”与 JMA 理论相转变率曲线比较

进行了定量的比较, 结果如图 18-35 所示。可以看出, 在转变的初期, 它们之间并不存在明显差异。随着相变的进行, 差异逐渐明显并持续至相转变的后期。

综合以上各种条件下相转变率-时间依赖关系的分析, 可以看到, 尽管 JMA 方程一般式不易达到描述各种相变的动力学全过程, 但其中的指数生长律对大多数情况下的线性生长是有效的, 并且也适用于扩散控制的早期生长。克利斯汀 (J.W.Christian) 曾用该方程模拟各种相变的进程, 得到了如表 18-1 所概括的不同情况下的

n 指数, 它对于估计具有不同相变机制的相变速率有一定的帮助。

表 18-1 JMA 方程在各种转变机制中的 n 值

①多型性转变, 不连续脱溶, 共析反应, 界面控制的生长等	条 件	n	②长程扩散控制型生长	条 件	n
	成核率随时间增加	>4		由小尺寸开始生长, 成核率增加	$>5/2$
	成核率不随时间变化	4		由小尺寸开始生长, 成核率为常数	$5/2$
	成核率随时间减少	$3 \sim 4$		由小尺寸开始生长, 成核率减少	$\frac{3}{2} \sim \frac{5}{2}$
	成核率为零(位置饱和)	3		由小尺寸开始生长, 成核率为零	$3/2$
	晶棱成核饱和后	2		初始尺寸较大颗粒的生长	$1 \sim 3/2$
	晶面成核饱和后	1		稀疏有限尺寸针状、片状沉淀生长	1
				长圆柱状沉淀加粗	1
				大片状沉淀增厚	$1/2$
				位错线上沉淀	$2/3$

(3) 时间-温度-转变(TTT)图 以上所讨论的相转变-时间动力学关系均是对某一恒定温度下而言的。实际上改变相变发生的温度, 将可以得到不同的转变率-时间关系。如果我们在不同的温度下, 通过改变时间使系统的相转变率达到某一组预先选定的值, 我们则可以在时间-温度平面上得到相应的一组“C”形曲线, 这就是通常所谓的 TTT 图。与平衡相图不同, TTT 图是描写某种材料在不同温度下、不同时间相变产物的结构及其数量的图, 是一种动力学相图。图 18-36 为一共析碳钢的 TTT 图。图中珠光体开始形成与最后完成所相应的曲线都具有“C”的形状。这是因为在较高温度转变时, 虽然扩散系数较大, 但系统的过冷度不大, 故转变速度较慢; 反之, 在较低的温度下, 系统的过冷度虽然较大, 但扩散系数会变得很小, 于是转变速度也将是较慢的。正如图中所示, 对于共析碳钢只有在 550°C 左右, 相变驱动力与扩散两个因素的综合作用才使系统的相转变速度达到了最大值。这一基本的规律对于各种受扩散控制的相变过程有着相当的普适性。

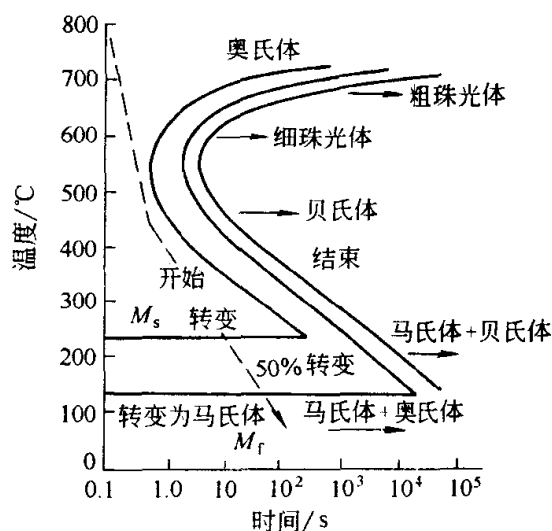


图 18-36 共析碳钢的 TTT 图

TTT 图在钢铁工业中钢铁和其他金属材料热处理方面有着重要的实际应用价值。这是因为通过恒温测定得到的 TTT 图可以作为变温相变过程描述的一种有效近似。在这个意义上, 将材料系统所经历的连续变温过程曲线 $T(t)$ 绘于其 TTT 图上, 根据变温曲线与等转变率曲线 $Q(t, T) = q$ 的位置关系, 便可得知系统经历变温过程后所具有的组织结构信息。另一方面, 利用 TTT 图人们可以设计和编制出一定的材料热处理程序以获得所需要的特定组织结构。如图 18-36 所示, 若期望获得马氏体组织, 在材料的冷却过程中, 就必须采用不与珠光体开始转变的那根等转变率曲线相交的冷却曲线。图 18-36 中所示的虚线是得到马氏体组织的一条快速冷却曲线。

TTT 图的另一大应用是由伍尔曼 (D.R. Uhlmann) 首先将 TTT 图的方法应用于无机氧化物玻璃形成动力学理论研究中而得到开发的, 这一方法在后来解决金属玻璃的形成和稳定性问题方面得到了同样极有意义的应用。伍尔曼认为, 判断一种物质能否形成玻璃, 首先必须确定玻璃中可以检测到的晶体的最小体积, 然后再考虑熔体需要多快的冷却速度才能防止这一结晶量的产生。实验表明, 当晶体混乱分布于熔体中时, 晶体不被检测到的

TTT 图的另一大应用是由伍尔曼 (D.R. Uhlmann) 首先将 TTT 图的方法应用于无机氧化物玻璃形成动力学理论研究中而得到开发的, 这一方法在后来解决金属玻璃的形成和稳定性问题方面得到了同样极有意义的应用。

伍尔曼认为, 判断一种物质能否形成玻璃, 首先必须确定玻璃中可以检测到的晶体的最小体积, 然后再考虑熔体需要多快的冷却速度才能防止这一结晶量的产生。实验表明, 当晶体混乱分布于熔体中时, 晶体不被检测到的

体积分数上限为 10^{-6} 。根据 JMA 形式理论, 当转变分数很小时有:

$$\zeta = \frac{\pi I_v}{3} u^3 t^4 \quad (18-93)$$

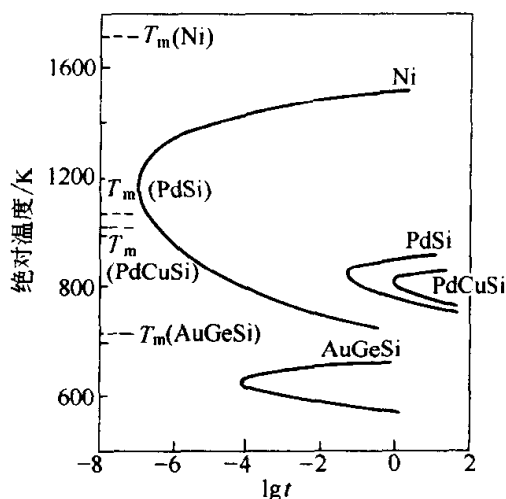


图 18-37 析晶体积分数为 10^{-6} 时不同熔点物质的 TTT

于是, 当知道熔体在不同温度下的成核率 I_v 和晶核生长速度 u , 就可利用上式并令 $\zeta = 10^{-6}$, 作出相应熔体结晶的 TTT 图线。图 18-37 给出了这类图线的实例, 图中“C”形曲线的内侧为熔体在一定过冷度下形成晶体的区域, 在“C”形曲线的外侧则是一定过冷度下仍为熔体的区域, “C”形曲线头部顶点对应于析出体积分数为 10^{-6} 晶体时的最短时间。因此, 为避免形成给定的晶体分数, 熔体的冷却速度必须大于某一临界速度 U_c , 其数值大小可由下式给出:

$$U_c = \frac{T_m - T_n}{t_n} \quad (18-94)$$

式中, T_m 、 T_n 分别为材料的熔点和 TTT 图中 C-曲线上“鼻尖”处对应的温度, t_n 为与“鼻尖”相应的时间。形成玻璃的临界冷却速度主要是随熔体组成而变化的。表 18-2 列举了几种化合物和金属形成玻璃态时的临界冷却速度和熔融温度时的粘度。

表 18-2 几种氧化物和金属形成玻璃的有关参数

性 能	化 合 物									
	SiO ₂	GeO ₂	B ₂ O ₃	Al ₂ O ₃	As ₂ O ₃	BeF ₂	ZnCl ₂	LiCl	Ni	Se
$T_m/^\circ\text{C}$	1710	1115	450	2050	280	540	320	613	1380	225
$\eta/\text{dPa}\cdot\text{s}$	10^7	10^6	10^5	0.6	10^5	10^6	30	0.02	0.01	10^3
$U_c/^\circ\text{C}\cdot\text{s}^{-1}$	10^{-6}	10^{-2}	10^{-6}	10^3	10^{-5}	10^{-6}	0.1	10^8	10^7	10^{-3}

由表 18-2 可以看出, 凡是熔体在熔点时具有很高的粘度, 并且粘度随温度的降低而急剧增加的氧化物 (如 SiO₂、B₂O₃ 等), 是很容易形成玻璃态结构的, 它们的临界冷却速度非常低。而相比之下, 在熔点时粘度很小的熔体 (如 Al₂O₃ 和金属 Ni), 却很容易结晶。故若要它们形成玻璃态结构, 就必须有很高的冷却速度才能避免结晶的发生。为了能在适当的冷却速度下获得玻璃态的材料, 考虑以下诸影响因素可以得到有效的启发。

① 增大结晶成核势垒，从而降低成核速率。为此应将熔体内导致非均匀成核的异相杂质尽量清除干净。此外由式(18-37)和式(18-38)可知，结晶潜热较小的材料具有较大的成核势垒，固-液界面能高的材料成核势垒也较大。因此这些材料易于获得玻璃态结构。

② 增大原子由熔体进入晶核所必须克服的长大动力学势垒，从而减小晶体生长的速度。其措施为选择熔体粘度高并随过冷度加大而快速增加的材料，或从该方面对材料进行改性，使之具备这些性质。

③ 如果结晶时涉及原子作较大程度的重新分配，例如共晶成分的熔体析晶，复杂的扩散因素将会对晶体的生长速度予以较大的限制，从而有利于材料获得玻璃态结构。

④ 熔点较低的材料有时易于形成非晶态。考虑到熔点 $T_m = \Delta H / \Delta S$ ，低的熔点意味着较小的能量可以造成较大的无序程度。这也意味着，熔体转变为晶态的驱动力较小，因而易于形成非晶态。

18.3.4 成核-生长和失稳分解过程

在本章 18.2.5 节已经提及，对二元（或多元）的匀相系统，当其组成-自由焓曲线在某一温度下出现如图 18-38 所示的情况时，组成位于 c_a 和 c_b 区间内的单相系统将变得不稳定而趋于出现所谓的失稳分相过程。这是因为组分在这个区间的均匀单相比分解成组成分别为 c_a 和 c_b 的两相有更高的自由能。根据系统具体组成点位置的不同，存在两种不同类型的分相机制：成分在 s_1 左侧和在 s_2 右侧的分相将经历成核-长大过程，而成分在 s_1s_2 间的分相将无需经历成核就可自发分解。这种失稳分相过程的可能性早在 19 世纪就被吉布斯在理论上所洞察，但直到 20 世纪 60 年代随着微结构观测分析技术的发展才终究获得证实。可能发生失稳分相过程的材料系统可以是结晶态的均匀固溶体和结构无序的均匀玻璃态物质。前者在金属合金系统中不乏其例，如 Al-Ag 和 Al-Cu 系统等。后者首先是在硼硅酸盐系统玻璃中被发现。对 $0.75\text{SiO}_2\cdot 0.20\text{B}_2\text{O}_3\cdot 0.05\text{Na}_2\text{O}$ 系统的玻璃作 $500\sim 600^\circ\text{C}$ 的热处理，将使原来均匀的玻璃分成截然不同的两相。其中一相几乎是纯 SiO_2 ，而另一相为富 Na_2O 的 B_2O_3 。这种分相的产物经酸处理后可制成含 $4\sim 15\text{nm}$ 微孔的纯 SiO_2 技术玻璃。

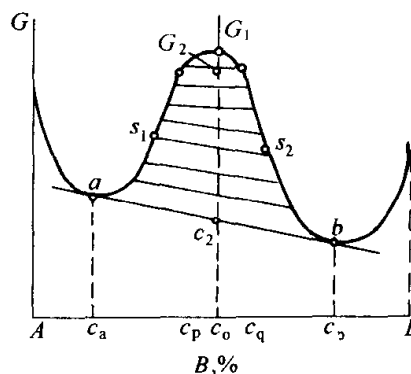


图 18-38 可分相二元固溶体
组成-自由能曲线

(1) 成核-生长机制的分相 如图 18-39 所示，有一组成位于 s_1 左侧 c_0

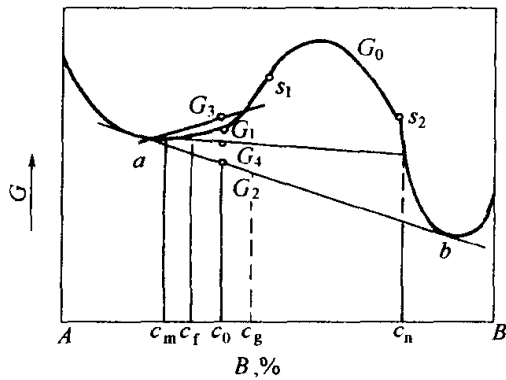


图 18-39 连续固体中成核-长大机制分相示意图

处的固溶体, 其自由能 G_1 大于分相后两相混合物自由能 G_2 , 故分相将有利于系统自由能的降低而达到更稳定的状态。若系统浓度起伏造成微区内成分偏离为 c_f 及 c_g , 容易看出, 由于在该区间 $(d^2G/dc^2) > 0$, 自由能曲线呈现下凹, 而使成分微小偏离后系统的自由能 G_3 高于 G_1 。因而这种不稳定的浓度涨落将自发地使系统恢复到原来的状态。为了使系统分相成功, 只有当浓度涨落足够大时 (如 c_m 、 c_n), 系

统的自由能才能降低 ($G_4 < G_1$), 分相方可自发进行直至达到平衡成分 c_a 和 c_b 止。下面就分相所需的最小浓度涨落和其分相过程特点作一简单的讨论。

如图 18-40 所示, 组成为 c_0 的系统涨落为 c_1 、 c_2 时, 系统的自由能变化由杠杆原理可得:

$$\begin{aligned}\Delta G &= (n_1 G_1 + n_2 G_2) - (n_1 + n_2) G_0 \\ &= n_2 \left[(G_2 - G_1) - (c_2 - c_0) \frac{G_0 - G_1}{c_0 - c_1} \right] \quad (18-95)\end{aligned}$$

当成分为 c_2 的分相产物很少时, $c_1 \rightarrow c_0$, 故 $(G_0 - G_1)/(c_0 - c_1) \rightarrow (dG/dc)_{c_0}$ 。于是, 上式改写成:

$$\Delta G_v = \Delta G / n_2 = \left[(G_2 - G_0) - (c_2 - c_0) \left(\frac{dG}{dc} \right)_{c_0} \right] \quad (18-96)$$

比较图 18-40 可得: $G_2 - G_1 = AE$, $(c_2 - c_1)(dG/dc)_{c_0} = BE$, $\Delta G_v = AE - BE = AB$ 。显然, 当 $\Delta G_v > 0$ 时, 任何浓度涨落将会很快地被消除, 仅当 $\Delta G_v < 0$ 时, 涨落才可能得到助长。因此对应的浓度涨落最小值应由

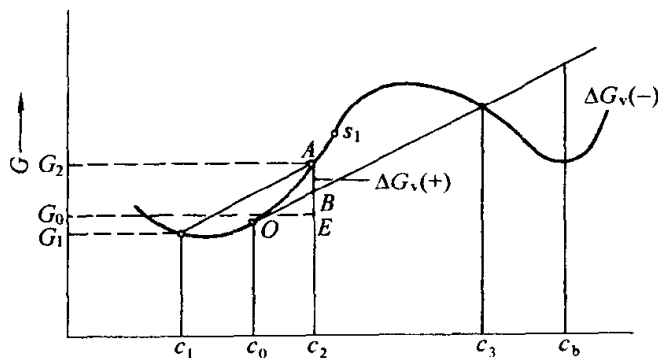


图 18-40 成核长大机制浓度涨落最小值

$\Delta G_v = 0$ 确定：过与组成 c_0 对应的 O 点作自由能-组成曲线的切线并延长与自由能曲线相交，得由交点所对应的组成点 c_3 。即，当浓度涨落大于 c_3 时，更大的浓度起伏将会由于 $\Delta G_v < 0$ 而被自发地得到发展直至达到组成点 c_b ，此时可形成组成为 c_b 的稳定胚核。胚核的形成将使其周围 B 组分浓度迅速下降，进而胚核的生长需要通过扩散从母相得到 B 组分。如图 18-41(a) 所示，这是一个由高浓度向低浓度的扩散，其结果将是胚核的不断长大直至分相的最后完成。不难理解，这一分相的特点是起始浓度涨落大而涉及的空间范围小，第二相从胚核的产生就具有恒定的组成并与母相间存在明确的界面，其成核-长大动力学过程可由前面所讨论过的理论进行描述。

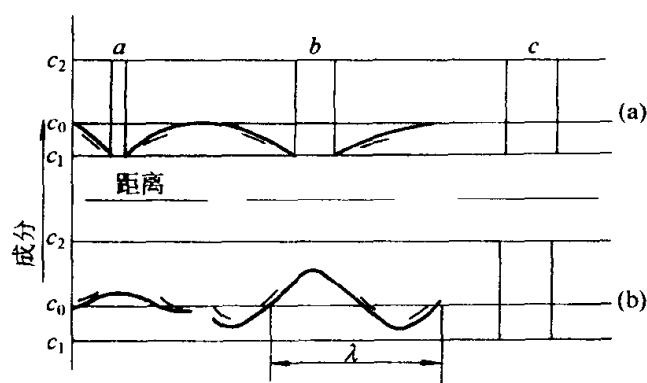


图 18-41 连续固溶体中成核长大、失稳分相浓度分布示意图

(a) 成核长大机制；(b) 失稳分相机制

(2) 失稳分解 如图 18-38 中所示组成-自由能曲线，其连续性要求存在二阶导数 $(d^2G/dc^2) = 0$ 的两个数学上定义为拐点或旋点的组成点 s_1 和 s_2 。当系统的组成点位于 s_1s_2 之间，分相过程将具有完全不同于上述的成核-长大机制。在此区间，系统具有 $(d^2G/dc^2) < 0$ 而使其自由能曲线呈现上凸。于是，组成为 c_0 的固溶体内部发生微小的浓度起伏或偏离 c_p 和 c_q ，系统的自由能将由 G_1 降为 G_3 。不难看出，在此区间内具有任何组成点的系统内部发生任何微小的浓度起伏都将导致系统自由能的下降。因此，此区间内的分相将无需任何临界胚核，浓度起伏一旦开始就将迅速席卷整个系统的全部范围，并逐步完成分相过程直至达到平衡浓度 c_a 和 c_b 为止。将不同温度下旋点的位置在相图中用虚线连接起来（如图 18-42 中所示）形成失稳线。过冷到失稳线以下的具有组成在失稳区内的固溶体发生的分相常称为失稳分解，该分相过程具有如下特点。

① 失稳分解相属于吉布斯第二类涨落所导致的相变。如图 18-41(b) 所示，在相变刚开始时，系统中存在程度甚微而范围甚广的浓度起伏。由于系统自由能随着此类涨落的发生而出现与其幅度成正比地连续减小，从而促使

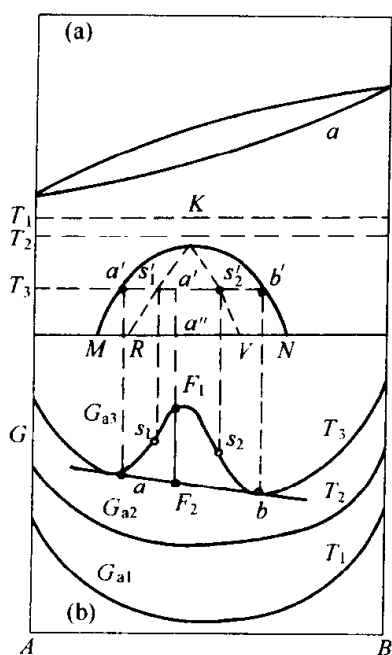


图 18-42 连续固溶体中失稳区的失稳分相

系统中发生幅度越来越大的成分涨落并最终分解成两相。在分相的过程中，成分在空间的涨落可以理解成一系列成分调制波的叠加，波峰与波谷所在的区域意味着分相过程中产生的两个不同的物相。在分相完成之前，它们之间无明晰的界面，系统中的某一成分将连续地从波谷向波峰扩散并使其在波峰区不断富集。这种扩散显然不同于通常负浓度梯度方向上的正扩散，故称之为负扩散或爬坡扩散。

② 实验观测表明，绝大多数可发生失稳分相的固溶体第二相与基相有着结构上共格的关系，尤其对于合金系统。但分相首先是带来成分上的差异，它不可避免地导致固溶体晶格参数的不同。所以，共格关系的保留必然产生晶格的畸变和产生不利于失稳分相进行的晶格应变能。图 18-43 示意地给出了分相过程中可能出现的三种不同晶格关系。一般情况下，由于晶格弹性模量的各向异性，晶格应变能在不同结晶学方向上的差异将可能导致失稳分相时成分偏聚具有一定的方向性。不难理解，晶格应变能的出现将要求失稳分相的发生有更大的过冷度。

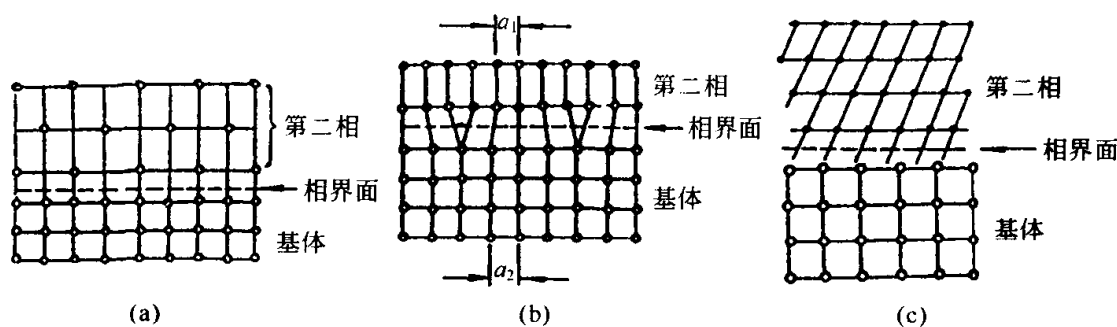


图 18-43 连续固溶体中可能出现的三种晶格关系

(a) 完全共格；(b) 部分共格；(c) 非共格

③ 失稳分相的发生使原来成分均匀的系统开始出现分布于整个系统的浓度梯度。在出现浓度梯度较大的地方正是系统中溶质浓度最大处过渡到最小处的区域。这些区域可以类比于两相之间的弥散界面，因而将出现相应于界面能的附加能量——浓度梯度能，它的出现将部分抵消相变的驱动力。成分调制波波长越短，梯度能越大，实际有效相变驱动力下降越多。

18.3.5 失稳分解的动力学方程

正如上节中所作的分析, 当均匀的固溶体或玻璃态物质的组成-自由能间关系满足 $(d^2G/dc^2) < 0$ 条件时, 系统对于吉布斯第二类涨落将失去稳定而出现具有成分调制结构的失稳分解。人们在 20 世纪 40 年代首次于 Cu-Ni-Fe 永磁合金中发现了这种分相的迹象, 随后卡恩等人发展并建立了失稳分解的动力学理论。这一理论成功地解释了失稳分解产生的成分调制结构, 阐明了材料的弹性各向异性对分相产物形态的影响, 在相变动力学研究方面标志着一个意义深远的重大进展。本节将对此理论作简略介绍。

(1) 非均匀连续介质的自由能密度 具有成分调制的系统中, 浓度不再是均匀恒定的量, 各组分浓度在空间的非均匀分布使浓度梯度成为研究系统性质的不可忽视的重要因素。这是因为系统组分的浓度分布联系着组分间的相互作用关系, 从而影响着系统的能量状态。现以一 AB 固溶体组成沿 x 方向发生变化为例, 考察其浓度分布对系统能量状态的附加影响。

如图 18-44 所示, 固溶体由一系列原子平面所组成, 各原子平面中 A 的浓度分别为 c_{p-1} 、 c_p 、 c_{p+1} ……。设 N 为原子面密度; Z 为每个原子和相邻原子成键数; E_{AA} 、 E_{BB} 、 E_{AB} 分别为 AA、BB 和 AB 原子对相互作用能。于是若仅考虑最邻近相互作用, 并令 $\Delta E = 2E_{AB} - E_{AA} - E_{BB}$, 可得系统组成浓度梯度引起的附加能量 E_S :

$$E_S = \frac{NZ}{2} a_0^2 \Delta E \left(\frac{c_p - c_{p+1}}{a_0} \right)^2 = K (\nabla c)^2 \quad (18-97)$$

式中, ∇c 为扩展到三维情况下的浓度梯度, 常数 $K = NZa_0^2 \Delta E / 2$ 。容易看出, 由于 $(\nabla c)^2$ 总为正值, 因而附加能量 E_S 的正负完全决定于常数 K 。当 $\Delta E > 0$ 即 $E_{AB} > (E_{AA} + E_{BB})/2$, 则有 $E_S > 0$ 。这说明当系统在某一温度下, 出现 AA 和 BB 原子间平均作用能小于 AB 原子间作用能时, 系统将趋于由原来均匀混合的 AB 单相分解成富 A 和富 B 的两相, 此时浓度梯度的存在将引起系统能量的增加; 反之若 $\Delta E < 0$, 则有 $E_S < 0$, 此时 $E_{AB} < (E_{AA} + E_{BB})/2$, 系统趋于自发均匀混合, 浓度梯度的存在使系统能量降低。

相对于成分均匀系统, 非均匀系统引起的附加能量还包括原子尺寸不同而产生的晶格畸变能。作为一级近似, 可以认为固溶体的晶格常数与其组成

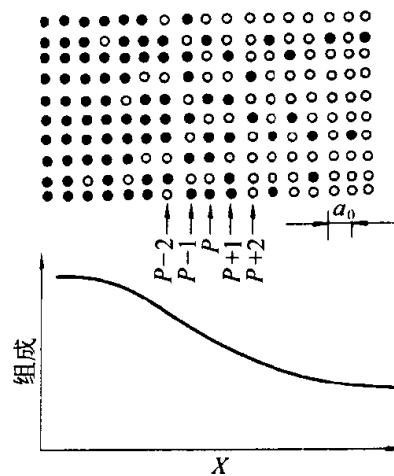


图 18-44 X 方向原子面结构与组成变化示意图

浓度成线性关系，即：

$$a - a_0 = (c - c_0) \eta a_0 \quad (18-98)$$

式中， a 为浓度为 c 时固溶体的晶格常数， a_0 为当系统具有平均浓度 c_0 时的晶格常数， η 为不同原子间的失配度。

为简单起见，考虑晶体具有弹性各向同性，并设沿 x 方向的成分按正弦波进行调制。此时， x 方向上晶格常数平均值为 a_0 ，因而不出现沿 x 方向的正应力，即 $e_{xx} = 0$ 。为保持共格， y 和 z 方向的正应变应正好抵消成分调制所引起的晶格常数的变化，即有：

$$e_{yy} = e_{zz} = -\frac{a - a_0}{a_0} = -\eta(c - c_0) \quad (18-99)$$

由对称性的考虑可知，系统中各切应力为零。根据上式及 $e_{xx} = 0$ ，由弹性力学的基本方程不难求出此时系统中共格畸变能量密度为：

$$E_T = \frac{E\eta^2}{1-\nu}(c - c_0)^2 \quad (18-100)$$

式中， E 为介质的杨氏模量， ν 为泊松比。综合考虑浓度梯度和共格畸变能对系统能量附加贡献后，可得非均匀连续介质总自由能密度应为：

$$g_s = g(c) + \frac{E\eta^2}{1-\nu}(c - c_0)^2 + K\left(\frac{dc}{dx}\right)^2 \quad (18-101)$$

上式中第一项为均匀介质的自由能密度，第二项为共格畸变能量密度，第三项则为浓度梯度对系统能量密度的贡献。

(2) 失稳分解动力学方程 在式(18-101)中，我们得到了非均匀连续介质的自由能密度。因此，系统的自由能总量应有下面积分式给出：

$$G_\Sigma = A \int \left[g(c) + \frac{E\eta^2}{1-\nu}(c - c_0)^2 + K\left(\frac{dc}{dx}\right)^2 \right] dx \quad (18-102)$$

当系统达到平衡时，上式自由能应取极小值。根据数学上变分原理，若考虑系统质量守恒条件 $\int (c - c_0) dx = 0$ 作为约束，上式的极值问题则相当于用拉格郎日不定乘子法解下面欧拉方程：

$$\frac{d}{dx} \left[\frac{\partial I}{\partial \left(\frac{dc}{dx} \right)} \right] - \frac{\partial I}{\partial c} = 0 \quad (18-103)$$

式中

$$I = g(c) + \frac{E\eta^2}{1-\nu}(c - c_0)^2 + K\left(\frac{dc}{dx}\right)^2 - \alpha(c - c_0) \quad (18-104)$$

其中 α 为拉格郎日不定乘子。在 $E\eta^2/(1-\nu)$ 和 K 均与浓度 c 无关的情况下，由式(18-103)可解得：

$$\alpha = \frac{\partial g(c)}{\partial c} + 2 \frac{E\eta^2}{1-\nu}(c - c_0) - 2K \frac{d^2c}{dx^2} \quad (18-105)$$

根据扩散热力学理论, 有:

$$\frac{\partial c}{\partial t} = \frac{M}{N} \frac{\partial}{\partial t} \left[\frac{\partial^2 g(c)}{\partial c^2} \right] \frac{\partial^2 c}{\partial x^2} \quad (18-106)$$

将式 (18-105) 中 α 代替上式中 $\partial^2 g / \partial c^2$, 便得共格非均匀介质中的扩散方程:

$$\begin{aligned} \frac{\partial c}{\partial t} &= \frac{M}{N} \frac{\partial}{\partial c} \left[\frac{\partial g(c)}{\partial c} + 2 \frac{E\eta^2}{1-\nu} (c - c_0) - 2K \frac{d^2 c}{dx^2} \right] \frac{\partial^2 c}{\partial x^2} \\ &= D^* \left\{ \left[1 + \frac{2E\eta^2}{1-\nu} \frac{1}{\left(\frac{\partial^2 g(c)}{\partial c^2} \right)} \right] \frac{d^2 c}{dx^2} - \frac{2K}{\left(\frac{\partial^2 g(c)}{\partial c^2} \right)} \frac{d^4 c}{dx^4} \right\} \quad (18-107) \end{aligned}$$

若假设上式中 $d^2 c / dx^2$ 和 $d^4 c / dx^4$ 的系数与浓度 c 无关, 则由傅里叶变换法求得方程 (18-107) 的特解:

$$c(x, t) - c_0 = A(\beta) \exp[R(\beta)t] \exp[i\beta x] \quad (18-108)$$

显然, 上式是振幅为 $A(\beta) \exp[R(\beta)t]$ 的正弦(或余弦)周期函数, 它的各种线性组合均是方程(18-107)的解。其中 $A(\beta)$ 为 $t=0$ 时波矢为 $\beta=2\pi/\lambda$ 浓度波的振幅; $R(\beta)$ 为增幅因子, 其表达式为:

$$R(\beta) = -\frac{M}{N} \beta^2 \left[\frac{\partial^2 g(c)}{\partial c^2} + \frac{2E\eta^2}{1-\nu} + 2K\beta^2 \right] \quad (18-109)$$

不难看出, 当增幅因子 $R(\beta) < 0$, 各种波长的浓度起伏均是不稳定的, 随着时间的延长将迅速衰减消失; 反之, 当 $R(\beta) > 0$, 各种波长的浓度起伏将随时间延长而迅速增大, 此时系统将发生自发失稳分解。为使增幅因子 $R(\beta) > 0$, 应有条件:

$$\frac{\partial^2 g(c)}{\partial c^2} + \frac{2E\eta^2}{1-\nu} + 2K\beta^2 < 0 \quad (18-110)$$

此式表明, 考虑共格畸变和浓度梯度对系统自由能的附加贡献后, 为了使失稳分解得以发生, 仅仅要求 $\partial^2 g / \partial c^2 < 0$ 是不够的, 而必须有:

$$\frac{\partial^2 g(c)}{\partial c^2} < -\left(\frac{2E\eta^2}{1-\nu} + 2K\beta^2 \right) \quad (18-111)$$

式中右边第一项正比于共格畸变能, 只有当两组分的失配度 $\eta=0$ 时, 共格畸变能消失为零; 第二项正比于浓度梯度能, 当常数 K 一定时, 浓度波波长越大, 梯度能越小。因此仅当 $K=0$ 或 $\lambda \rightarrow \infty$ 时, 浓度梯度能附加贡献为零。所以, 在实际的分相系统中, 由于共格畸变能和浓度梯度能的存在而使失稳分相的有效驱动力有所下降, 它们在一定程度上缩小了分相的范围和程度。

由式(18-109) 还可看出, 欲使失稳分解得以进行, 式中右边括号项必须小于或等于零, 于是容易推得:

$$\beta \leq \beta_c = \sqrt{\frac{1}{2K} \left| \frac{\partial^2 g}{\partial c^2} + \frac{2E\eta^2}{1-\nu} \right|} \quad (18-112)$$

即存在一临界波矢 β_c ，当浓度起伏波长 $\lambda < \lambda_c$ 时，起伏将受到抑制而迅速衰减，只有当 $\lambda > \lambda_c$ 的浓度起伏才可得到加强。同时应该看到，临界波矢的大小正比于系统共格弹性畸变附加能而反比于浓度梯度附加能。因此，当

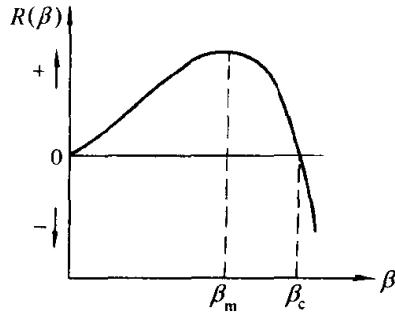


图 18-45 振幅因子 $R(\beta)$ 与波矢 β 的关系

反映浓度梯度能 K 常数足够小时，临界波长可能短至系统的晶格常数，从而可能促使晶体中出现 $ABABAB\cdots$ 的有序化现象。

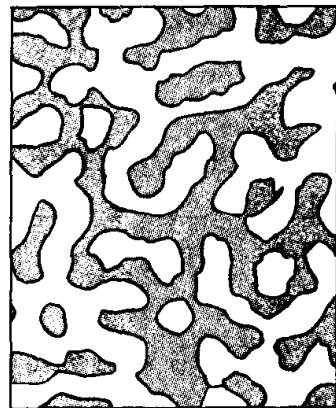
将式(18-112)代回式(18-109)并关于 β 求极值，可得当 $\beta = \beta_m = \beta_c/\sqrt{2}$ 时， $R(\beta)$ 取极大值。 $R(\beta)$ 随 β 的依赖关系如图 18-45 所示。由于浓度波振幅因子 $A(\beta, t)$ 对 β 的变化具有指数关系而极为敏感，故波矢为 β_m 的组分调制波振幅将迅速上升而处于支配地位。利用 $1/R(\beta)$ 关系可以估计失稳分解达到极大振幅所需要的最短时间。注意到式(18-109)关系以及互扩散系数 $D^* = M(\partial^2 g / \partial c^2) / N$ ，不难

推得：

$$t_m = -1/\beta_c^2 \left(\frac{ME\eta^2}{2N(1-\nu)} + \frac{D^*}{4} \right) = - \frac{\lambda_m^2}{8\pi^2 \left[\frac{ME\eta^2}{2N(1-\nu)} + \frac{D^*}{4} \right]} \quad (18-113)$$

式中， λ_m 为极大振幅浓度波所对应的波长。

对于弹性各向同性的晶态固溶体或均匀的玻璃体，失稳分相所产生的成分调制的波矢将无一定的择优方向，处于完全的随机分布。卡恩曾对这一情况的失稳分相产物形态进行了计算机模拟，他假设所有浓度波波矢大小为 β_m 、方向和振幅分别为随机分布和高斯分布。对 100 个正弦波叠加的结果表明，失稳分相的产物并不具有明显的周期，而是互相连通的海绵状组织。图 18-46 中示出了模拟分相产物的截面结构，图中阴影区为高浓度相区，而空白区域为低浓度相区。



然而，对于弹性各向异性的固溶体，除非不同原子间的失配度 η 很小以至于式(18-109)中第二项可以忽略不计，失稳分相的成分调制一般具有择优方向。从式(18-109)中不难看出，那

图 18-46 各向同性介质失稳分相计算机模拟结果

些具有较小弹性模量 E 的方向的波矢将有较大的增幅因子 $R(\beta)$, 因而将优先得到发展和处于主导地位。对于立方晶系的材料, 最常见的情况是沿 $\langle 100 \rangle$ 方向杨氏模量数值最小, 故在该方向上成分的调制将优先发生。定量的计算结果表明, 此时失稳分相的结果类似于“晶格常数”为 λ 的 CsCl 结构, 其中 Cs 离子的位置相应于 B 原子浓度极大位置, Cl 离子的位置相应于 B 原子浓度极小的位置。 B 原子的空间分布可近似表达为:

$$c - c_0 = A \left[\sin\left(\frac{2\pi x}{\lambda}\right) + \sin\left(\frac{2\pi y}{\lambda}\right) + \sin\left(\frac{2\pi z}{\lambda}\right) \right] \quad (18-114)$$

式中, A 为与时间有关的振幅。

应该指出, 本节到此为止对失稳分相动力学的理论处理结果仅仅是方程 (18-107) 作线性化近似后所得到的, 因此这些结果仅适用于描述分相过程的早期。为了使方程的解能够适用分相过程的中后期, 必须寻求方程 (18-107) 本身的非线性解。其中最为有效的途径是方程的数值解。D. de Fontaine 等人曾对于在 100°C 经不同时间时效的 Al-22.5%Zn 合金中失稳分相产物浓度分布作了数值解, 其结果如图 18-47 所示。由图中可见, 实际浓度分布与正弦波有明显的差异, 分相后期平衡的两相界面仍然显得弥散, 似乎存在组织的粗化迹象。

熔体、玻璃和聚合物也可能发生失稳分解, 作为失稳分解热力学判据的表达式 $\partial^2 G / \partial c^2 < 0$ 原则上也适用于这些系统, 关于聚合混合物的失稳分解将在 21.7 节中做一些介绍。这里简单介绍玻璃熔体冷却过程中的失稳分解现象及其应用。

Wiltzius 等报道^[7], 硼硅酸盐玻璃熔在冷却至分相温度以下时便分解为富含 SiO_2 的相和富含 B_2O_3 的相。两相分别形成连续的网络。由于富含 B_2O_3 的相可溶于酸, 将冷却获得的两相玻璃在酸中浸泡, 其中 B_2O_3 的相可逐渐溶解, 只留下富含 SiO_2 的玻璃相。这种单相玻璃具有互通的多孔结构。这种方法便被用来制造有广泛用途的多孔玻璃。

图 18-48 为这种多孔玻璃的小角中子散射测量结果, 显示其特征尺寸在

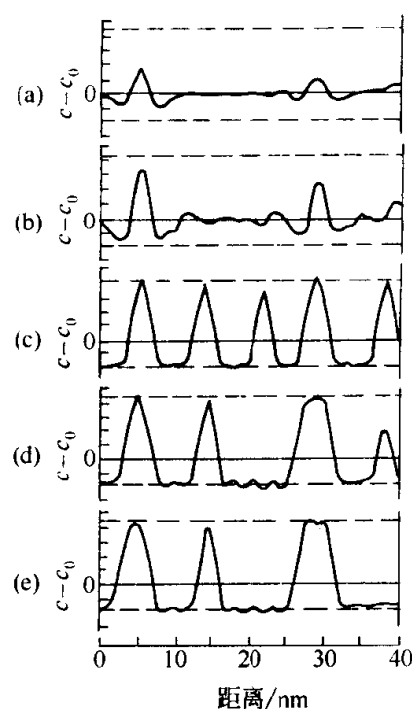


图 18-47 Al-22.5% Zn 合金 100°C 时效失稳分解产物浓度分布

(a) 10min; (b) 133min; (c) 60min;
(d) 400min; (e) 416min

几十纳米量级。Wiltzius 等还对该多孔玻璃的失稳分解过程利用 Cahn 的方程进行了理论计算, 结果如图 18-48 中的实线所示。显然, 小角中子散射的结果与理论计算的结果非常符合。这一实验也被认为是对 Cahn 的一次重要证明。

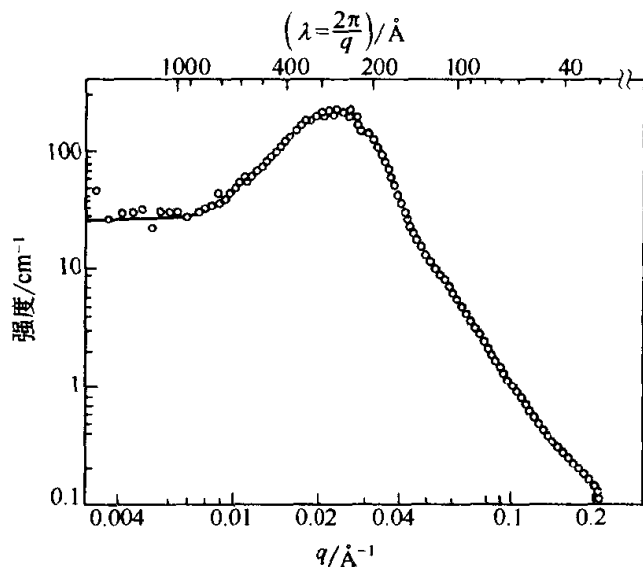


图 18-48 硼硅酸盐玻璃的小角中子散射结果及理论计算结果

注: $1\text{\AA} = 0.1\text{nm}$

18.4 结语

综上所述, 我们可以对涉及的相变进行分类。相变的类型可以从三个不同的角度 (即按热力学关系、按结构变化和按动力学关系) 来进行讨论。

相变的热力学规律是非常清楚的, 在按热力学关系讨论相变问题时, 系统的吉布斯自由能起了热力学势的作用。如果在相变点系统的热力学势的第 $(n-1)$ 导数保持连续, 而其 n 阶导数不连续, 则系统被定义为 n 级相变。一级相变的自由能的一阶导数在相变点是不连续的, 因而熵和体积的变化不连续, 说明它有相变潜热。而二级相变中, 熵和体积在相变点是连续的, 而自由能的二阶导数所确定的一些响应函数, 如比热容、压缩率和膨胀率则有不连续的变化。在自然界中观察到的相变多数是一级相变, 合金和金属中的相变也是如此。属二级相变的往往是一些比较特殊的相变: 如在临界点的气液相变, 铁磁相变, 部分超导相变, 超流相变, 部分合金的有序-无序相变, 部分铁电相变等。按照物理学界的习惯, 将二级和高级相变通称为连续相变或临界现象, 一级相变称为不连续相变。

从晶体学的观点, 阐明母相与新相在晶体结构上的差异, 即按结构变化对相变进行分类, 是对用热力学关系进行分类的一个重要补充。结构相变可

以分重构型、位移型和有序无序型三种基本类型。重构型相变中，大量化学键被破坏，在重新组合后，新相和母相之间在晶体学上没有明确的位向关系，而且原子的近邻的拓扑关系也产生显著的变化。这类相变经历了很高的势垒，相变潜热很大，过程缓慢。这类相变属于强一级相变。当然，液-固相变和气-固相变也必然是重构型的。另外，还有位移型相变，在相变前后原子的近邻的拓扑关系仍保持不变，相变过程不涉及化学键的破坏，新相与母相之间存在明确的晶体学位向关系，它经历的势垒很小，相变潜热也很小甚至完全消失。因此位移型相变可能是二级相变或弱一级相变。还有一种位移相变，它以晶格切变为主，也可能涉及晶胞内原子的相对位移，这就是人们通常说的马氏体相变，也是强一级相变。有序-无序相变在结构上往往涉及多组元固溶体中两种或多种原子在晶格点阵上排列的有序化。这可以是二级相变或弱一级相变。

从动力学机制上，相变可以分为匀相转变和非匀相转变。前者没有明确的相界，相变是在整体中均匀进行，其相变过程中的涨落程度很小而空间范围很大。二级相变总是按匀相转变方式进行的。也有一些一级相变是按匀相转变方式进行的，例如失稳分解，即在原始均匀固溶体中形成长波长周期性

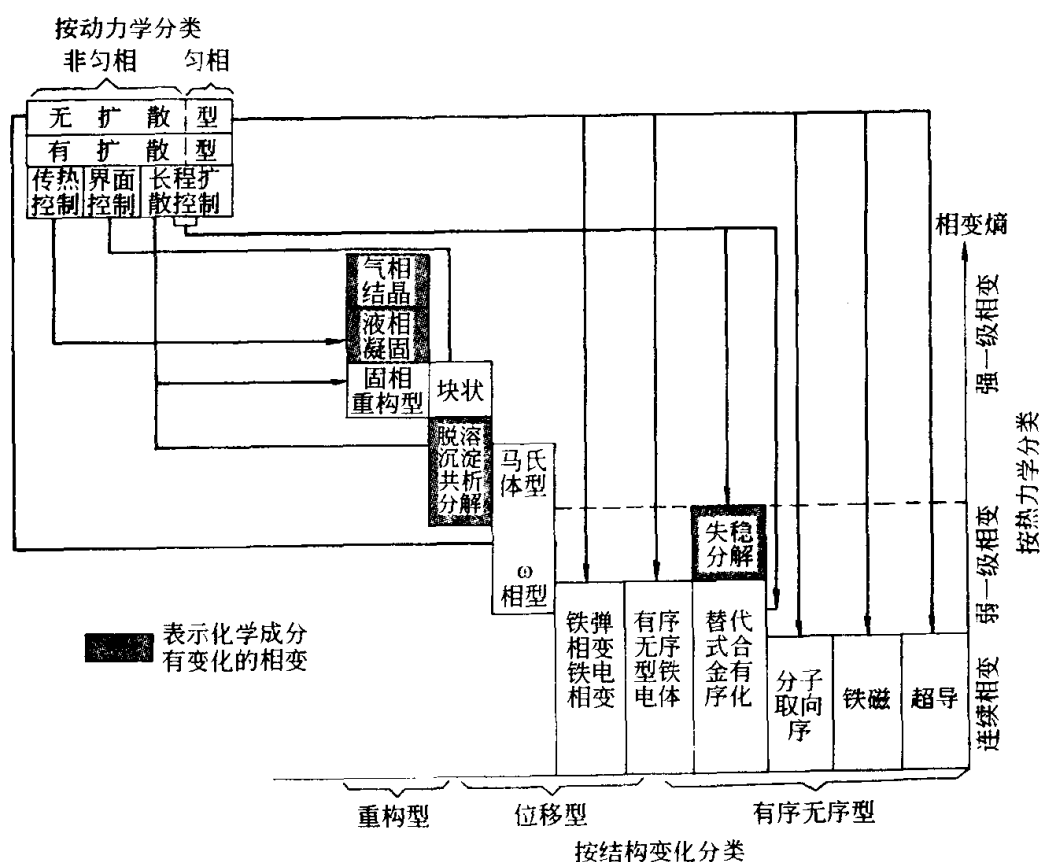


图 18-49 相变的分类

变化的分解过程。更常见的则是通过新相的成核生长来实现的，相变中母相与新相共存。其相变过程中，涨落的程度很大而空间范围很小。在非匀相转变过程中，如果成核生长不涉及原子扩散，就被称为无扩散相变，反之为有扩散相变。

我们将不同角度对于各种相变的分类总结在一张图上（见图 18-49），作为本章的总结。

参 考 文 献

- 1 冯端等. 金属物理学. 第二卷 相变. 北京: 科学出版社, 1990
- 2 哈森 .P. 物理冶金学. 北京: 科学出版社, 1984
- 3 陆佩文等编. 无机材料科学基础. 武汉: 武汉工业大学出版社, 1996
- 4 邓永瑞等. 固态相变. 北京: 冶金工业出版社, 1996
- 5 Khachaturyan A.G. Theory of Structural Transformations in Solids, New York: John Wiley & Sons, Inc, 1983
- 6 Rao C.N.R. et.al. Phase Transitions in Solids, New York: MacGraw-Hill, 1978
- 7 Wiltzius P. et.al. Phys. Rev. A. 1987, 36, 2991~2994

第 19 章 薄膜的外延生长

撰稿人 张 荣

外延技术在当代半导体器件制造工艺中占有核心地位。无论是作为计算机心脏的超大规模集成电路，还是现代通信技术中必不可少的光电子器件，其核心结构都涉及一层或多层单晶薄膜材料，必须依靠外延技术来制备。从物理上看，与晶体生长一样，外延也是一种相变过程，牵涉到相变的条件、速度等。但是外延与体单晶生长又有所不同，特别是因为外延层与衬底不同带来许多新的特点，产生了新的现象和规律，这些正是本章要讨论的问题。本章包括 3 个小节，第 1 节介绍外延的基本物理过程，主要讨论外延的机理及物理规律。第 2 节和第 3 节分别介绍应用上十分重要的同质和异质外延。

19.1 外延的基本物理过程

19.1.1 基本概念

外延 (epitaxy) 一词的使用已经有一段历史，从语源上来看，它是由希腊语的 epi (意为在……之上) 和 taxy (意为有序) 组合而成。一般将一定条件下，一种单晶体表面上另一种单晶体的定向生长称为外延，外延部分称为外延层，提供外延基础的晶体称为衬底。外延层与衬底间的差别可能是化学组成上的，也可能是结构上的。对于非晶体来说，一般只是短程有序，在比较大的尺度上，材料中原子的排列呈无序状态。对于多晶体来说，最大的特征是材料中不同部分的晶体取向不同，即使在有些情况下可以获得在某一晶向 (一般是生长方向) 上比较一致的排列 (这种情况称为择优生长，该晶向称为择优方向，这种结构有时称为织构)，但在垂直于该晶向的晶面内，各部分晶体的晶向一般呈无序排列。而外延则要求生长的晶体材料在各个晶向上呈比较严格的规则排列。同时，这一定义也将外延与块状晶体生长区分开来，如果“外延层”与衬底无论从什么意义上说完全是同种晶体，那么“外延层”的形成过程实际上与通常意义上的晶体生长过程没有本质的差别，虽然这时的生长条件与一般同种材料的块状晶体生长条件可能有较大的差别，这样的过程一般不称之为外延。外延生长区别于块状晶体生长的另一个特征是生长条件不同。最常用的外延方法是气相外延，即在衬底表面造成生长物原子的过饱和，驱使气相中的生长物原子并入固相，在衬底表面外延生

长出晶体薄膜。气相外延的温度通常远低于一般情况下同种材料块状晶体的生长温度，生长时的过饱和度与块状晶体生长相比也比较低，这导致外延生长的速率一般情况下远低于块状晶体的生长速率。这一特点决定了外延方法更多的是作为一种薄膜材料的制备手段得到应用。

一个典型的外延过程是掺硼的硅（100）面上不掺杂硅的外延生长，这里硅外延层的结构及取向与衬底的结构与取向完全相同，具体地说就是外延硅的（100）面平行于衬底的（100）面，（100）面内外延层的 $\langle 110 \rangle$ 晶向也与衬底的 $\langle 110 \rangle$ 晶向完全一致，记为 $(100) \langle 110 \rangle_{\text{Si}} \parallel (100) \langle 110 \rangle_{\text{Si:B}}$ 。这种外延层结构及取向与衬底完全一致的情况在外延生长中极为普遍，绝大多数半导体薄膜材料的外延生长都遵从这一规律。

理论与实验结果都证明，外延层的结构与取向主要取决于外延层与衬底间界面的自由能，稳定的外延层初始晶体结构及取向是由最低界面自由能决定的。对于外延层材料与衬底不同的外延体系来说，界面自由能选择的结果可能导致对前述外延规律的偏离，引起外延层取向或结构与衬底不同。这方面的一个例子是 MgO（100）面上 PbS 或 PbTe 的外延生长，两种情况下外延层取向都与衬底不同，而是 $(100) \langle 110 \rangle_{\text{PbS或PbTe}} \parallel (100) \langle 100 \rangle_{\text{MgO}}$ 。也就是说，外延层的生长面与衬底晶面一致，均为（100）面，但生长面内的取向不同，外延层的 $\langle 110 \rangle$ 晶向与衬底的 $\langle 100 \rangle$ 晶向平行，外延层与衬底的晶格相互间旋转了 45° 。类似的情形也发生在蓝宝石（0001）面上外延生长 GaN 的结构中，这时界面自由能选择的结果导致 GaN 外延层与蓝宝石衬底的晶格相互间旋转了 30° 。

界面自由能与界面所处的晶面取向有一定的依赖关系，这种依赖关系有时会导致外延层生长面取向对衬底生长面取向的偏离。这种情况可以由 GaAs（00 $\bar{1}$ ）面上 Ag 外延生长的例子来说明。在低于 200°C 的外延生长温度下，GaAs As 面 [即（00 $\bar{1}$ ）面] 上 Ag 的生长面是（011），而 Ag 的 $\langle 111 \rangle$ 晶向平行于 GaAs（00 $\bar{1}$ ）面上 $\langle \bar{1}10 \rangle$ 晶向的 As 悬挂键方向，取向关系是 $(011) \langle 111 \rangle_{\text{Ag}} \parallel (00\bar{1}) \langle \bar{1}10 \rangle_{\text{GaAs}}$ 。

当界面自由能最小化对外延层生长面结构提出极限要求时，这种要求甚至可能改变外延层的晶体结构。下面这个非常有趣的关于 Ag（001）面上外延生长铜的例子可以充分说明界面自由能的选择对外延层晶体结构的影响。Ag 和 Cu 的体材料均具有面心立方（fcc）结构，在 Ag（111）面上外延生长 Cu 的取向关系是 $(111) \langle 110 \rangle_{\text{Cu}} \parallel (111) \langle 110 \rangle_{\text{Ag}}$ 。但当很薄的 Cu 外延生长在 Ag（001）面上时，外延的 Cu 薄膜具有体心立方（bcc）结构，而晶体取向关系是 $(001) \langle \bar{1}\bar{1}0 \rangle_{\text{bcc Cu}} \parallel (001) \langle 010 \rangle_{\text{fcc Ag}}$ 。Cu 的 bcc 结构只在外延薄膜中发现过，自然界中没有存在 Cu bcc 结构的证据。

通过上面的讨论，我们不难看出外延层的初始稳定晶体取向主要取决于外延材料与衬底材料的晶体结构及跨越界面的化学键的特性，而不依赖于具体的生长过程。这提供了一种将外延界面的平衡结构问题同外延生长动力学问题分开处理的可能，前者通过研究外延界面自由能解决，后者则具体研究生长物质的运输、表面吸附、脱附、反应、迁移及并入晶格等过程。

从热力学的角度来看，通常外延材料与衬底材料间存在化学键能和几何结构两方面的差异，或者说两者的化学势差存在化学键能与几何结构两种因素的贡献。再来看前面提到过的掺硼硅衬底上不掺杂硅外延的例子。由于衬底掺的杂质很少，可以忽略掺杂对整个外延材料中化学键能的影响，因此外延材料与衬底材料的化学键能差异对整体化学势差几乎没有贡献。相反，由于外来杂质的引入，应力会引起掺杂衬底与未掺杂外延层晶格常数间有一定的差别，外延层必须纳入一定的应变来匹配衬底的晶格，这种情况下整体化学势差主要来自于每个原子应变能的贡献。从这个例子的讨论中可以看出，这类外延的基本特点是化学键能因素对外延材料与衬底材料间化学势差的贡献很小，化学势差主要来源于两种晶格间的失配。我们称这类外延为同质外延。不具备上述特点的外延则称为异质外延。

19.1.2 表面成核

对外延材料结构有最大影响的阶段是生长的最初阶段，这个阶段叫做成核。当衬底表面只吸附少量生长物原子时，这些原子是不稳定的，很容易挣脱衬底原子的吸引，离开衬底表面。所以，要想在衬底表面实现外延材料的生长，首先由欲生长材料的原子（或分子）形成原子团，然后这些原子团不断吸收新的原子加入而逐渐长大成晶核。它们再进一步相互结合形成连续的单晶薄层。其过程如图 19-1 所示。

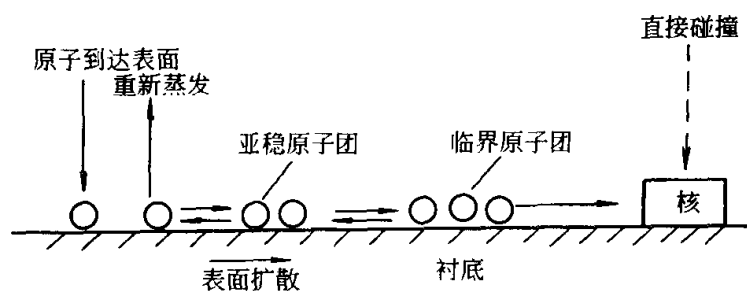


图 19-1 成核与生长过程示意图

气相外延是最常用的外延手段（参见第 15 章），这里重点讨论气相外延中的表面成核问题。在气相外延过程中，根据经典的热力学成核理论，气相中出现过饱和，即生长物的气相分压大于生长物的饱和蒸气压时，生长物原子可能发生气相成核，也称为均匀成核。由气相中形成晶核的临界半径 r^* 为

$$r^* = \frac{2\sigma V}{k_B T \ln(p/p_0)} \quad (19-1)$$

式中, σ 为表面自由能, V 为原子体积, p 为气相中反应物的分压, p_0 为衬底上生长物的平衡压力。在气相外延过程中, 气相中生长物的过饱和度都比较高, 加上生长是在衬底表面进行的, 是非均匀成核。在这种情况下, 包括有表面自由能项的经典热力学理论不再适用, 必须用原子统计理论来描述。

假设在单位衬底表面上有 N_0 个吸附原子的位置, 它们吸附了一些原子, 其中只吸附一个原子的有 N_1 个, 吸附两个原子组成的原子对的有 N_2 个, 吸附三个原子组成的原子团的有 N_3 个, 吸附 n 个原子组成的原子团的有 N_n 个。假定整个过程偏离热力学平衡不远, 可以近似按平衡态处理, 那么, 通过计算可得出分布为

$$\frac{N_n}{N_0} = \left(\frac{N_1}{N_0} \right)^n \exp(E_n/k_B T) \quad (19-2)$$

式中, E_n 是由 n 个原子形成原子团时的结合能。

临界晶核应该是生长几率等于 $1/2$ 的原子团, 一旦它接受一个原子之后, 它就取得了大于 $1/2$ 的生长几率, 成为稳定的晶核。由 n 个原子组成的临界晶核可通过与气相原子直接撞击或捕获吸附在衬底表面上的原子而实现稳定化。稳定原子团的生成率就是成核率。单位面积上的成核率可写为

$$I = N_n^* \omega_n \quad (19-3)$$

式中, ω_n 是单个原子加在大小为 n 的临界晶核上的速率, N_n^* 是衬底表面大小为 n 的临界晶核的面密度。对于由捕获表面扩散原子而成长起的晶核, 可推导出成核率的表达式

$$I = \sigma_n^* \nu \tau_0^{n+1} \frac{J^{n+1}}{N_0^n} \exp\left(\frac{E_n + (n+1)E_\tau - E_d}{k_B T}\right) \quad (19-4)$$

式中, σ_n^* 是临界晶核周围的吸附位置数 (也可以看成是临界晶核的捕获截面), J 是输入的反应物流, ν 是一个吸附原子的振动频率, E_d 是吸附原子表面扩散的活化能, τ_0 是一个与吸附原子在表面停留时间有关的常数, E_τ 是原子吸附能。

在不同的实验条件下, 原子吸附能和成核率是不同的。对于 E_τ 值较小的, 形成晶核要求较高的输入流。图 19-2 是气相外延生长中原子吸附能 E_τ 和表面扩散能 E_d 的示意图。

临界晶核在衬底上的构型与许多因素有关, 如果衬底为面心立方结构的

晶体, 当温度较低时, 单个原子就可以成为一个临界晶核, 它与任一原子都可以组成原子对, 以后再加入单原子形成三原子团, 如图 19-3(a)、(b)、(c) 所示。由于原子加入原子团的方向带有任意性, 因此, 原子团的构型将受原子加入的随机方式来决定, 原子集团间方向并不完全相同。当温度升高时, 双键相连才是稳定的, 如图 19-3(d)、(e)、(f)。

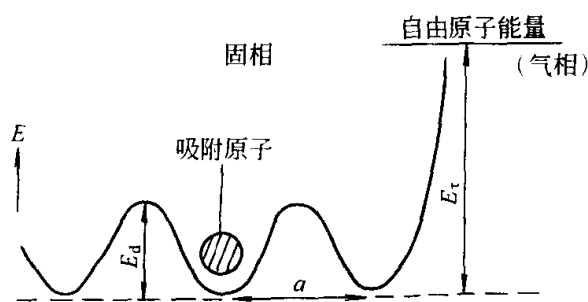


图 19-2 气相外延生长时生长物原子的吸附能 E_d 和表面扩散能 E_a 示意图

晶核形成和生长的方式还与生长层及衬底的性质有关。如果生长层的吸附原子与衬底原子间的粘着力大于吸附原子间的内聚力, 则可发生层状生长。即成核后, 其他原子迁移到台阶和扭折处, 通过台阶的侧向移动来进行层状生长。此时衬底取向对原子密排面的偏离越大, 在表面上提供的台阶数越多, 单个原子加入也较容易。对于衬底和外延材料不同的异质外延, 有时生长物的吸附原子与衬底间的粘着力小于吸附原子间的内聚力。这时, 吸附原子将先形成三维核岛, 而后核岛一边生长的同时一边又出现新核, 最后在核岛间吸附原子浓度下降到不能生成新核时, 它们相互连结形成膜。

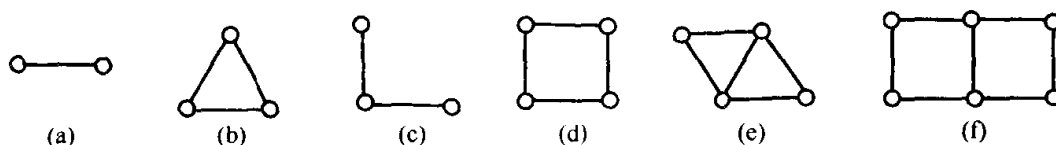


图 19-3 二维原子集团

生长温度及生长物的过饱和度对成核过程有很大的影响。若降低生长温度, 则原子在晶体表面上的迁移速率和距离也都变小, 原子难以顺利到达正常晶格位置排列。如生长物蒸气的过饱和度较大、成核较多, 就会生成多晶。如生长物的过饱和度进一步增大而且生长温度更低时, 吸附的原子迁移率更小。一般它停留在原来吸附位置附近, 而不能运动到晶体表面台阶等能量最低的位置去排列。而新的吸附原子团又不断生成。若新生的原子团生成速率大于它的迁移速率, 就生成无定型结构。

如果淀积物的过饱和度非常高, 气相的温度也高时, 在气相中分子的碰撞频率增大, 此时产生均匀成核。这种三维成核对材料外延来说是非常有害的, 应当尽力避免。

19.1.3 表面动力学

在反应物被输运到衬底上方后，通常会发生下列过程：①反应物扩散到衬底表面；②反应物吸附到衬底表面；③表面过程，包括表面化学反应、表面迁移及并入晶格等；④表面反应的附加产物从表面脱附；⑤附加产物扩散离开表面。这里的每个步骤都有特定的激活能，因此在不同的外延温度下对生长速率的影响也不同。

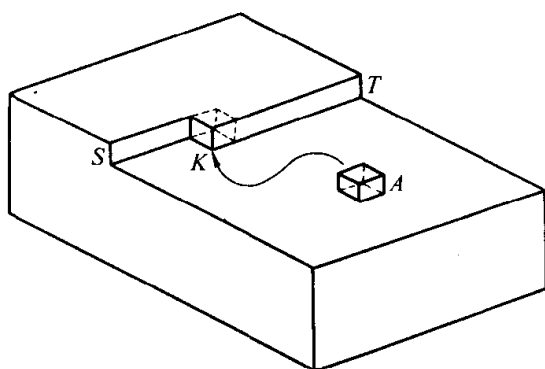


图 19-4 具有单原子层高度的表面台阶示意图

A 为到达表面的生长物原子，经过表面迁移后在 K 位置并入晶格

如果不考虑生长速率，仅从外延质量来看，步骤③是非常重要的。淀积到衬底表面上的原子通常去寻找合适的位置，使得系统的总能量降至最低，对于一个完美表面，这些位置是与由衬底形成的较深势阱有关的。对于实际表面来说，像表面台阶之类的表面缺陷是原子并入晶格的最佳位置，如图 19-4 所示。对于在表面上存在许多淀积原子的情况，它们除了在表面缺陷处键合外，还相互结合以进一步减少自由键的数目。外来的淀积

原子不断加入这个小的原子群并形成大的聚集体。很显然，当这些原子团继续生长时，它们自己就被看作是提供高结合能位置的表面缺陷，在淀积过程中进一步聚集原子。

原子团的生长主要有两个途径：熟化和聚合。在衬底表面上大量的原子团可以通过包含有质量输运的熟化过程，借助原子的表面扩散由小的原子团生长成大的原子团（如图 19-5 所示）。所谓熟化实际上是大原子团通过消耗小原子团而逐渐生长的过程。因为在有限的温度下，原子总是具有逃逸原子团的概率，并且通过表面扩散附着到另一个原子团上。这种从原子团逃逸出的原子通量的大小取决于原子团本身的尺寸。当然也存在着由大原子团向小原子团的相反的输运，但是总趋向是从小原子团流向大原子团，最终增加了大原子团的平均尺寸。

在熟化过程中主要的驱动力来自材料的浓度梯度，而该梯度是由不同尺寸的原子团上的气压差异产生的。通常是小原子团上的气压大于大原子团上的气

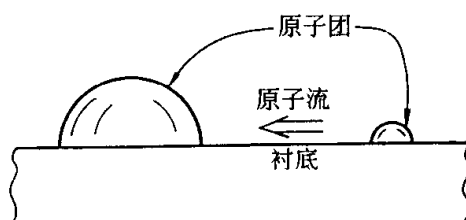


图 19-5 原子从表面的小原子团输运到大原子团示意图

压，当原子经过表面扩散并附着到大原子团上的过程中，气压差必定保持不变。因为来自小原子团的原子不断地供给，从而保持了其上的高压。就这样通过原子输运的过程，小原子团逐渐变小，而大原子团逐渐变大。

原子团的另一种生长机制是聚合过程，即把两个小的原子团合并成一个大的原子团。因为结合新原子团的表面积小于两个分开的小原子团表面积之和，因此结合后原子团总的表面能减小了，就像当两个水滴靠拢到一起时会形成一个大水滴一样（图 19-6）。

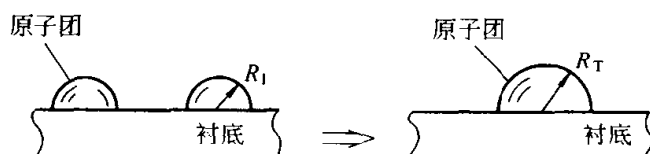


图 19-6 衬底上存在两个半径为 R_1 的小原子团（左图）和在相同衬底上存在一个体积为两个小原子团之和、半径为 R_T 的大原子团的比较

19.1.4 外延界面的结构

界面是分隔两个相邻体相的空间区域。外延界面是外延层与衬底两块单晶体的边界，在界面两侧，两块晶体保持特定的晶向对应关系。原理上讲，外延界面的结构与组成界面的两种晶体的化学键合性质、晶格结构、晶格参数及相应体材料的化学性质有很大的依赖关系。一般将外延界面分成 5 类。

(1) 晶格失配完全通过外延层应变来补偿 这已经在许多外延结构中被证实，特别在晶格失配比较小的外延结构中。例如，Cu (111) 衬底上 Ni 的外延和 GaAs (100) 衬底上 Ge 的外延。近年来的理论和实验研究都证明，这种界面的稳定性与晶格失配度及外延层厚度均有非常密切的关系，对特定的外延系统来说，外延层厚度存在一个临界值，只有当外延层厚度不超过这个临界值时，晶格失配才能够通过外延层应变来完全补偿，否则将在界面产生位错。临界厚度的值与外延层-衬底间的晶格失配及两种材料的力学性质密切相关，晶格失配越小，则临界厚度越大，材料越硬，则临界厚度越小。

(2) 晶格失配通过产生失配位错加以补偿 这存在于许多晶格失配的外延结构中，特别在晶格失配较大或外延层较厚的情况下。这种位错就是通常所称的失配位错，界面出现位错的晶体结构也称为晶格弛豫结构。在晶格失配比较大的情况下，失配位错密度可能达到很大的数值，例如，在蓝宝石 (0001) 面上外延 GaN 时，典型的失配位错密度在 $10^8 \sim 10^{10}/\text{cm}^2$ 量级。

(3) 有序的吸附原子或离子层 例如，W (100) 衬底上 K 外延结构的界面就包括一个带负电的 W 离子单层、带正电的 K 离子单层和 K 双层组成的偶极层。

(4) 界面处存在外延层/衬底原子互扩散区或固溶体区, 或亚稳相区
例如当 Pb 生长在 Au (100) 面上时会在界面处形成金属间化合物 Au_2Pb 。

(5) 界面处存在外延材料与衬底间的化合物层 这方面的典型例子是金属硅化物的形成, 当在 Si (111) 或 (100) 衬底上生长金属时, 在界面处会出现新的物相, 这就是金属硅化物层。

在大多数外延过程中, (1)、(2) 所代表的过程是最主要的, 在同一个外延系统中, 这两种情况都可能发生, 这时决定外延界面结构的机制是位错能与外延层应变所带来的弹性能的总和应当为最小。在材料体系确定后, 最主要的决定因素是外延层的厚度。通常厚度在临界值以下时, 外延层内蕴积的应变弹性能小于产生位错的能量, 晶格失配基本由外延层应变来补偿。当外延层厚度超过临界值时, 弹性能已经有了相当的积累, 在外延界面产生位错的能量已不再“高不可攀”, 而一旦在外延界面产生位错, 则弹性能会有大幅度的降低, 从总能量的角度来看, 以位错来补偿晶格失配比通过外延层应变来补偿在能量上更为有利。界面结构的这种转变在很多情形可能通过很多步来实现, 换句话说, 就是存在很多中间状态, 在这些状态, 既存在界面上的位错, 也存在外延层的应变, 两者对补偿晶格失配的贡献比例随外延层厚度增加而变化, 逐步从以应变补偿为主, 变化到以位错补偿为主, 相应的晶体结构从应变结构, 变成部分弛豫结构, 最后变为完全弛豫结构。

当材料所处的环境温度改变时, 晶格会随着环境温度的变化发生膨胀或收缩, 这种膨胀或收缩的程度由材料的热膨胀系数决定。对外延系统来说, 外延材料与衬底材料的热膨胀系数一般并不相同, 这种热膨胀系数的差异称为热失配。热失配有时很大, 这时如果发生环境温度的变化, 那么晶格失配度就会有较大的改变, 其结果是在外延体系中附加新的应力, 改变了弹性能在总能量中的比例, 引发新位错的产生。通常外延温度总是远高于室温, 因此当外延样品从生长温度降到室温时, 热失配有时会导致新位错的产生。

外延系统中的位错能与外延层应变弹性能均与两种材料的结构及物理性质有密切的依赖关系, 而材料的结构及物理性质又与温度有关。因此温度的变化能够引起外延系统中能量关系的变化, 从而改变界面结构。例如, 对外延样品在比外延温度更高的温度下进行退火处理, 会改变样品的微结构, 有时会导致外延层应变的弛豫。

上述讨论都是以衬底为半无限厚材料的假定为基础的。这一假定的直接后果是所有的晶格失配都必须通过外延层的应变或位错来补偿。大多数外延系统中, 衬底的典型厚度为 $100 \sim 500 \mu\text{m}$, 外延层的厚度一般在数纳米至十微米, 远小于衬底厚度, 上述假定基本上是合理的。但当衬底厚度很小时, 外延层的临界厚度大幅度增加, 因为这时衬底本身发生应变。如果衬底的厚

度远小于外延层的厚度，则衬底是主要的晶格失配承担方，此时衬底厚度也存在一个临界值，当衬底厚度小于此临界值时，外延层的厚度再大，外延界面上也不会产生位错。这正是自协变（compliant）衬底技术的物理基础。

19.2 同质外延

19.2.1 同质外延的生长模式

同质外延的生长模式主要由衬底温度决定。不同衬底温度条件下，淀积薄膜的不同结构组态表示在图 19-7 中。首先考虑最低温度区域（室温或更低）的生长情况。一般得到的是非晶态结构的薄膜，这是因为在如此低的温度下原子的表面迁移能力很弱。到达表面的原子随即被陷入到非晶的聚集体中，在这种情况下必须进行热处理，使这些原子重新排列才能形成外延薄膜。对于较高的生长温度，到达表面的生长物原子具有足够大的表面迁移率使其调整自己的位置在表面上外延生长，此时可能形成外延岛状区。在足够高的温度区域，外延生长主要是借助台阶的横向生长。图 19-7 中标出了温度转变区。从原理上讲，在高真空系统中，如果淀积速率足够低，即使在较低的温度下也能进行以台阶为媒体的外延生长。

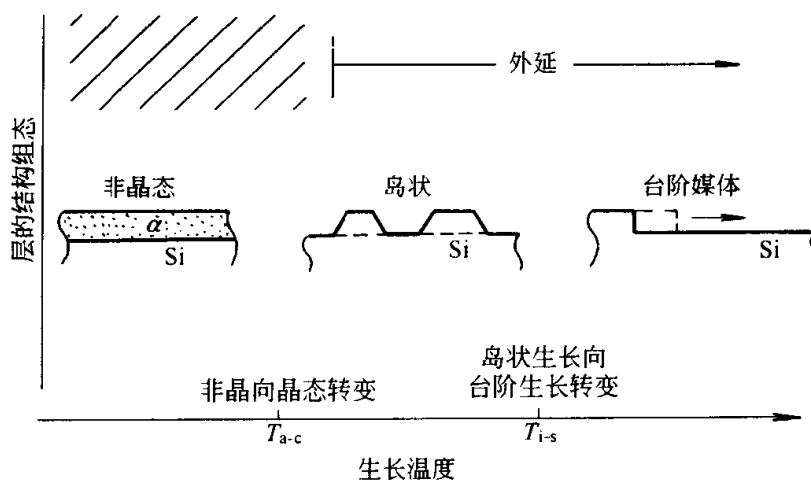


图 19-7 不同生长温度下各种生长层的结构组态示意图

图中标出了转变温度范围。实际的转变温度还与生长条件有关

基于表面扩散系数的考虑，可以比较定量地来分析外延生长的温度依赖关系。如果表面扩散系数 D_s 由下式表示：

$$D_s = \lambda^2 \nu_s \exp(-\Delta G_s / k_B T) \quad (19-5)$$

那么原子跳跃一次的平均时间 τ_s 可由下式给出：

$$\lambda^2 = D_s \tau_s$$

$$\tau_s = \frac{1}{\nu_s} \exp(\Delta G_s / k_B T) \quad (19-6)$$

如取 $\nu_s \approx 10^{13}/s$, $\Delta G_s = 1.0\text{eV}$, 则得室温下的 $\tau_s \approx 5 \times 10^3 s \approx 1.4\text{h}$, 如果淀积速率为 R (单原子层/s), 那么其他原子将在 $1/R$ 时间内跌落到初始淀积原子的邻近区域, 在其他原子到来前我们至少需要跳跃一次的时间, 让原有原子找到合适位置。小时量级的原子平均跳跃时间意味着在这样的条件下外延实际上是不可能发生的。实际的外延温度要比室温高出许多, 这时的原子平均跳跃时间要小得多。

19.2.2 杂质分凝与控制

外延过程中, 保持不变的生长条件下外延的材料中杂质在表面或界面附近呈不均匀分布的现象称为杂质的表面分凝。表面分凝的存在会影响外延界面的杂质分布, 甚至影响掺杂效率, 从而影响外延系统的物理性质。对于以外延界面为主要工作区的器件来说, 界面处的杂质分凝可能是实际器件的性能与设计性能偏离的主要原因。

外延系统中的杂质分凝问题与溶液生长过程中溶质的分凝有本质的不同。后者的物理原因是溶质越过位垒进入溶液所需的激活能和进入固溶体所需的激活能不同, 而前者则是表面自由能选择的结果。固体表面晶格中出现杂质原子时, 表面自由能会发生变化。如果杂质的引入增加了表面自由能, 则杂质倾向于离开界面(表面), 如果杂质的引入降低了表面自由能, 则杂质倾向于在界面(表面)处聚集。

杂质的分凝与外延温度有很大的关系, 原则上在较低的温度下生长有利于降低乃至消除杂质的分凝。但过低的生长温度会导致外延材料质量的劣化, 因为低温下生长物原子的表面迁移能力急剧下降, 原子难以到达势能最低的晶位。实际应用中控制杂质分凝更多地采用通过表面活性剂降低表面分凝的方法。表面活性剂又分为两类, 一类是恒表面活性剂, 如 Si 和 GaAs 外延中的 Sb, 其特点是表面活性剂的引入能大幅度降低表面自由能, 因此它总是倾向于在表面聚集, 其结果是生长表面永远被这类表面活性剂所覆盖, 随着外延过程的进行, 表面活性剂不断被推到表面, 而其下的杂质呈均匀分布。这类活性剂除上述要求外, 还要求在生长温度下有极低的平衡蒸气压, 即不易挥发。另一类是气相表面活性剂, 如 SiGe 衬底上 Si 外延中的 H 原子, 其特点是与外延材料中的某些成分优先成键, 在 Si/SiGe 的外延中优先与表面 Si 原子成键, 因此能有效抑制 SiGe 衬底中 Ge 原子向表面分凝。因为这种表面活性剂弥漫于整个生长气氛中, 所以整个衬底表面都被覆盖, 同时与外延方法兼容, 也不会引起新的玷污。

19.2.3 Si 和 GaAs 的同质外延

Si 是最重要的半导体材料, 目前有 90% 以上的半导体产品都是用 Si 制

作的。Si 的同质外延包括低掺杂衬底上高掺杂层的外延、高掺杂衬底上低掺杂层的外延和掺杂衬底上不同种类掺杂外延层的外延等几种情况。Si 的同质外延面临的核心问题有两个，一个是杂质的控制，包括分凝效应和生长环境中非故意掺杂效应的抑制，另一个是外延层结构的控制。杂质的问题除上述手段外，主要是采用低温、低压的生长技术，如超低压化学气相淀积、超高真空化学气相淀积和分子束外延等技术。解决结构问题主要靠所谓的完美 Si 外延技术。由于杂质原子与 Si 原子的共价半径不同，杂质的引入会改变外延层的晶格常数，致使整个 Si 片发生翘曲。当翘曲程度超过弹性范围，会产生位错来缓和内部的应力。针对这种情况，发展了完美 Si 外延技术，这一技术的要点是同时掺入另一种对内部应力有抵消作用的元素，通过共掺杂实现应力补偿，减小或避免晶格畸变。

GaAs 是化合物半导体材料的代表，在微波、光通信等许多领域均有极其重要的应用。GaAs 同质外延除面临与 Si 同质外延相同的问题外，还有一些其特有的问题。其中最突出的问题就是 As 的饱和蒸气压较大，生长及热处理时极易挥发，在材料中留下 As 空位。GaAs 的外延生长一般要求在富 As 的气氛下进行，如第 15 章中讨论 GaAs 的分子束外延时所介绍的那样。GaAs 的大规模外延生长常用的手段是低压金属有机化学气相淀积 (LP-MOCVD) 方法，其核心是采用液态的金属有机化合物源，这种源分解比较完全，而且流量易于控制，从而有利于对整个生长过程实现更加有效的控制。在 LP-MOCVD 方法中，一般要求远大于 1 的 V 族源流量/Ⅲ族源流量比，以满足生长气氛富 As 的要求。

19.3 异质外延

19.3.1 异质外延系统的结构与分类

异质外延生长是指外延材料与衬底材料互不相同条件下的外延生长，最常见的例子是生长在 GaAs 上的 AlGaAsⅢ~V 族化合物系统。 $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ 符号对应着外延薄膜/衬底，也就是说在该例中，AlGaAs 是外延薄膜， x 和 $(1-x)$ 分别代表 Al 和 Ga 的相对含量。对于 AlAs 和 GaAs 两种Ⅲ~V 族化合物，具有几乎相同的晶格常数，0.5661nm (AlAs) 和 0.5654nm (GaAs)，它们的晶格失配度仅为 0.1%。在实际应用中，晶格失配甚至可小于由这两种原始材料引起的 0.1% 的失配度，这是因为我们可利用像 $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}/\text{GaAs}$ 这样的合金组分，而 AlGaAs 合金的晶格常数一般在两个原始材料的晶格常数之间。所谓超晶格结构就是这些外延层在生长方向上的周期排列。作为一个例子就是在 GaAs 衬底上一个由 100nm 的 $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$ 层和 10nm 的 GaAs 层组成的重复结构，用符号来表示就是 $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}(100\text{nm})/$

GaAs (10nm)/……/Al_{0.5}As(100nm)/GaAs(10nm)/GaAs。因为 AlGaAs 组分在生长方向具有周期性，所以该结构被称为超晶格结构。对于 Ge 在 Si 上的外延生长是大晶格失配下外延生长的一个例子。这两种原始材料的晶格常数约有 4% 的差异 ($a_{\text{Ge}} = 0.5646\text{nm}$, $a_{\text{Si}} = 0.5431\text{nm}$)，这在外延生长中是很大的晶格失配了。在实际应用中，也常采用合金组分的外延生长，例如 $\text{Ge}_x\text{Si}_{1-x}/\text{Si}$ 。

异质外延生长能够形成自然界中没有的人工结构材料。在半导体应用中，材料的选择是根据其能隙与晶格常数的关系来决定的。图 19-8 中列出了最常见的几类半导体 (IV 族, III~V 族, II~VI 族) 材料的晶格常数和能隙。图中位于同一垂直线上的两种材料 (相同的晶体结构和晶格常数) 提供了进行晶格匹配异质外延生长的机会。例如，可以利用图 19-8，找到 $\text{In}_x\text{Ga}_{1-x}\text{As}$ 合金与重要的半导体材料 InP 之间有很好的晶格匹配。

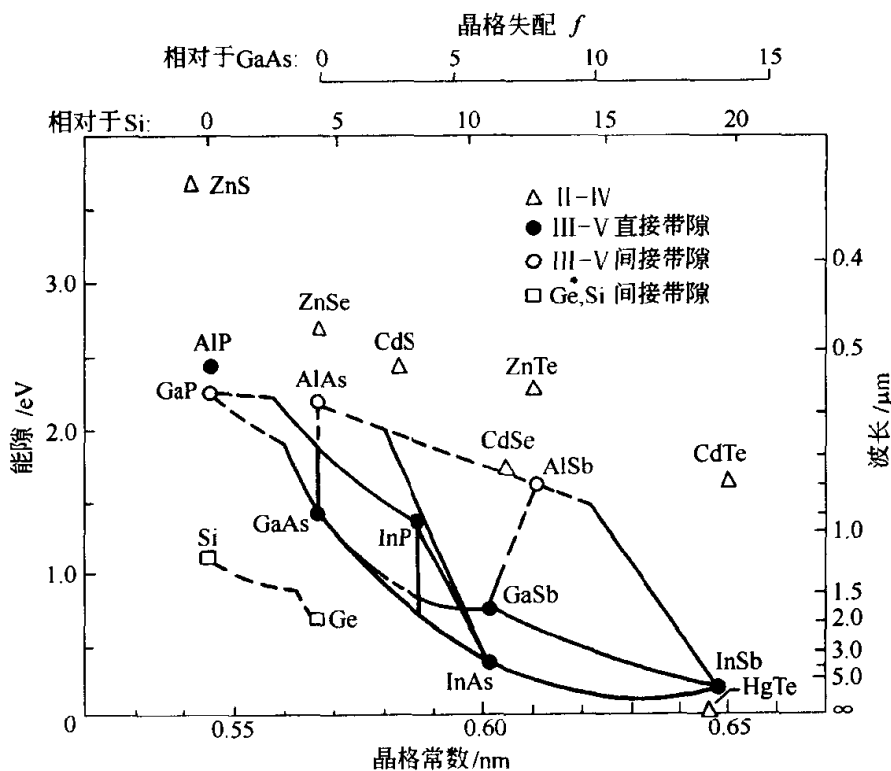


图 19-8 常见半导体材料的能隙和晶格常数的关系

图中空心符号代表间接带隙材料，实心符号代表直接带隙材料。

材料符号间的连线给出了三元合金材料的能隙和晶格常数。

相对于 Si 和 GaAs 材料的晶格失配分别表示在图的顶部

图 19-9 中示意地表示了晶格失配异质外延的剖面结构。衬底与薄膜都具有立方结构，晶格常数分别为 a_s 和 a_f 。该外延薄膜可以是应变的，也可

以是弛豫的。在赝晶生长模式下，外延层中平行于界面的晶格常数必须等于衬底材料的晶格常数，这样单胞将发生畸变，畸变的大小由泊松比决定。立方结构的单胞畸变为四方结构的晶胞。如果外延材料的晶格常数小于衬底材料的晶格常数，那么晶胞在平行于生长面方向上是拉长的，而在其高度上是缩短的。

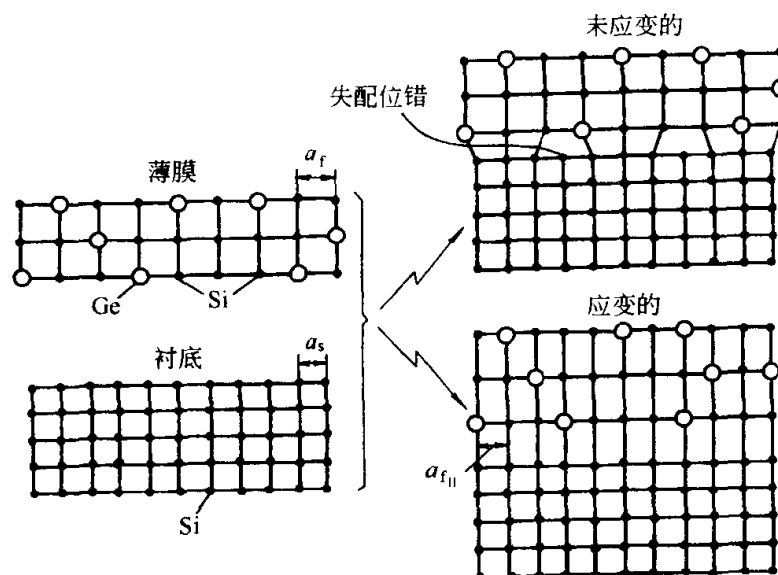


图 19-9 Si 单晶衬底上有 $\text{Ge}_x\text{Si}_{1-x}$ 薄膜形成的晶格失配异质结构。这种异质结构可以是应变的，也可以是带有失配位错而未应变的

我们定义生长面内的应变 $\epsilon_{||}$ ：

$$\epsilon_{||} = \frac{a_{f||} - a_{||}}{a_{||}} \quad (19-7)$$

式中的 $a_{f||}$ 表示外延材料平行于界面的或生长面内的晶格常数， $a_{||}$ 表示与体材料相同的未应变状态薄膜的晶格常数。对于赝晶体的外延材料， $a_{f||} = a_s$ ，其应变等于晶格失配。

对于弛豫的或未应变的系统，它们仍保持原有立方晶体的对称性。但在许多情况中，一个薄膜-衬底系统不一定是全部应变的或全部弛豫的，它们有可能包含位错的部分补偿。

从材料物理的观点来看，外延、应变或弛豫等过程指的是薄膜和衬底在晶体取向上的对准。但从电子材料的观点来看，应变的和弛豫的材料在电学性质上有极大的差异。材料中的位错会影响电荷输运性质和发光性质。从一定意义上说，研究晶格失配半导体系统的目标就是要生长只含有最小位错密度的赝晶体薄膜，因为这样的结构是器件中电子通过异质外延界面所必

需的。

19.3.2 应变层的稳定性

位错的产生提供了调节晶格失配材料的另一种机制（与应变机制相反的）。当外延层厚度超过临界值时，系统的弹性能超过了产生位错的能量，就会在外延层中产生位错。

对于赝晶体外延薄膜的临界厚度 h_c ，可由应变能和位错能与厚度的依赖关系以及总能量最小化原理推导得到。对处于任意应变 ϵ_{\parallel} 状态的系统，其应变能和位错能之和可由式 (19-8) 给出：

$$E_{\text{总}} = \epsilon_{\parallel}^2 Bh + \frac{\mu_f b}{2\pi(1-\nu)}(f - \epsilon_{\parallel}) \left[\ln\left(\frac{h}{b}\right) + 1 \right] \quad (19-8)$$

由 $dE_{\text{总}}/d\epsilon_{\parallel} = 0$ 可求得总能量处于极小值时的临界应变 ϵ_{\parallel}^* ；即由

$$\frac{dE_{\text{总}}}{d\epsilon_{\parallel}} = 2\epsilon_{\parallel}^* Bh - \frac{\mu_f b}{2\pi(1-\nu)} \left[\ln\left(\frac{h}{b}\right) + 1 \right] = 0 \quad (19-9)$$

那么

$$\epsilon_{\parallel}^* = \frac{\mu_f b}{4\pi(1-\nu)hB} \left[\ln\left(\frac{h}{b}\right) + 1 \right] \quad (19-10)$$

这里， B 是弹性常数的复杂函数，取决于晶体的对称性和生长方向。对于弹性各向同性薄膜， B 可简化为

$$B = 2\mu_f \frac{(1+\nu)}{(1-\nu)} \quad (19-11)$$

利用式 (19-11) 代入，得到

$$\epsilon_{\parallel}^* = \frac{b}{8\pi(1+\nu)h} \left[\ln\left(\frac{h}{b}\right) + 1 \right] \quad h > h_c \quad (19-12)$$

$$\text{和} \quad \epsilon_{\parallel}^* = f \quad h \leq h_c \quad (19-13)$$

式中的 h_c 由式 (19-14) 定义。

临界应变 ϵ_{\parallel}^* 可能的最大值是失配度 f 。从能量角度来分析，当薄膜在能量上允许产生第一个失配位错时的厚度，即是应变薄膜的临界厚度 h_c ，从这个意义上讲， h_c 就是当薄膜的 $\epsilon_{\parallel}^* = f$ 时的厚度。因此

$$h_c = \frac{b}{8\pi(1+\nu)f} \left[\ln\left(\frac{h_c}{b}\right) + 1 \right] \quad (19-14)$$

图 19-10 表示在 Si (100) 衬底上生长的 $\text{Ge}_x\text{Si}_{1-x}$ 应变层的临界厚度与 Ge 浓度百分比的关系。在应变层中改变 Ge 的百分比等价于改变失配度 f 。在室温下： $x=0$ ， $f=0$ ，当 $x=1$ 时， $f=0.042$ 。图 19-10 中的曲线是计算值，而数据点是实验值。图中实验值稍高于理论计算值是因为被测量的薄膜处在亚稳态，尚未完全弛豫到平衡状态。

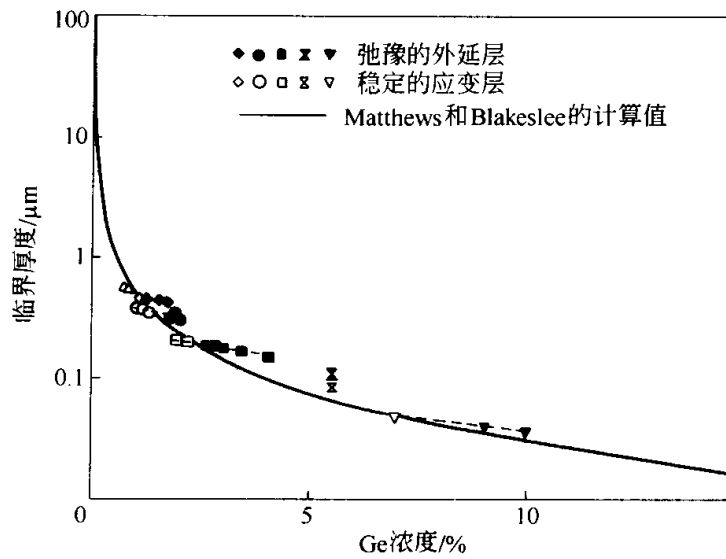


图 19-10 在 Si (100) 衬底上生长的 $\text{Ge}_x\text{Si}_{1-x}$ 应变层的临界厚度 h_c 与 Ge 浓度百分比的关系

19.3.3 异质外延的生长模式

异质外延存在三种主要的生长模式。首先是 Stranski-Krastanov (S-K) 生长模式，在该模式中淀积的起始层或以后的几层是保持平滑的，而原子团只在表面形成。在这种情况下，薄膜与衬底之间的界面能很强，使得生长的薄膜能延展充满整个衬底表面。原子团只能在外延薄膜的最顶上的几层原子间形成。

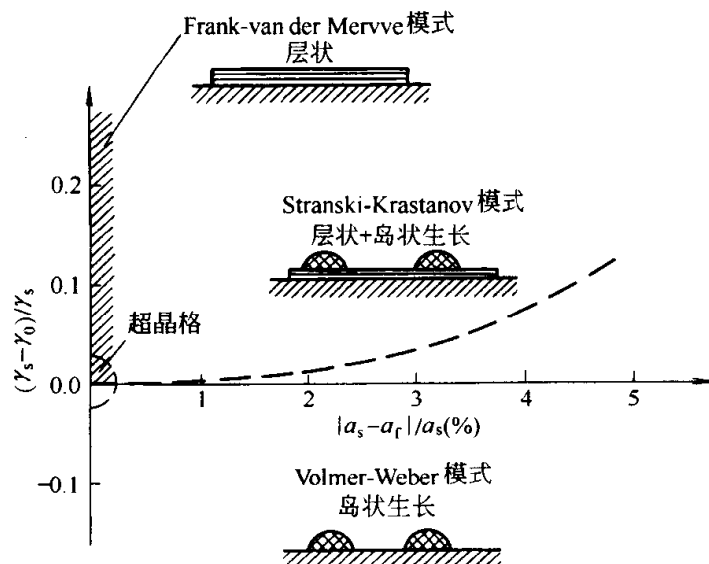


图 19-11 三种主要生长模式示意图及其能量比 W (衬底对薄膜) 和晶格失配的关系

W 定义为 $(\gamma_s - \gamma_0)/\gamma_s$, γ_s 和 γ_0 分别表示衬底和薄膜的表面自由能

图 19-11 是三种主要生长模式的示意图, 并标出了相对势能比 W (衬底对薄膜) 与晶格失配的关系。在合适的条件下, 逐层生长 (Frank-Van der Merwe 模式) 是能够实现的, 在这种结构中晶格失配为零。同质外延生长中, 表面能之比等于 1。如果外延薄膜的单位面积表面能大于衬底, 那么薄膜按 Volmer-Weber 生长模式形成原子团, 随着晶格失配的增加, 即使在 $W > 0$ (如图 19-11 中虚线表示) 的情况下也能形成原子团。

本章着重讨论了半导体的外延生长, 其原理也同样适用金属薄膜或氧化物薄膜的外延生长。当然在具体生长技术上是有些差异的, 这里就不细加讨论了。

参 考 文 献

- 1 闵乃本. 晶体生长的物理基础. 上海: 上海科学技术出版社, 1980
- 2 杜经宁, J.W. 迈耶, L.C. 费尔德曼. 黄信凡, 杜家方, 陈坤基译. 电子薄膜科学. 北京: 科学出版社, 1997
- 3 Ivan V Markov. Crystal Growth For Beginners. Singapore. World Scientific Publishing Co. Pte. Ltd., 1995
- 4 杨树人, 丁墨元. 外延生长技术. 北京: 国防工业出版社, 1992
- 5 吴自勤, 王兵. 薄膜生长. 北京: 科学出版社, 2001

第 20 章 生长界面的稳定性 ——枝晶生长和分形生长

撰稿人 王 牧

20.1 引言

当材料在远离平衡条件下生长时，其生长形态往往呈树枝状。对于这些类似树枝的形态，人们根据希腊语“树”（dendron）的拼写，起了一个名字 dendrite，中文翻译成枝晶。日常生活中最常见到的这类形态是雪花。人们常说没有两片雪花是一样的，这从一个侧面反应了枝晶生长的复杂性。对雪花的系统研究可以追溯到 100 多年前美国费尔蒙特州的一个叫杰理奇（Jericho）的偏僻小镇。在那里自学成材的农民本特利（Wilson A. Bentley, 1865~1931）对雪花产生了浓厚的兴趣，经过好几年的尝试和失败，拍摄了 5000 多幅雪花照片。他于 1931 年出版了“雪晶”（Snow Crystals）一书^[1]。书中收集了 2400 多幅精美的雪花照片。图 20-1 是部分雪花的图案。从一些图片中人们可以隐约看出雪花的界面生长和形态演变过程。在生长初期，生长界面一般为正六边形。随着界面的推进，多面体形态的稳定性被破坏。六面体的顶角处生长出树枝状的枝杈。面对这些漂亮的雪花，人们不禁要问：雪花究竟是怎样生成的？为什么会出现这些规则对称的枝杈？晶体的顶角处

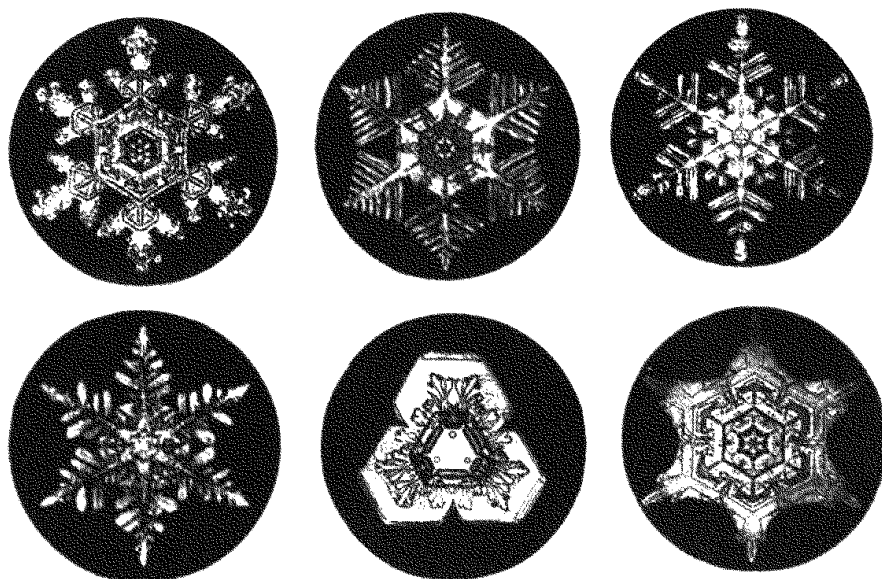


图 20-1 显微镜下拍摄的雪花照片

为什么容易失稳？

人们对枝晶生长的兴趣不仅仅在于自然界存在雪花这一类奇特的形态，还在于枝晶生长也是金属材料凝固和界面生长过程中经常遇到的现象，在金属

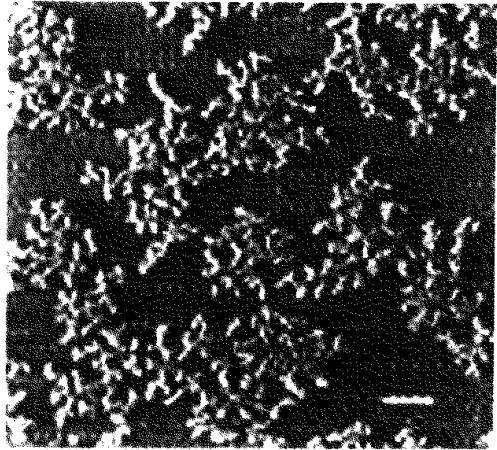


图 20-2 薄膜生长初期分形状的界面形态

属的铸锭、焊接等过程中出现。枝晶的生长往往伴有溶质的分凝和偏析，即使通过热处理也难将其痕迹完全消除，因而对金属材料的物理性能产生影响^[2]。人们为了在材料制备中避免出现枝晶就需要理解枝晶生长的过程和机制。

在薄膜材料的生长初期人们还经常观测到另一类形态，见图 20-2 所示。与枝晶相比，该形态没有十分明显的几何对称性，枝杈在生长中无规分裂，形成一种随机性很强的聚集簇。这一类聚集簇不具备严格周期结构特有的平移对称性，而是遵循标度不变性的规律。我们将这一类具有标度不变性的形态称为分形（fractal）。需要说明的是分形聚集体一般不是单晶，而由多晶聚集而成。分形聚集是远离平

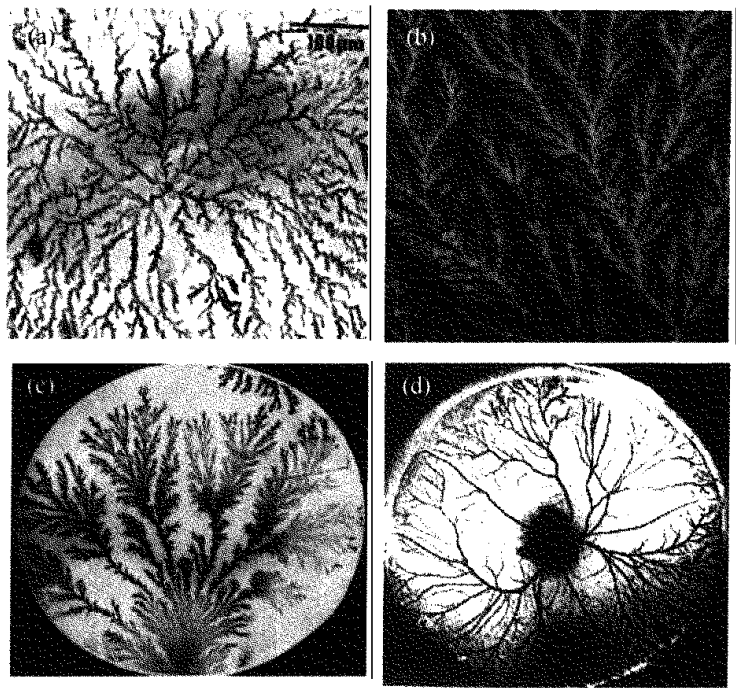


图 20-3 在不同实验系统中观测到的分形形态
(a) 晶体生长中的分形聚集体；(b) 电化学沉积中金属铜的聚集体；
(c) 岩石中矿物的分形沉积物；(d) 胚胎发育中血管分布

衡生长中较常出现的形态，存在于很多完全不同的生长体系中。图 20-3 显示在晶体生长、电化学沉积、矿物和生物体系中观测到的具有分形特征的形态。分形作为远离平衡条件下的生长形态之一，它们是怎样产生和演变的？它们的聚集过程与枝晶生长有什么异同？这也是本章要涉及的问题。

在本章的 20.2 节中我们介绍平界面的失稳及其判据；在 20.3 节介绍影响枝晶生长的几个重要因素和枝晶生长理论；随后在 20.4 节介绍分形的概念、表征方法和控制分形生长的几个物理学因素。

20.2 生长界面的稳定条件

20.2.1 温度梯度对界面稳定性的影响

假设固液界面初始时为一平面，我们来考虑温度分布对平界面稳定性的影响。在固液界面前的熔体中，其温度分布通常可以设想有 3 种形式，如图 20-4 所示。第一种是越离开界面温度越高，也就是说温度梯度是正的。由于我们假设固液界面的温度为凝固点 T_m ，故熔体的温度高于凝固点，因此称熔体为过热熔体 [图 20-4 (a)]。第二种温度分布如图 20-4 (b) 所示，温度梯度是负的，熔体为过冷熔体。第三种温度分布的特征是，熔体中的温度不是单调地改变，虽然远离固液界面的熔体仍为过热熔体，但是在固液界面附近却出现了一个狭小的过冷区，如图 20-4 (c)。

对过热熔体来说，如果在偶然的因素干扰下在平固液界面上出现某些凸

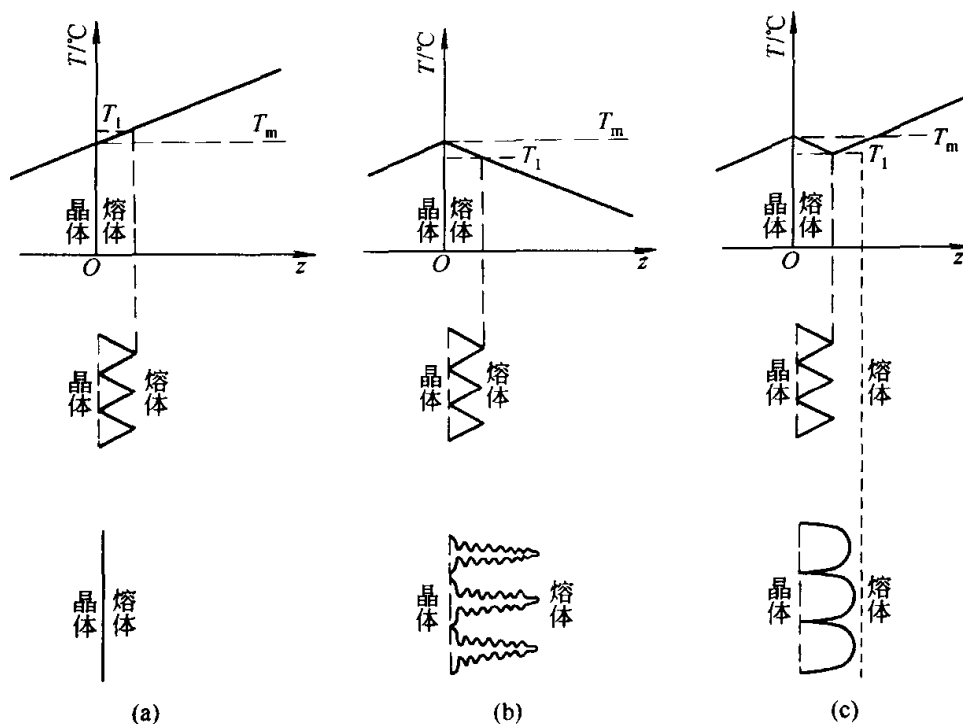


图 20-4 温度梯度对界面稳定性的影响

缘, 由于温度梯度是正的, 这些凸入到熔体内部的凸缘尖端处于更高的温度 T_1 , 见图 20-4 (a)。于是凸缘尖端的生长速率明显地下降, 或是被后面的固液界面所追及, 或是被熔化掉, 最终凸缘消失, 界面恢复平坦。这就是说, 当熔体中温度梯度为正值时, 平界面是稳定的。只有当界面在稳定的条件下, 材料生长的速率才是可以控制的 (调整工艺参量控制界面向熔体中推进的速率), 这是大多数生长过程所希望的条件。如果熔体中的温度梯度是负的, 则因干扰而产生的凸缘的尖端处于较低温度 T_1 , 即 $T_1 < T_m$, 见图 20-4 (b)。因而凸缘尖端的生长速度更高, 凸缘愈来愈大, 于是原先平坦界面上就出现了很多尺寸不断增长的凸缘, 我们说这种情况下平坦界面是不稳定的。其时凸缘尖端的生长速度越来越大, 生长变得不可控制了; 同时凸缘本身也会因干扰而出现分支, 从而出现枝晶的生长。在第三种温度分布的情况下, 由于固液界面前存在一狭窄的过冷区, 因而在平坦界面上因干扰而出现的凸缘能够保存, 但是由于远离固液界面处仍为过热熔体, 这些凸缘不能无限制地发展, 故保持一稳定的尺寸。此时界面的几何形状就像在平坦界面上长出了很多胞, 故称胞状界面 (cellular interface), 如图 20-4 (c) 所示。在这种情况下, 平界面是不稳定的, 而胞状界面却是稳定的^[3]。

20.2.2 浓度梯度对界面稳定性的影响

从前面的讨论中, 我们知道如果熔体中温度梯度是负的, 平坦的固液界面是不稳定的; 而如果熔体中温度梯度是正的, 平坦的固液界面就是稳定的界面。但是如果考虑了溶质浓度梯度的影响, 即使熔体中温度梯度是正的, 平坦界面也不一定是稳定的固液界面。在熔体中不同的温度梯度的情况下, 溶质浓度梯度对光滑界面稳定性的影响是需要具体分析。我们先来定性地讨论溶质浓度梯度对平坦界面稳定性的破坏作用。如果熔体中的温度分布如图 20-5 (a), 其温度梯度是正的。如果没有溶质影响, 平坦界面当然是稳定界面。但是如果熔体中存在平衡分凝系数 $k_0 < 1$ 的溶质, 在晶体生长的同时, 这些溶质不断地被排泄出来, 在生长界面前形成溶质边界层。溶质边界层中溶质分布如图 20-5 (b) 所示。分凝系数 $k_0 < 1$ 的溶质的第二个效应是, 对于很多种溶质, 溶液 (熔体) 的凝固点随溶质浓度增加而降低, 如图 20-5 (c) 所示。由于溶质边界层中溶质浓度随距界面的距离 z 的增加而减小, 故边界层中的凝固点也将随 z 的增加而上升。边界层中凝固点关于距离 z 的变化表示于图 20-5 (d) 中。在 $z = 0$ 处, 边界层中溶质浓度最高, 其值为 $c_L(0)$, 见图 20-5 (b); 相应的凝固点 $T(0)$ 也是最低, 见图 20-5 (c)、(d); 此后随着 z 的增加, 溶质浓度降低, 故凝固点随之升高, 至 $z = \delta$ 处, 浓度达到平均浓度 c_L , 故其凝固点也升高到相应的 T_m 。在边界层外, 浓度是均匀的, 故其凝固点也恒为 T_m 。在晶体生长过程中, 我们首先假定固液界面

温度为凝固点 T_m ，这样晶体才能继续生长。当溶质边界层建立后，界面外溶液的凝固点由原来的 T_m 降低到 $T(0)$ ，于是界面不再继续生长。我们可调整生长控制条件，将界面温度降至 $T(0)$ ，维持晶体生长。通常将界面温度降至 $T(0)$ 时，并不改变生长系统中熔体内的温度梯度，因而温度梯度仍为正值，且大小不变。如果熔体中没有溶质边界层，熔体中各处的凝固点都是恒定的。如果实际温度分布具有正温度梯度，则只有界面温度为凝固点，熔体中其他地方都高于凝固点，因而熔体处于过热状态。过热的程度随离界面的距离 z 而增加，故界面是稳定的。如果熔体中存有溶质，当溶质边界层建立后，在边界层内各点的凝固点不等，见图 20-5 (d)。虽然界面实际温度仍为凝固点，而且当离开界面进入熔体时，熔体的实际温度上升，但在图 20-5 (d) 之阴影线的区域内，熔体的实际温度低于凝固点，这意味着熔体处于过冷状态。一旦在平坦界面上因干扰产生的凸缘，其尖端处于过冷度较大的熔体中。其结果是尖端的生长速率比界面快，凸缘不能自动消失，于是平坦界面的稳定性被破坏了。固液界面初始时为过热熔体，因溶质分凝的出现而产生一过冷区，这种因组分变化而产生的过冷现象被拉特夫 (J. W. Ruttelv) 和查默斯 (B. Chalmers) 称为组分过冷 (constitutional supercooling)。

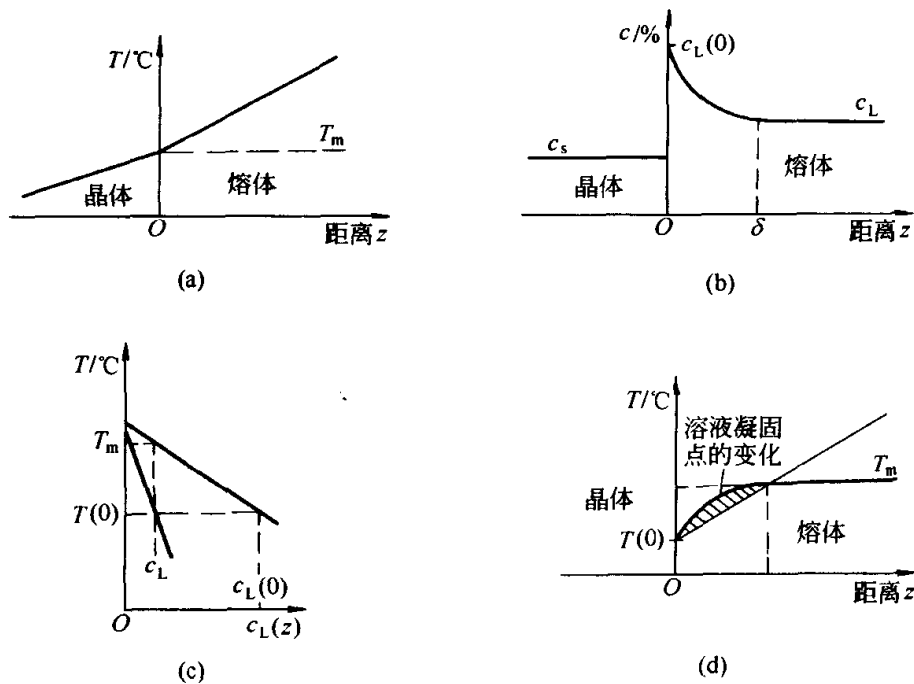


图 20-5 溶质分布对界面稳定性的影响

- (a) 固液界面邻近的温度分布; (b) 固液界面处的溶质分布 (溶质边界层);
 (c) 凝固点与浓度的关系; (d) 溶液凝固点分布以及组分过冷区的形成

组分过冷在晶体生长中是十分重要的现象，它最初是在 1937 年被斯米尔诺夫斯基 (Smialoweki) 发现的。在 20 世纪 50 年代查默斯等人进行了系统地研究，并提出了组分过冷的概念。组分过冷现象和枝晶生长都是推动界面稳定性理论发展的主要实验依据。发生组分过冷与如图 20-4 (c) 所示的温度非正常分布十分类似，在固液界面前沿都存在一狭窄的过冷区，而远离界面处的熔体都是过热的。因而在这两种情况下平坦界面都是不稳定的，都将转变为胞状界面。然而值得注意的是，上述两种情况下产生胞状界面的原因在本质上是完全不同的，前者由于溶质改变了溶液的凝固点以及溶质边界层的形成，后者由于实际温度的反常分布。

负温度梯度 [图 20-4 (b)] 与组分过冷同样能破坏平坦界面的稳定性，但两者也有明显的不同。负温度梯度的情况下，整个熔体处于过冷态，界面上的凸缘能自由地高速地向熔体中伸展，难以人为地控制生长。而组分过冷的情况下，由于组分过冷区有一定的厚度（约等于溶质边界层的厚度），因而凸缘只能被限制在组分过冷区内。

20.2.3 界面能对界面稳定性的影响

固液界面在偶然因素干扰下产生凸缘，因而增加了固液界面的面积，这就使固液界面的总界面能增加了。我们知道，界面能的增加就提高了系统的自由能。而在平衡条件下系统的自由能有缩小的趋势，于是固液界面面积就趋于缩小，这将促使平坦界面上的凸缘趋于消失。因而界面能是促进生长界面稳定的因素。理论分析表明，如果干扰较小（在干扰初期），凸缘的尺寸小于微米数量级，则界面能对界面稳定性的贡献较大。如果凸缘已经长大了，其尺寸超过微米数量级，则表面能的作用就不大了。

20.2.4 平界面的失稳条件

在晶体生长系统中，界面失稳问题总是和质量传输或热量传输问题联系在一起的。因此需要求解扩散方程和热传导方程。这两种方程的数学形式是相同的，它们的解可以相互套用。虽然温度场和浓度场在性质上有许多相似之处，但也有差异之处，这表现在界面的边值条件上。在固液界面上温度是连续的，而溶质浓度却是不连续的，其差异取决于分凝系数 k_0 。如果采用化学势来代替溶质浓度，就可以消除这种差异。令 μ 表示溶质原子相对于溶剂原子的化学势， $\bar{\mu}$ 表示 μ 和温度为 T^* 时两相平衡的化学势 $\mu_c(T^*)$ 的差异，即：

$$\bar{\mu} = \mu - \mu_c(T^*)$$

如和平衡偏离不大，可以写为

$$\bar{\mu} = \delta c \left(\frac{\partial \mu}{\partial c} \right)_{c=c_e} \quad (\text{液相内})$$

$$\tilde{\mu}' = \delta c' \left(\frac{\partial \mu}{\partial c} \right)_{c=c'_e} \quad (\text{固相内})$$

这里的 δc 表示局域浓度偏离。这样, 扩散方程就可以写为

$$D_c \nabla^2 \tilde{\mu} = \frac{\partial \tilde{\mu}}{\partial t}$$

这和热传导方程

$$D_T \nabla^2 T = \frac{\partial T}{\partial t}$$

完全相似, 而

$$D = M \frac{\partial \mu}{\partial c}$$

这里的 M 和迁移率成正比, 这样, 固液界面处质量守恒关系

$$D'_c (\nabla c' \cdot \mathbf{n}) - D_c (\nabla c \cdot \mathbf{n}) = \Delta c (\mathbf{v} \cdot \mathbf{n})$$

就可以改写为

$$\mathbf{n} \cdot (M' \nabla \tilde{\mu}' - M \nabla \tilde{\mu}) = \Delta c v_n$$

其中, Δc 为界面上液相和固相的浓度差。而 $\tilde{\mu}$ 在界面上是连续的, 其数值为

$$\tilde{\mu}^* = - \left(\frac{\gamma}{\Delta c} \right) \kappa$$

γ 为表面张力, κ 为界面曲率。

为了强调传热和传质的等效性, 我们引入无量纲的量 u

$$u = \begin{cases} \frac{T - T_m}{L/c_p} & (\text{传热模型}) \\ \frac{\tilde{\mu}}{\Delta c \left(\frac{\partial \mu}{\partial c} \right)} & (\text{传质模型}) \end{cases}$$

连续性条件可表示为 $v_n = -D \nabla \mu$, 而界面的吉布斯-汤姆逊关系可表示为 $u^* = -d_0 \kappa$, 其中 d_0 为表面张力效应的毛细长度。

有了上述定义后, 我们可以着手解传输方程。在以速度 v 运动的坐标系中, 扩散方程可写成

$$D \nabla^2 u + v \frac{\partial u}{\partial x} = 0 \quad (20-1)$$

定义扩散长度 $l = \frac{D}{v}$, 上式可改写为

$$\nabla^2 u + \frac{1}{l} \frac{\partial u}{\partial x} = 0 \quad (20-2)$$

在边界条件 $u|_{x=0} = 0$; $u|_{x=\infty} = -1$ 下, 方程的解可表示为

$$u = \begin{cases} \exp\left(-\frac{x}{l}\right) - 1, & x \geq 0 \\ 0 & x \leq 0 \end{cases} \quad (20-3)$$

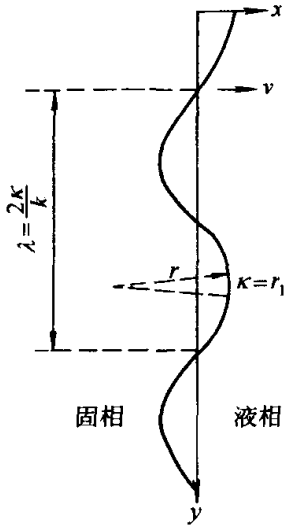


图 20-6 界面干扰波的几何关系

在得到扩散方程的稳态解后，我们在界面上引入微扰，随后考察这个解在物理上是否是稳定的。如果干扰会自发消失，这意味着界面在物理上是稳定的；如果干扰不仅不消失，反而随着时间的推移而增长，那么界面就是不稳定的。考虑如图 20-6 所示的界面，干扰前界面为 $x = 0$ 。由于界面上的任何干扰都可以用傅里叶级数或积分来展开，假设干扰后界面变为

$$x(y, t) = \delta(y, t) = \delta_k \exp(iky + a_k t) \quad (20-4)$$

干扰波的波矢 k 与 x 轴垂直。我们需要特别关注参量

$$a_k = \frac{\partial \delta}{\partial t} = \frac{\dot{\delta}}{\delta} \quad (20-5)$$

的符号。如果 a_k 是正值，干扰振幅与时俱增，界面随着时间的演变是不稳定的；如果 a_k 取负值，干扰振幅随时间而趋于零，因而界面是稳定的。为简单起见，我们假设对应于界面的干扰 $\delta(y, t)$ ，生长界面两侧的固液两相中的浓度场的响应分别为

$$\text{液相: } u \cong \exp\left(-\frac{x}{l}\right) - 1 + u_k \exp(iky - qx + a_k t) \quad (x > 0) \quad (20-6)$$

$$\text{固相: } u' \cong u'_k \exp(iky + q'x + a_k t) \quad (x < 0) \quad (20-7)$$

上式中 q 和 q' 项前的符号正好相反，这反映了干扰对于 u 的影响将随 x 点远离界面而衰减。将上式代入稳态扩散方程 (20-2)，可以求得 q 和 q' 所满足的方程，即

$$-\frac{q}{l} + q^2 - k^2 = 0 \quad (20-8)$$

$$\frac{q'}{l} + q'^2 - k^2 = 0 \quad (20-9)$$

q 和 q' 应为此方程正值的根。利用边界处 u 的值 $u|_{x=0}$ 和 $u|_{x=\infty}$ 可以确定剩下待定的系数 u_k 和 u'_k 。在 $x = \delta_k$ 处界面的曲率为 $\kappa \approx k^2 \delta_k$,

$$u|_{x=\delta_k} = u'|_{x=\delta_k} = -d_0 k^2 \delta_k \quad (20-10)$$

其中 d_0 为毛细长度。考虑到 u_k , u'_k 和 δ_k 都是小量，我们可以在级数展开中舍弃高次项，即

$$-\frac{\delta_k}{l} + u_k = u'_k = -d_0 k^2 \delta_k \quad (20-11)$$

界面在受到干扰后，其运动速度由原来的 v 变成 $v + \frac{\partial \delta_k}{\partial t}$ 。在 $x = \delta_k$ 处，根据连续性条件，有

$$v + a_k \delta_k = -D \frac{\partial u}{\partial x} \Big|_{x=\delta_k} \quad (20-12)$$

在利用式 (20-8) 和式 (20-9) 求 $\frac{\partial u}{\partial x}$ 的值时采用了线性化处理，得到

$$a_k \delta_k = -\left(\frac{v}{l}\right) \delta_k + Dq u_k \quad (20-13)$$

由式 (20-11) 消去 u_k 和 δ_k ，得到

$$a_k = v \left[q - \left(\frac{1}{l} \right) \right] - Dq d_0 k^2 \quad (20-14)$$

在界面速度较低的情况下， $k \gg \frac{v}{2D}$ 。此时 $q \approx q' = k$ ，上式可简化为

$$a_k = \frac{\delta}{\delta} \approx kv(1 - d_0 l k^2) \quad (20-15)$$

a_k 的符号直接影响到界面的稳定性。 a_k 中第一项为正，它与界面运动速度成正比，其结果是导致界面失稳；第二项为负，它与界面能有关，起稳定界面形态的作用。 $a_k = 0$ 的一个根为 $k = 0$ ，表明对于长波干扰界面是不稳定的。另外两个根由 $1 - d_0 l k^2 = 0$ 决定。所对应的临界干扰波长是

$$\lambda_c = \frac{2\pi}{k_c} = 2\pi \sqrt{\frac{d_0 D}{v}} \quad (20-16)$$

临界干扰波长确定了失稳图像的特征尺寸，尤其是在界面失稳的初期。应该指出的是，我们这里只考虑了扩散对平界面失稳的影响。在实际生长系统，特别是合金的凝固系统中必须要考虑热量和质量的混合传输。具体分析可参见文献 [4]。

20.3 枝晶生长

枝晶是远离平衡生长中常见的一种形态，它一般具有一针状主干及其对称或不对称的侧枝。自由生长的枝晶由于材料内禀各向异性的作用具有较为明显的对称性，比如雪花具有六次对称。对枝晶生长机制的研究可以追溯到 A. Papapetrou 在 20 世纪 30 年代的工作和 G.P. Ivantsov 在 40 年代在理论上所作的尝试。Papapetrou 对枝晶生长的观测表明抛物线形状的针状主干的生长速率是一个常数，它是由扩散场决定的。Ivantsov 在忽略界面张力的条件下得到了枝晶针状主干的解析解。对应于一定的生长驱动力，主干生长速度与尖端曲率半径的乘积为一常数。这个理论指出对应给定的生长驱动力存在一簇解，其数目是无穷多的。然而在特定的实验条件下，人们只

观测到一个特定的枝晶形状和生长速度。1980 年前后人们通过一些简化了的枝晶生长模型,发现各向异性的表面张力可以帮助确定枝晶的生长速度。此后人们开展了大量的解析和数值计算,对枝晶生长有了较为深刻、定量的理解。本节我们介绍控制枝晶生长的几个主要因素以及对枝晶生长机制的最新认识。

20.3.1 控制枝晶生长的几个物理因素

(1) 表面张力 我们通过一过饱和的二元合金体系中的等温结晶过程来理解表面张力对枝晶生长的作用。系统中组元 A 的含量远大于组元 B (溶质) 的含量。在结晶过程中组元 B 由于分凝效应被部分地排出晶体外,因而 B 在晶体中的含量远小于在熔体中的含量,如图 20-7 (a) 所示。我们定

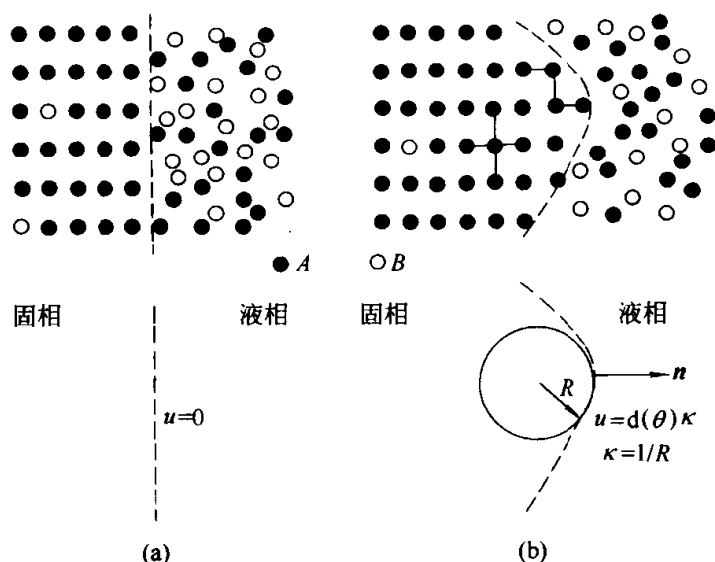


图 20-7 固液界面的微观图像及其与浓度场的关系

义溶质 B 在位置 x , 时刻 t 的浓度为 $c(x, t)$; 定义 $c_l^0(c_s^0)$ 为平固液界面液相 (固相) 一侧的平衡浓度。如前所述 $c_s^0 \ll c_l^0$ 。为表示简洁, 我们在本节中定义液相中无量纲浓度场

$$u(x, t) = \frac{c_l^0 - c(x, t)}{c_l^0 - c_s^0} \quad (20-17)$$

$u(x, t)$ 在 A 的过饱和液相中恒为正, 而当平衡时为 0。假设 A 的晶体向液相一侧凸起, 固液界面变成一个弯曲界面, 如图 20-7 (b) 所示。假设 A 的晶体具有正方点阵。考虑简单的二维情况。此时晶体内部每个原子周围都有 4 个最近邻的键合, 而在弯曲处的键合数将小于 4。换句话说, 在凸起的界面处, 晶体 A 的原子受到较少的束缚, 将较为容易地进入到液相中, 与平界面相比, 凸起界面前 A 原子的浓度较高 ($u > 0$)。很容易看出, u 应该

是局域表面曲率 κ 的增函数（对于晶体向液相凸起的界面我们定义 κ 为正）。对于界面弯曲不是十分剧烈的界面（ $\kappa a \ll 1$ ，其中 a 是晶格常数），假设线性关系

$$u = d(\theta) \kappa \quad (20-18)$$

成立。方程（20-18）实际上就是著名的吉布斯-汤姆孙关系，其中 $d(\theta)$ 是微观毛细长度，它与分子间相互作用力程在同一量级。微观毛细长度 $d(\theta)$ 是晶体结晶学取向 θ 的函数。对于正方晶格，考虑到对称性，我们假设 $d(\theta)$ 具有下列形式

$$d(\theta) = d_0(1 - 15\epsilon \cos 4\theta) \quad (20-19)$$

另一方面，由热力学我们可以推出表面张力与 $d(\theta)$ 的关系。对于稀薄固溶体

$$d(\theta) \approx \frac{\gamma + \gamma''}{k_B T c_1^0} \quad (20-20)$$

其中 $\gamma = \gamma_0(1 + \epsilon \cos 4\theta)$ 是晶体与熔体的界面能； γ'' 是 γ 相对于 θ 的二次导数，而 $\gamma + \gamma''$ 一般称为界面劲度或界面刚度。 k_B 是玻尔兹曼常数， T 为等温生长的温度。由此我们可以导出

$$d_0 \approx \frac{\gamma_0}{k_B T c_1^0} \quad (20-21)$$

大多数情况下，枝晶的主干沿着最大表面张力的方向生长。在二维情况下，其等效于最小毛细长度的方向。对于典型的有机晶体，各向异性系数 ϵ 的值在0.005~0.05之间。

(2) 扩散 当生长界面前存在浓度梯度 $\frac{\partial c}{\partial x}$ 时，相应的质量流 j_x 可写成

$$j_x = -D \frac{\partial c}{\partial x} \quad (20-22)$$

这意味着当平界面上引入一凸起后，单位时间内由扩散效应到达该凸起的质量会大于其他地方，从而导致其凸起的进一步增长。这体现了扩散场对生长界面失稳的贡献。由式（20-22）出发，考虑一单位空间的质量守恒，我们很容易得到

$$\frac{\partial c}{\partial t} = -\frac{\partial j_x}{\partial x} = D \frac{\partial^2 c}{\partial x^2} \quad (20-23)$$

对于二维体系，我们可以写成

$$\frac{\partial c}{\partial t} = D \nabla^2 c \quad (20-24)$$

或

$$\frac{\partial u}{\partial t} = D \nabla^2 u \quad (20-25)$$

其中

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

假设界面的法向生长速率为 v_n ，由质量守恒我们知道，单位时间内晶相生长的原子（分子）数必须等于由界面法向的质量流所输运的原子数，即

$$v_n = D \frac{\partial u}{\partial n} \quad (20-26)$$

式 (20-25) 和式 (20-26) 加上吉布斯-汤姆逊关系 [式 (20-18)] 可以描述晶体的界面演变了。需要指出的是吉布斯-汤姆逊关系只在近平衡条件下，或者说是生长驱动力较小的系统中成立。当枝晶的生长速度很大时，我们必须要考虑非平衡效应。

(3) 噪声 当空间两点的浓度相等时，我们说两点间没有净的粒子流。但是当有限数量的粒子在等浓度的两点间以相同的几率向左或向右无规行走时，由于无规行走的随机性，两点间随时有净的流量，其符号和幅度是随机涨落的。在输运过程中随机性可通过加入涨落项 q_x 引入，即

$$j_x = -D \frac{\partial c}{\partial x} + q_x \quad (20-27)$$

在 y 方向可以引入同样的方程。由于分子或原子在不同方向的无规行走是独立的，因此 q_x 和 q_y 可以看成是没有时空关联的随机变量，它们都满足高斯分布。噪声对枝晶侧枝的产生有重要贡献。

20.3.2 枝晶主干的生长——针状晶体的生长

为简单起见，人们将枝晶的主干生长与侧枝的生长独立开来考虑。这种理论上的处理方式的前提是侧枝的生长与主干生长是相互没有干扰和耦合的。对于侧枝的主干尖端与侧枝间距离远大于 D/v 的情况，这样的考虑是完全合理的。

枝晶主干生长速率与生长驱动力之间的关系最早是由伊万佐夫 (Ivantsov) 得到的^[5]。他的模型中忽略了界面张力并假设枝晶主干为旋转抛物面。在给定的生长驱动力（过饱和度） Δ 下三维生长界面满足

$$\Delta = p \exp(p) \int_p^\infty dz \frac{\exp(-z)}{z} \quad (20-28)$$

其中 $p = \rho v / 2D$ ， ρ 为枝晶主干尖端的曲率半径。显然，对应于一给定生长驱动力， ρ 与生长速度 v 的乘积 ρv 是一常数。换言之伊万佐夫的解尚不能确定 ρ 和 v 。要独立地给出 ρ 和 v ，还必须引进另外的判据。下面我们分别从宏观和微观角度考虑枝晶主干的生长。

(1) 由质量输运给出的枝晶主干的生长条件 我们首先从较为宏观的尺

度研究枝晶的形状与生长速度的关系。假设枝晶的形状可以由

$$x = \frac{y^2}{2\rho} \quad (20-29)$$

表示, 并且针状晶体的长度为 L 。根据式 (20-29) 枝晶的宽度 w 约为 $\sqrt{\rho L}$, 其宽度与长度的比 w/L 约为 $\sqrt{\rho/L}$ 。当 $L \gg \rho$ 时, $\frac{w}{L} \rightarrow 0$ 。这就是说, 在比主干尖端曲率半径 ρ 大很多的尺度上研究扩散流对主干生长的贡献时, 针状主干实际上可以看成是一条细线。下面我们通过扩散方程式 (20-25) 来求解到达主干尖端的质量流。这里边界条件设为 $u = 0$, 因为在这样的尺度上界面张力的效应是可以忽略的。因此问题就归结到在针状晶体对应的细线的顶端附近求解 Laplace 方程。 $\nabla^2 u = 0$ 的解是已知的, 并且满足

$$|\nabla u(\mathbf{r})| \sim \frac{1}{\sqrt{|\mathbf{r}|}} \quad (20-30)$$

其中 $|\mathbf{r}|$ 是空间一点到针状晶体尖端的距离。由此可知, 垂直于细线方向的浓度梯度应具有下列形式

$$\left. \frac{\partial u}{\partial y} \right|_{y=0} = \frac{F}{\sqrt{d_0 x}} \quad (\text{当 } x \rightarrow 0 \text{ 时}) \quad (20-31)$$

其中 d_0 的引入是为了保证 F 是一个无量纲的表征供应针状晶体生长质量流的量。这一点很容易通过对式 (20-31) 的积分看出,

$$\int_0^a dx \left. \frac{\partial u}{\partial y} \right|_{y=0} = 2F \sqrt{\frac{a}{d_0}} \quad (20-32)$$

其中 a 为在针状晶体尖端以后的与尖端曲率相当的任一距离。另外, F 主要

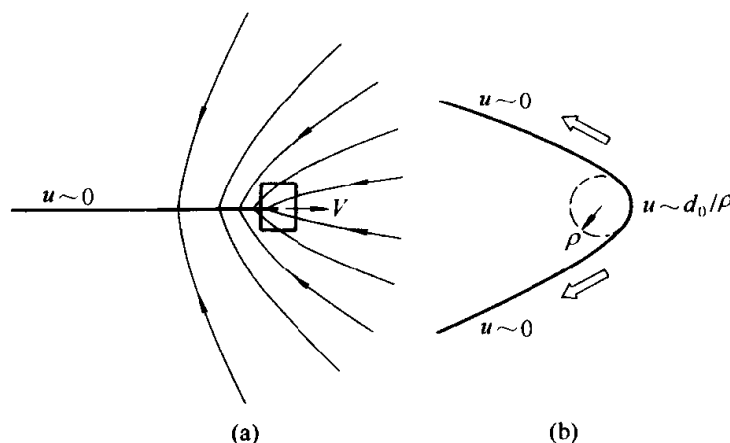


图 20-8 从两个不同的尺度考虑枝晶主干的生长

(a) 从宏观输运角度考虑枝晶尖端的生长, 进而确定 ρV^2 ;

(b) 从微观表面张力的角度考虑枝晶尖端的生长, 进而确定 $\rho^2 V$

依赖于远离枝晶处 u 的性质。由质量守恒我们知道

$$y_i(x) = D \int_0^x dt \frac{\partial u}{\partial y} \Big|_{y=0} = D \int_0^x \frac{dx'}{V} \frac{\partial u}{\partial y} \Big|_{y=0} = 2F \left(\frac{D}{V} \right) \sqrt{\frac{x}{d_0}} \quad (20-33)$$

这里 y_i 是枝晶主干的半宽度， V 为枝晶主干的生长速度。由针状晶体的抛物线形状我们得到

$$\rho = 2 \left(\frac{FD}{V} \right)^2 \frac{1}{d_0}$$

即

$$\rho V^2 = \frac{2(DF)^2}{d_0} \quad (20-34)$$

(2) 由表面张力决定的枝晶尖端生长条件 本节从微观角度研究枝晶的生长。在枝晶主干尖端曲率半径 ρ 的尺度上，扩散输运和表面张力效应的竞争对枝晶的生长起重要作用，如图 20-8 (b) 所示。一方面扩散流将质量输运到针状晶体的尖端，以维持枝晶的主干以速度 V 生长；另一方面由弯曲界面造成的生长界面上的浓度差又将通过面扩散造成界面上的质量转移。这两者的竞争直接影响枝晶的生长速度。在稳态条件下，由体扩散输运到生长界面前的质量

$$j_d \sim D \frac{\partial u}{\partial n} \sim V \quad (20-35)$$

而表面张力梯度造成的质量流倾向于将尖端界面上的质量转移到针状晶体的两侧，从而限制了针状晶体的不断变尖。由表面张力引起的质量流可表示为

$$j_s \sim D \frac{u_{\text{tip}} - u_{\text{side}}}{\rho} \quad (20-36)$$

其 u_{tip} 和 u_{side} 分别表示 u 在针状晶体尖端和侧面的值，两点间的距离为 ρ 的量级。考虑到针状晶体侧面曲率很小 $u_{\text{side}} \approx 0$ ，而 $u_{\text{tip}} \approx \frac{d_0}{\rho}$ ，由此我们得到 $j_s = \frac{Dd_0}{\rho^2}$ 。这两个流的强度之比用无量纲的参数 σ 表示，

$$\sigma = \frac{j_s}{\frac{1}{2} j_d} = \frac{2Dd_0}{V\rho^2} \quad (20-37)$$

由此我们得到了 V 和 ρ 的第二个关系式

$$\rho^2 V = \frac{2Dd_0}{\sigma} \quad (20-38)$$

根据 1980 年以后人们提出的一系列设想，以及其后解析和数值的研究，我们现在认识到 σ 取决于描述枝晶以一固定形态和恒定速度 V 运动的自洽解。为了决定这样的解， σ 必须取满足微分方程可解性条件的形式。这在物理上实际上对应于枝晶头部的形状必须是光滑的。对于完全各向同性的表面张力

$[\epsilon = 0, \epsilon$ 的定义参见式 (20-19)], 方程组是无解的。在这种情况下, 动力学模拟和实验都表明此时针状晶体的尖端失稳, 出现分叉形态。对于非零而有限的各向异性常数 ϵ , 只有当 σ 取一系列的分立值时方程才有解, 而这些解中只有对应于最大 σ (亦即最大生长速率) 的解对于外界无穷小的扰动才是线性稳定的, 因而在真实实验系统中可以被观测到, 而其他的解均会产生尖端失稳而出现分叉现象。对应于稳定针状晶体尖端的 $\sigma(\epsilon)$ 是 ϵ 的单调增函数。当 ϵ 较小时, 我们有

$$\sigma(\epsilon) \sim \epsilon^{\frac{7}{4}} \quad (20-39)$$

另一方面, 由式 (20-34) 和式 (20-38) 我们可以独立地决定针状晶体的生长速度和尖端曲率半径

$$v = \frac{D}{d_0} (2\sigma)^{1/3} F^{\frac{4}{3}} \quad (20-40)$$

$$\rho = d_0 \left(\frac{2}{\sigma^2} \right)^{1/3} F^{-2/3} \quad (20-41)$$

由此可见只要计算出了 F 的值, 枝晶生长的解就给定了。一般情况下, F 是时间 t 的函数。对于 F 的计算需要分别考虑枝晶的生长初期的过渡状态和到达稳定时的两种情况。假设处于生长初期枝晶的尺寸为 $L(t)$, 生长驱动力 $\Delta \ll 1$, 此时

$$F(t) = \frac{\Delta}{|\ln \Delta|} \sqrt{\frac{d_0}{L(t)}}$$

当针状晶体的生长已赶上了向外扩展的圆形扩散场时, 初始的过渡态结束, 枝晶的生长进入稳态。对于已达到稳定生长的枝晶, 我们可以推出

$$F(t) = \Delta \sqrt{\frac{d_0 v}{\pi D}}$$

将稳态条件下 F 的表达式代入式 (20-40) 和式 (20-41), 我们得到当 $\Delta \ll 1$ 时 ρ 和 v 应满足的关系

$$\begin{aligned} \frac{v d_0}{D} &\approx \frac{2\sigma \Delta^4}{\pi^2} \\ \frac{\rho}{d_0} &\approx \frac{\pi}{\sigma \Delta^2} \end{aligned} \quad (20-42)$$

这两个公式分别给出了枝晶主干生长速率和曲率半径与生长驱动力的关系。由这两个公式得到

$$\Delta = \sqrt{\pi p} \quad (20-43)$$

其中 p 与式 (20-28) 中定义的完全一样。在 $\Delta \ll 1$ 情况下, 这个解与二维 Ivantsov 的解是一致的, 即

$$\Delta = 2\sqrt{Pe^p} \int_{\sqrt{p}}^{\infty} dz e^{-z^2}$$

20.4 分形的概念和分形聚集

迄今为止,人们对一类具有平移不变性的物质,如单晶体的物理性质已经作了深入地研究,有了较为透彻的了解。然而,对另一类具有标度不变性的物质或形态,如具有分形特征的尘埃聚集体、宏观尺度上河道和海岸线的分布等,则研究得不多。物质世界的很多形态具有分形的特征,大到星云,小到微米甚至纳米尺度的晶核,都能观测到分形的存在。物理学家大量地介入分形形态的研究始于20世纪80年代初,当时T. A. Witten和L. M. Sander开创性地提出了扩散限制聚集(diffusion-limited aggregation)的分形生长模型。这个模型很快被物理学家所接受并用来解释各种与分形有关的生长和聚集现象。物理学家对分形的兴趣主要因为两方面的原因。首先,在很多

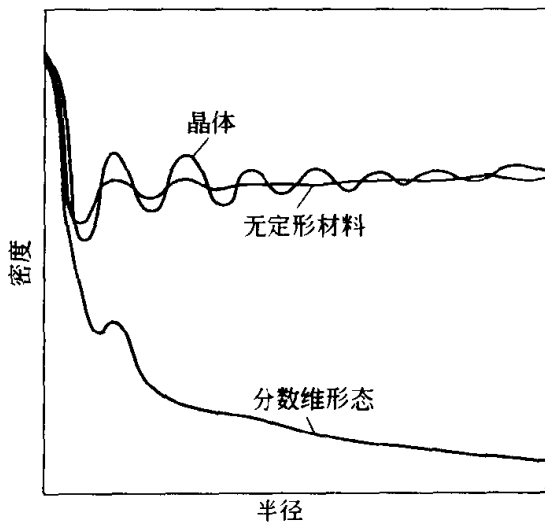


图 20-9 不同形态的物质密度与测量尺寸间的关系

多一级相变过程中,新相的成核和生长都是以分形为特征的,尤其是在生长驱动力较大的情况下更是如此(见图20-2)。研究分形的形态特征、生长过程和生长界面的动力学行为,对研究相变中涉及的形态学和动力学问题显然是很有帮助的。另一方面,具有分形特征的物体的一些性质使得这类材料可能具有潜在的应用。众所周知,通常晶体或非晶物体的密度在一定尺度以上具有一恒定的值,它不再随尺度的改变而变化;然而对分形物体,

其质量 $M \propto R^{D_f}$, 因而其密度 $\rho \propto \frac{R^{D_f}}{R^D} = R^{D_f-D}$ 。随着 $R \rightarrow \infty$, 密度是趋于0的, 见图20-9所示。此外, 对具有分形特征的物质来说, 其他的物理性质, 如原子散射因子, 振动谱等都可能发生改变, 从材料研究的角度看, 这或许是设计和制备新材料的一个可能途径。

20.4.1 分形形态的表征

分形是具有自相似对称性的几何对象。它指的是: 对一类具有无穷嵌套的几何对象, 适当地取出一部分, 并加以放大, 观测者看到的结果与整体对象完全相同。分形可以分为两类。一类是规则分形, 又称决定论的分形。它

是按一定规则构造出的具有严格自相似性的形态。图 20-10 显示了嵌置在二维欧氏空间的塞尔宾斯基铺垫 (Sierpinski Gasket)。它是用如下迭代过程得到的：先取一单位长度的等边三角形 (对应 $k=0$ 级)，将每边中点连接起来，形成四个较小的等边三角形，然后将中间一个挖去。这样余下的三个边长为 $1/2$ 的等边三角形构成 $k=1$ 级结构。继之，对这些三角形实施同样的连接、分割和取舍操作，便得到 9 个边长为 $1/4$ 的等边三角形。如此重复，最终形成具有各种大小空隙的等边三角形集合。其他具有严格自相似性的分形的例子还有康托线段以及考克曲线等。另一类分形是所谓的无规分形，比如在界面生长和许多物理问题中产生的一些图案或构造。其特点是不具有严格的自相似性，只是在统计意义上是自相似的。一个典型的例子是连续相变的卡丹诺夫 (L.P. Kadanoff) 图像：考虑一个伊辛 (Ising) 自旋系统。自旋间存在着各种尺度 (从晶格长度量级到无穷大) 的涨落区。这些涨落区互相嵌套，犬齿交错。如果用不同倍数的显微镜去观测，你会发现每个局部区域的内部结构与整体相似^[6]。

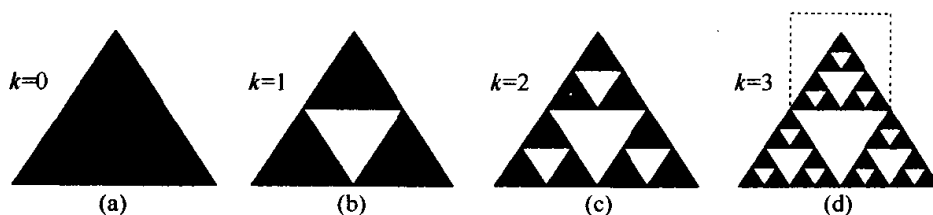


图 20-10 二维欧氏空间的塞尔宾斯基铺垫

分形可以通过其维数来表征。在 d 维欧氏空间，如果我们将尺度放大 l 倍 (或缩小 $1/l$ 倍)，对于 D 维的几何图像将能获得 k 个原来的图形。豪斯道夫 (Hausdorff) 用

$$D = \frac{\ln k}{\ln l} \quad (20-44)$$

来定义几何图形的维数。 D 一般被称为豪斯道夫维数，它不一定限于整数。

对于前面提到的塞尔宾斯基铺垫，利用上述定义，我们可以得到 $D = \frac{\ln 3}{\ln 2} \approx 1.585$ 。对于无规分形，我们几乎没有可能利用上述定义来表征它。通常采用检验标度不变性的办法来检验一个不规则形态是否是分形。为此我们定义密度-密度相关函数

$$C(r) = \frac{1}{V} \sum_r \rho(r + r') \rho(r') \quad (20-45)$$

它代表相距为 r 的两个点的密度关联，它是数学期望值。设集团内有 N 个粒子 (粒子数等于体积)。式 (20-45) 给出了如果在 r' 处存在一个粒子，则

在 $r + r'$ 处找到一个粒子的概率。如果 r 属于集团, 则应有 $\rho(r) = 1$, 否则便为 0。通常的分形是各向同性的, 关联不依赖于方向, 因此密度-密度关联函数只与 r 有关。如果关联函数在一个任意因子 b 的长度重标下是不变的, 这个对象便具有标度不变性, 即

$$C(br) \propto b^{-\alpha} C(r) \quad (20-46)$$

此处 α 是大于 0 小于欧氏维数 d 的非整数。显然只有当 $C(r) \propto r^{-\alpha}$ 上述关系才能满足。假设一个研究对象中半径为 L 的范围内的粒子数为 $N(L)$

$$N(L) \propto \int_0^L C(r) d^d r \propto L^{d-\alpha}$$

由此我们很容易看出物体的维数 D 可表为 $D = d - \alpha$ 。这也就是说, 对于 d 维空间的一个分形, 我们只要通过密度-密度关联函数计算出 α , 就可以得到其分形维数。

20.4.2 由扩散控制的分形聚集过程^[7]

20 世纪 80 年代初 Witten 和 Sander 提出了扩散限制聚集 (DLA) 分形生长模型。考虑一个二维晶格。首先选定一个种子 (核)。以种子为中心, 一定尺寸 R 为半径设定一个边界。此边界即为粒子源。在边界上随机选取一点, 从该点发射出一个粒子, 通过随机行走向中心处的核逼近。随机行走的步长对各个方向都是恒定的。一旦粒子到达核即停止行走并成为核的一部分。这时再次从边界 R 处随机选点, 发射粒子……。这样的过程不断重复下去, 当核长大到一定程度后, 我们可以发现它形成了无规分叉的分形形态。图 20-11 (a) 显示了计算机生成的一个由 3000 个粒子组成的分形聚集体。它和枝晶有相似之处, 但缺乏明显的各向异性。为定量地检验形态的标度不变性, 对该分形的密度-密度关联函数进行测量。对于多个簇的测量结果, 如图 20-11 (b) 所示, 数据可用 r 的非整数幂函数来拟合。如上节所述密度-密度关联函数的幂指数与分形的维数有线性关系 ($D = d - \alpha$)。由于无规行走的粒子一到达聚集体即成为聚集体的一部分, 这相当于生长驱动力极大、系统远离平衡态的情形。真实的生长系统中有可能出现非线性的生长几率分布。因此, 在后来的模拟中, 人们引进了局域生长几率函数。生长几率函数可以反映粒子间的相互作用、附着点的最近邻占据数、生长界面的曲率等参量, 进而可以模拟吉布斯-汤姆逊效应对生长形态的影响。肖容福等人对 DLA 模型作了进一步地改进, 引进了界面动力学过程, 表面扩散以及生长界面对流等因素, 使得这个计算机模型更接近实际晶体生长过程。此外最初的 DLA 模拟只考虑一个聚集体的生长。真实系统中当然不只存在一个聚集体。相邻聚集集团之间的相互作用可能会对生长形态和生长动力学关系等产生影响。为此人们又提出了多于一个聚集体的 Cluster-Cluster

DLA 模型。在 Cluster-Cluster DLA 中，不仅考虑粒子的行走，而且也考虑不同聚集体之间相互移动和相互作用。

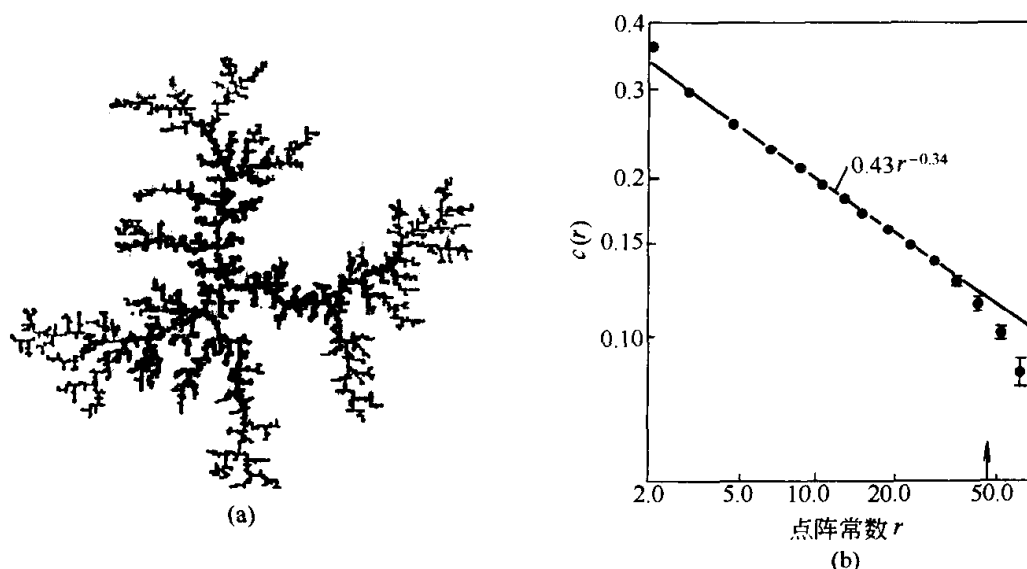


图 20-11

(a) 显示 DLA 模型生成的一个 3000 个粒子组成的分形聚集体；(b) 分形的密度相关函数

DLA 模型之所以具有如此之大的影响，主要在于它实际上描述了一类扩散场中的生长。在具体的形态生长过程中，无论是晶体生长（浓度场），还是电化学沉积（电场 + 浓度场），或是电介质击穿（静电场），都是受某个拉普拉斯场或扩散场控制的。事实上，在 DLA 中，粒子是通过随机行走向中心的核逼近的。某粒子在行走 k 步后在 r 处被发现的几率可写成 $U(r, k\tau)$,

$$U(r, k\tau) = \frac{1}{c} \sum_a U[r + a, (k-1)\tau] \quad (20-47)$$

其中 c 为某格点的近邻格点数。对于三角点阵 $c=6$ ，而对正方点阵 $c=4$ 。 τ 是每行走一步所需要的时间， a 代表粒子可能行走的几个方向。很显然公式 (20-47) 是扩散方程 $\frac{\partial U}{\partial t} = D \nabla^2 U$ 的离散形式，对于质量传输系统，

$U(r, t)$ 可以看作是规一化的浓度场 $U(r, t) = \frac{C(r, t)}{c_{\max}}$ ，其中 c_{\max} 是粒子源

处的浓度。当 $U(r, t)$ 随时间变化缓慢时， $\frac{\partial U}{\partial t} = 0$ ，故有 $\nabla^2 U(r, t) = 0$ 。

描述无规分叉形态生长的另一个模型是由 Niemeyer, Pietronero 和 Wiesmann 提出的静电击穿模型 (DBM)。这个模型最初是为了描述复杂的放电图案，后来人们发现它具有更深刻的物理含义。DBM 中第 i 点的生长几率 P_i 可表示为

$$P_i = \frac{|\Psi_i|^\eta}{\sum_i |\Psi_i|^\eta}$$

其中 Ψ_i 是第 i 点处的电势, η 是一个参数。当 $\eta=1$ 时 DBM 就变成 DLA; 而当 $\eta=0$ 时 DBM 变成人们熟悉的 Eden 模型。

20.4.3 界面能对聚集形态的影响

维塞克 (T. Vicsek) 在 DLA 模型的基础上, 在界面生长中引入表面张力, 考虑了吉布斯-汤姆逊效应对界面生长的影响。当两平衡的界面是曲面时, 其平衡温度和浓度等参量是界面曲率的函数。对于一个随机起伏、不断分叉的生长界面, 其界面上各点的化学势将是不等的, 因而导致沿着生长界面的质量和能量的输运。其结果是生长界面上一些细微的沟槽和港湾被填平, 整个界面趋向平滑。在具体的计算机模拟实验中, 维塞克让粒子以无规行走的方式 (对应于非局域的扩散场) 来到聚集簇附近, 这对应于 DLA 的传输过程。但在界面过程中补充了两个规则: ①假设粒子的粘着几率是与表面的局域曲率有关, 如果粒子不粘着于表面, 就将继续扩散; ②当粒子一旦着陆以后, 就将弛豫到附近的能量更低的坐位 (具有最大数目被占近邻的坐位)。这一方法实际上结合了两种方法: 一是对于个别聚集粒子的扩散运动的直接模拟; 另一是对于凝固问题的连续统方程的蒙特卡罗模拟。这里的粘着几率与界面曲率有关 (规则①) 相当于连续统理论中吉布斯-汤姆逊关系式确定的边值条件; 而表面弛豫 (规则②) 用来模拟凝固表面上分子的调整位置的动力学过程, 这样可以导致空洞的消失, 更加致密的基体和光滑的表面。

为了模拟定向凝固, 人们可以让粒子沉积在一条线上 (而不是一个点上); 为了模拟界面以一速率运动, 可以引入有偏向的无规行走, 即定义朝向界面运动的几率 p 大于离开界面运动的几率, 所得结果如图 20-12 所示。这里呈现的胞状界面的图像与实验观测的结果颇为相似。

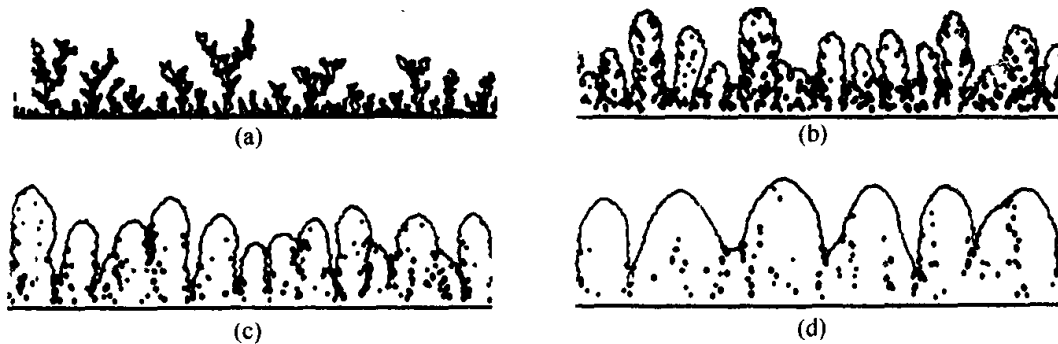


图 20-12 不同界面能条件下的生长形态

从 (a) 到 (d) 界面能逐步变大, 相应生长界面越来越光滑

参 考 文 献

- 1 Bentley W. A. and Humphreys W. J., Snow Crystals, New York: MacGraw-Hill, 1931
- 2 冯端等. 金属物理学. 第二卷. 北京: 科学出版社, 1990
- 3 Dynamics of Curved Fronts, ed. By P. Pelce. New York: Academic Press, 1988
- 4 闵乃本. 晶体生长的物理基础. 上海: 上海科技出版社, 1982
- 5 Ivantsov G. P., Dokl. Akad. Nauk SSSR, 1947, 58: 567
- 6 杨展如. 分形物理学. 上海: 上海科技教育出版社, 1996
- 7 Vicsek T., Fractal Growth Phenomena. Singapore: World Scientific. 1992

第 21 章 软物质的自组织

撰稿人 马余强

21.1 引言

物质的各种相互作用决定了自然界丰富多样的物质结构和生活，有简单的也有复杂的。通常我们认为物质结构的简单和复杂程度是由物质本身相互作用竞争的复杂因素来决定。随着系统相互作用复杂程度的增加，对抗力的出现以及由此产生的能量竞争导致系统出现失措（frustration）行为，这时系统的自由度很难同时得到满足，系统相应的状态往往出现多种可供选择的方式，分别对应于不同但相近的基态能量。因此失措的出现导致能量地形凹凸不平，在某一条件下，系统会择优地选取能量最低的状态，但也有可能出现在能量相对高些的亚稳态。如果物质是软的，易于在热涨落和外界影响下发生结构的变化。那么，决定物质结构的因素将会变得更复杂，仅仅有相互作用还不能决定物质结构的最后状态。原因是对于软物质，动力学效应和熵效应将起不可忽视的作用。因此，软物质的自组织（self-organization）行为将比通常的硬物质自组织更为复杂，这也是许多复杂软物质的自组织难以理解的原因。日常许多材料以及生命物质都是通过自组装（self-assembly）、自组织得到的。动力学效应和熵效应常常克服复杂相互作用产生的失措所引起的拓扑缺陷等，以辅助自组装、自组织过程顺利进行。由于动力学效应和熵效应的存在，自组织形成的结构转变不仅包括熵致相变，还包括流致相变和场致相变等动力学相变，分别依赖于自组装/自组织转变过程中是熵还是动力学效应起主要作用。本章系统地介绍了软物质体系中可能发生的各种自组织行为。我们从软物质的基本特性出发，系统地介绍了典型软物质系统在各种竞争因素下的自组织现象，试图从丰富的自组织现象中寻找共同规律和竞争机理，讨论如何来控制 and 设计自组装/自组织有序结构。

21.1.1 什么是软物质

顾名思义，软物质是一类柔软的物质，与固体硬物质相比，其形状容易发生变化：一方面容易受温度的影响，熵作用特别重要，而熵是刻画系统有序程度的物理量，因而软物质相有序程度的改变特别明显；另一方面容易受外力的影响，其结构或聚集体在外力作用下会发生奇特的变化，从而有可能导致材料性质发生根本的变化。这些特性决定了软物质在生物、化工技术和

日常生活中具有重要的应用前景。软物质的根本特征是在外界（包括温度和外力等）有微小的作用下，会产生显著的宏观效果，所谓小的影响大的效果是软物质体系的基本特征。实际上，这种微弱的作用会带来强烈变化的后果在日常生活中非常普遍，如在墨汁中加一点阿拉伯胶就能使之稳定时间大大延长，一点红卤就能使豆浆变成豆腐，几滴洗洁精会产生一大堆泡沫，一颗纽扣电池可以驱动液晶手表工作几年等等。在外界作用下软物质结构之所以发生根本的变化从而导致戏剧性的效果，表明通常所指的软物质是有其内部结构的。人们也通常用复杂液体或结构液体来表示软物质，说明软物质在其柔软的背后存在着复杂的、有结构的特性。软物质大多来自于有机物质，其结构常常介于固体和液体之间，虽然从宏观尺度看没有像晶体结构有周期性，从原子、分子尺度看也是完全无序，但在介观尺寸下存在规则的结构。软物质之所以表现出与固态和液态不同的特性在于介观尺度下这种有序结构的出现：一方面决定流体的热涨落和动力学相互作用支配着系统的行为；另一方面介观尺度下受约束结构显示出类似于固体的行为。其共同的作用支配和操纵了软物质独特的性质。简单液体并不是软物质，其原因是这类液体在任何条件下不会形成一定有序程度的结构，从而也很难在外界作用下引起结构的变化而导致性质的变化。一般说来，对于两种以上物质组成的混合物材料，有可能自身或外界作用下形成某种特定的结构。对于一种物质组成的软物质，其所组成的分子必须是有某种特定的结构或功能——所谓超分子，例如液晶分子和一些活化剂以及生物大分子等。这些分子自身可以通过自组装而形成各种有序的结构。实际上，与简单液体相比，软物质除了简单液体所特有的平移自由度外，常常还有一个或更多的附加自由度。如对于高分子聚合物，会出现附加的构像自由度；对于液晶，则存在附加的取向自由度。因此，这些分子由于附加自由度引起自身可以自组织，最后形成附加自由度为有序的结构。然而这种附加自由度引起的作用（如构像熵或取向熵）与其他相互作用竞争往往使得软物质结构研究变得非常复杂，这也是许多复杂软物质的自组织难以理解的原因。常见的软物质如高分子聚合物、液晶、胶体、乳状液和超分子材料（如双亲分子）以及一大类生命物质等等。由于材料中超分子或大分子材料的聚集或自组装，常常形成各种有序程度不同的结构，显示出许多独特的功能。软物质由于各种复杂的有序结构形成，使得系统性能发生很大变化，如油水系统加入双亲分子将大大改善这类系统的流动和湿润特性。下面举一些常见的软物质体系。

(1) 聚合物 一些简单的反应性碳氢化合物单体通过重复多次聚合反应而形成的长分子链结构，参见 3.3 节。单体可以相同，也可以不同，从而形成丰富多样的聚合物和共聚物（见图 3-17）。共聚物中最简单的是两嵌段共聚

物，即 A、B 两种不同单体构成的聚合物子链（链长数分别为 N_A 和 N_B ）以共价化学键相连形成 A_fB_{1-f} ，这里 $f =$

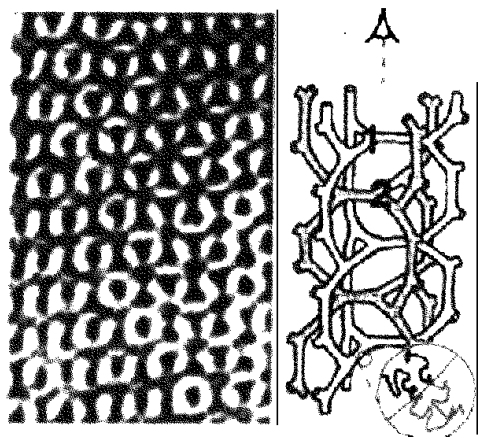


图 21-1 嵌段共聚物的双连续畴结构
(引自 Hasegawa H et al. *Macromolecules*.
1987, 20: 1651。畴尺寸为 100nm 左右。
右边草图为这种三维结构形成的示意图)

$\frac{N_A}{N_A + N_B}$ 。由于聚合物包含很多结构单元，因此，结构单元间的复杂相互作用

焓以及聚合物链的构像熵之间的相互竞争，决定了高分子链复杂的构形。同时，高分子聚合物溶液由于流动性导致的动力学行为，使得高分子系统成为极富挑战的研究对象。不同的高分子聚合物常常不相溶，在一定条件下不同聚合物一般要发生相分离。但如果是相分离材料构成的嵌段共聚物（如 A_fB_{1-f} ），由于不同聚合物的共价连接使宏观相分离得到抑制，出现微观小尺度下的条纹结构，

如图 21-1。不管怎样，这些条纹的花样结构取向常常是无规的，如何形成有很好取向的层状结构，从而实现有用的器件是材料设计的一个重要问题。

(2) 胶体 胶体系统通常认为是由某类物质构成的微小颗粒分散在另一类溶液中。颗粒的尺寸一般在几纳米到微米之间，有球状、柱状或其他形状。

这样的颗粒比原子的尺度大很多，使得量子效应并不重要，但同时又足够小，在常温下可以出现布朗运动，从而不会在引力的作用下很快沉淀。胶体系统广泛存在于自然界，它们可以是固体颗粒在溶液中，如颜料；一种液体在另一种液体中，如牛奶；固体颗粒在气体中，如烟雾。上世纪初开始的第一个关于胶体的实验，表明胶体粒子分散在溶液中的行为很像理想气体。人们知道通过胶体形成晶体的一个最著名的例子是珍珠蛋白石，蛋白石由规则堆积的二氧化硅粒子组成，其晶格粒子间距大约为光波长，由此而产生的光学相干效应形成蛋白石丰富而美丽的颜色。实际上，

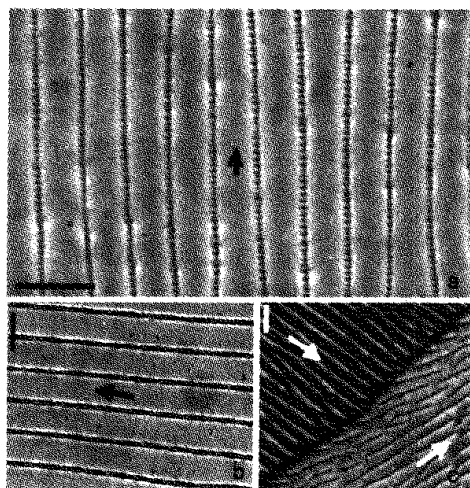


图 21-2 硅油胶体分散在液晶中
形成高度有序的长链

(a) 液晶 98%，硅油 2%，长链宽为 $50\mu\text{m}$ ；
(b) 液晶 98.4%，硅油 1.6%，长链宽为 $35\mu\text{m}$ ；
(c) 通过衬底形成不同方向的长链。液晶 98.4%，硅油 1.6%，长链宽为 $60\mu\text{m}$ (引自 Stutzmann N et al.

Nature.2000, 407: 613)

胶体系统的相行为研究是一项重要的技术问题，其主要原因是胶体粒子之间的相互作用可以通过调节粒子的表面以及改变溶液的特性来控制。另一方面，自然出现或人工合成的胶体粒子有不同范围的形状和尺寸可供选择。近来胶体晶化已成为自组装纳米和亚微米材料制备的一个重要过程。图 21-2 给出了在各向异性的液晶体系中胶体粒子自组装形成的纳米线结构。与原子构成的系统相比，胶体的相变和自组织研究有明显的优越性：胶体粒子的尺寸允许有原子分辨率的光学显微镜直接进行观察；由于同周围溶液的摩擦，胶体粒子运动很容易慢化，其弛豫过程便于实时观察。

(3) 活化剂 活化剂易于活跃在不同相界面，以减少界面能。典型的活化剂如双亲分子，其两端（极性头和碳氢尾巴）分别具有亲水和疏水特性，通过共价化学键牢固地结合在一起，参见 3.4 节。两嵌段共聚物把不相溶的两种聚合物子链 A 和 B 结合在一起，并各自分别浸润 A 和 B 聚合物，因而也可以看成一类活化剂。在溶液中由于活化剂的存在会自组装形成各种有序的结构，如胶团（见图 3-27）、泡（vesicle）、微乳状液（microemulsion）和膜（membrane）以及随超分子浓度不断增加，出现更复杂的结构，如六角（hexagonal）柱状和层状（lamellae）结构等。图 3-26 已给出了随着双亲分子浓度不断增加形成各种不同有序结构的示意图。我们知道，有极性分子组成的物质（如水）和有机分子组成的物质（如油）不相溶，会发生相分离，但如果在这类系统中，加入适量的双亲分子活化剂，会形成介观尺度下油水活化剂构成的微乳液结构，从而引起浸润等性质发生根本变化，完全不同于简单液体或混合物。

下面，我们将以聚合物、胶体和活化剂 3 种典型的软物质为代表，说明软物质在互作用能量竞争，动力学相互作用和熵驱动下出现的各种丰富的自组织现象。我们将特别强调软物质自组装/自组织过程中，由于动力学效应和熵驱动对软物质结构所带来的戏剧性的变化。

21.1.2 软物质怎样表现

日常生活中，人们面对的是分子所构成的聚集态，而且物质的性能往往通过这样的分子集合体来表现。从某种意义上讲，分子结构只是间接影响物质的性能，而正是这种介观尺度下的聚集态才是直接影响其性质的因素。构成软物质的基本分子由于其特有的结构，容易自组装聚集成介观尺度以下的有序结构，如层状、膜和液晶态等，从而体现出相应的宏观独特行为。不同分子由于自组装导致的聚集，会产生其相应的功能，即所谓物质自组装产生某种特定的功能。因此，决定软物质性质的，不仅仅是组成它的分子本身，更大程度上依赖于这些分子所经过的自组装过程：性质和功能来自于自组装过程。所谓自组装，就是在一定条件下，分子在溶液中通过空间自组织自发

地产生一个结构确定、具有一定功能的微观有序结构的过程。实际上,软物质的一个非常重要的响应就是通过分子自身自组装、外界驱动或者熵作用下在空间区域形成一种相干的有序结构,即所谓的空间自组织结构。常见的自组织有序结构有空间取向或周期有序出现,如液晶。另一类自组织有序结构是所谓的标度对称性(参见第7章),即空间自相似结构,这类有序结构广泛存在于非平衡的自组织现象里,特别发生在软物质自组装/自组织演化后期。一个典型例子是柔性聚合物在溶液中自由伸展,最后会形成空间自相似的结构。

材料的用途取决于它的性质,而性质则是由它形成的结构决定的。所以,研究软物质的自组装、自组织是研究它的结构形态、性质和功能、应用的基础。软物质由于柔软、易于自组装、设计和控制的特点,对新材料制备提供了大量的挑战和机遇。但不像硬物质,软物质自组装形成的结构稳定性将会成为软物质的一个重要问题。一方面由于动力学效应和熵驱动会影响软物质的重构,甚至有可能形成一种完全不同的新结构;另一方面在外界驱动下形成的结构,一旦外界作用消失,是否仍会继续保持长时间的稳定。在软物质自组装/自组织趋向有序的途径中影响的因素极其复杂,与固态硬物质相比,软物质如复杂液体具有高的流动性和热涨落,动力学效应和熵效应的共同参与,使得软物质自组织结构既复杂又丰富。特别是人们通常认为动力学流动、熵效应和涨落这些因素是不利于系统出现有序结构的,而恰恰在软物质系统中,我们往往观察到由于一定程度的动力学效应、熵效应和外界涨落的引入,软物质系统在纳米和亚微米尺度下形成高度有序的结构。

21.2 自组织的原理和控制

21.2.1 自组织趋向有序途径

热力学第二定律要求封闭的物理和化学系统往熵不断增加的方向,即往无序方向发展,最后的状态对应于自由能取极小值。这类极值原理对自然的解释是非常有用的。尽管如此,我们周围的大多数系统如物理、生物、经济和社会系统都是开放的。因此,这类系统通常由自组织结果来描述,其最后的自组织结构也许不是最稳定的模式,如在生长和聚集过程中所看到的许多现象。软物质的最显著特征是它能自组装/自组织形成一种空间相干的有序结构。软物质在各种相互作用能、熵作用和动力学效应下自组装/自组织形成的各种有序结构广泛存在于平衡和非平衡条件中。在趋向有序的途径中,与固体硬物质相比,无论在平衡态或非平衡态条件下,其自组装/自组织过程中熵与熵作用的竞争有明显不同。本书1.3.3节已对有序-无序转变过程中能量和熵的竞争作了详细地分析和描述。这里我们将分别从平衡态和非平衡态情形讨论能量和熵竞争对软物质和硬物质在自组织过程中趋向有序的不同行为。

(1) 平衡态 对于有相互作用的系统，热力学平衡下的最后状态由自由能 $F = H - TS$ 最小决定。这里 H 是体系的焓， T 是温度， S 是熵。在固体硬物质中，与熵 S 相比，焓 H 常常起主导作用，某种近似下可以认为焓决定系统的平衡态有序结构，所谓能致相变。与固态硬材料相反，软物质体系内能与 TS 相比或许很小，或许在软物质状态变化过程中几乎保持不变，因而在热力学平衡下，确定平衡态有序结构的自由能最小要求熵最大而不是焓最小，即所谓熵致相变（图 21-3）。例如，对于胶体溶液的相行为，我们可通过计算系统的自由能最小来决定。如果胶体颗粒看成是硬球粒子，只要粒子不接触，互作用能等于零。因此，胶体溶液的相行为可完全看成由熵来决定。熵最大决定了胶体系统的相行为。偏离熵最大引起熵的驱动产生熵致相变，导致系统出现相应的胶体晶化。如果系统是液晶，构成液晶的分子可看成是一种硬棒，即棒状分子，它可能会存在两种有序构形：分子位置的有序和分子排列取向的有序性。我们在下节会分析，发现在熵驱动下也会形成液晶有序相，即所有分子沿着一定的取向排列。此外，熵致相变也会发生在更复杂的软物质体系，如由不同形状如硬棒和硬球组成的混合系统。由此可见，与传统固体物理相比，熵在软物质物理特别是自组织结构有序的研究中具有特殊的地位。

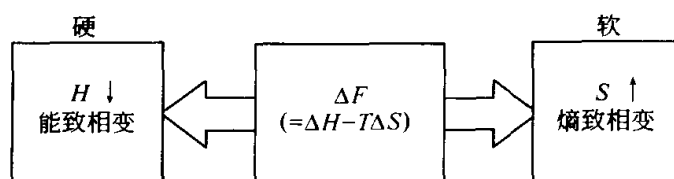


图 21-3 热力学平衡下能量和熵的竞争

(2) 非平衡态 从非平衡动力学角度分析，设想系统通过来自外部热涨落或外力的能源接触，平衡态被破坏，如图 21-4，对于硬物质，由于能量 ΔE 的加入，导致体系焓增加，于是要求非平衡动力学演变过程中熵增加以达到最后的自组织状态。因此硬物质系统由于能量加入系统往无序方向发展。例如一块晶体的平衡状态由自由能 $F = H - TS$ 的极小值决定。正如 3.3 节分析，如果加上外力来拉伸晶体，就偏离其平衡态，自由能增大。外加拉力迫使晶体的原子间距增大，即外力对抗了原子间作用做了功，因此外界提供的能量分布在晶格之间的谐振子中导致其固体焓 H 增加，而原子间距不同程度的拉伸所产生的构形对应于熵增加，即往无序方向发展；对于软物质，由于能量的加入，内能几乎没有变化，在动力学演变过程中势必要求熵变小以实现最后的自组织状态。因此软物质系统由于平衡态的打破，在动力学演变过程中常常以无序的方式却往有序的方向（熵变小）发展。如对橡胶拉伸情况与晶体拉伸就不一样（见图 3-21），拉力并没有迫使橡胶分子中

的原子间距改变，而是使交联点间的分子线段变直，即外力无法对体系内能有贡献，惟一的办法是调整聚合物构形，使弯曲的分子线团拉直，从而使分子线段的位形熵减小，也就是说拉伸的结果是使有序度增加。这种偏离平衡位置（熵最大）引起的熵改变会产生一种有效力，所谓熵力，与来源于外界的力同样真实。我们用手拉伸橡皮筋时，其恢复力主要来源就是这种熵力。而外力对抗熵力作了功。这里橡胶弹性形变是和熵联系在一起的，外力做功导致熵的减小，基于熵变化的弹性理论，我们通过聚合物链作为例子来计算这种熵力大小。聚合物可以看成由 N 个长度为 b 的单体结合的长链，这类长链分子通常具有高度的柔软性并因此而具有很大的熵值，大多数情形下每个单体除连接部分外相互之间不存在相互作用，而且连接部分的相互作用能与它们的相对夹角无关，因此系统在自由状态下的位形由熵的极大值决定。

若将链的两端分开 R 的间隔，将导致系统熵减小 $\Delta S = -\frac{kR^2}{2Nb^2}$ （这里 k/N 可以看成聚合物的熵弹性常数，在 Gauss 型链情形， $k=3$ ），即拉伸长链分子会导致系统熵减小，从而得到自由能增加 $\Delta F = -T\Delta S = \frac{kTR^2}{2Nb^2}$ ，系统产生的恢复力——熵力为 $f = -\frac{\partial F}{\partial R} = -\frac{kTR}{Nb^2}$ 。对于更复杂的生命聚合物如 DNA（参见图 1-27），拉伸对抗熵力做功，将会引起更丰富和有趣的现象。

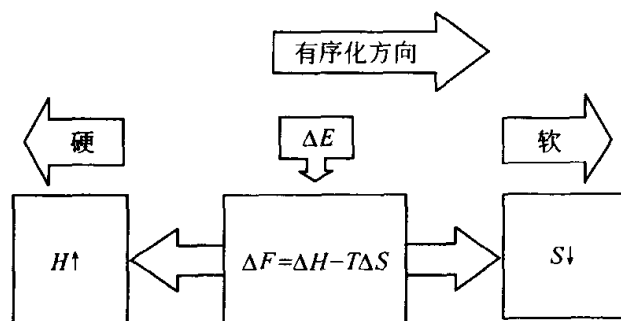


图 21-4 软物质和硬物质在外部驱动下由于体系焓和熵的竞争引起的自组织有序过程

此外，在软物质系统的非平衡过程中，随着外界能源的加入，也有可能导致软物质结构重组形成高度有序新结构的同时，其总能量也相应升高。这种亚稳态能否维持长时间稳定是构成软物质自组织结构稳定性研究的重要问题。总之，在外界驱动下结构或聚集体进一步有序化表明软物质趋向有序途径的方式有其特有的特征。

21.2.2 软物质自组织的形式

由于软物质的柔软性以及构成软物质单元本身的自组装能力，使得软物

质在相互作用、熵和外力驱动下显示出丰富多彩的自组织现象。常见的有：超分子如双亲分子等活化剂的自组装和聚集；单（大）分子自组织，如高分子构象和生物大分子折叠；熵驱动下的自组织（熵力和熵致相变），如胶体聚集和液晶相变；非平衡动力学自组织，如场致相变、流致相变、远离平衡自组织临界等。

(1) 超分子自组装 一些分子如双亲分子具有自组装能力，在溶液中会形成介观尺寸以下的结构。当溶液中双亲分子少量时，由于混合熵效应，不会形成胶束。随着双亲分子单体浓度不断增加，双亲分子由于碳氢尾巴疏水作用的影响，双亲分子容易跑到溶液表面或容器壁以尽可能避免碳氢尾巴与水接触。当双亲分子浓度达到某一临界值，即临界胶束形成的浓度（CMC）时，胶束开始出现：双亲分子疏水尾巴聚集在胶束内避免与水接触，而极性头在胶球外面与水接触，于是在水溶液中形成由胶束组成的超分子溶液。

我们考虑溶液中胶束形成的临界浓度以及有限尺寸的胶束在溶液中的稳定性。假定在溶液中已形成这些聚集的胶束，并与双亲分子单体共存维持平衡（如图 21-5），这要求双亲分子的化学势 μ 在单体和不同的聚合体中相等，

$$\mu = \mu_1^0 + kT \lg X_1 = \mu_N^0 + \frac{kT}{N} \lg \left(\frac{X_N}{N} \right) = \text{常数}, N = 1, 2, 3, \dots \quad (21-1)$$

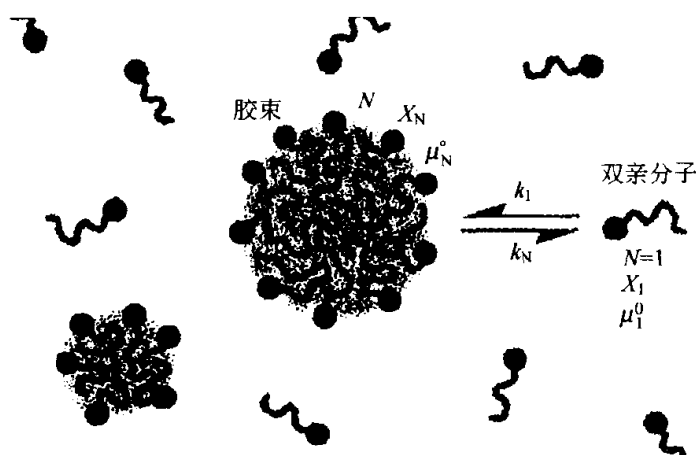


图 21-5 N 个双亲分子联结形成一个胶束，
胶束与两亲分子共存的示意图

这里 μ_N^0 是 N 个分子构成的聚集体中的分子标准化学势（每一个分子的平均相互作用能）， X_N 是 N 个分子构成的聚集体中分子浓度（严格地说是激活度）。实际上，从方程（21-1）不难看出，化学势贡献来自内能和熵两部分。当溶液中双亲分子单体与 N 个单体构成的胶束共存达到平衡时，式（21-1）可以写成

$$X_N = N \left\{ X_1 \exp \left[\frac{\mu_1^0 - \mu_N^0}{kT} \right] \right\}^N \quad (21-2)$$

如果在不同尺寸下的分子和周围环境经历相同的相互作用，即 μ_N^0 的值与 N 无关，于是式 (21-2) 简化为

$$X_N = NX_1^N \quad (21-3)$$

由于浓度 $X_1 < 1$ ，因而从式 (21-3) 发现 $X_N \ll X_1$ ，因此分子只能以单体形式存在。只有当分子在单体和聚集体有不同的结合能时，有可能形成聚集的胶束存在。从式 (21-2) 可以发现，聚集的胶束形成的必要条件是 μ_N^0 随着 N 增加必须减少，才有可能形成条件

$$(X_1)_{\text{CMC}} \cdot \exp \left[\frac{\mu_1^0 - \mu_N^0}{kT} \right] \approx 1 \quad (21-4)$$

相应地，聚集的胶束形成时对应的单体临界浓度是

$$(X_1)_{\text{CMC}} = \exp \left[\frac{\mu_N^0 - \mu_1^0}{kT} \right] \quad (21-5)$$

式 (21-5) 给出了双亲分子溶液中形成胶束时双亲分子单体所需要达到的临界浓度。我们知道，如果单体之间仅仅是简单的吸引相互作用会使 μ_N^0 随着 N 增加单调减少，导致单体在溶液中不断聚集，发生相分离，最后形成由单体组成的无限大聚集体。例如，在水溶液中加入油或者链烷分子到临界浓度，就会从水溶液中分离出体相，形成油或者链烷球滴，随着油或者链烷浓度的不断增加，会不断聚集，直到最后即发生宏观相分离。那么，是什么力使得这类双亲分子构成的超分子溶液中形成的聚集体尺寸是有限的呢？详细分析发现双亲分子形成的胶束尺寸有限，不是因为它们的溶解性和 CMC 值，而是因为双亲分子亲水和疏水作用构成的自组装能力：随着 N 增加， μ_N^0 单调减少，当 N 达到某一值，由于双亲分子的亲水和疏水特性， μ_N^0 将不再继续降低，会存在一个最小值或者不变值，对应于聚集体不再继续堆积，从而使得有可能形成有限的聚集体。与固体胶球或刚性 DNA 分子不同，即使是这些胶束也是软的和柔韧的，因为维持双亲分子形成的胶束并不是强的共价或离子键作用，而是来自弱的范德瓦尔斯力和氢键作用。随着溶液条件的改变，不仅会改变溶液中胶束之间的相互作用，而且会影响每一个胶束中分子间相互作用力，导致结构本身发生变化。随着双亲分子浓度不断增加，也会出现图 3-26 所示更为丰富的其他有序结构。

(2) 单（大）分子自组织 自组织可以在一个简单的分子（例如聚合物或 DNA 分子等）内自身实现。如 DNA 分子在自身抵抗扭曲和弯曲的过程中会出现自组织形式。如果这类分子头尾连接成环，自组织结构是平面内圆形环，但如果是扭曲的环，出现的形状不再是平面的圆环，而是分子围绕自

身扭曲形成的非平面内盘绕的环状结构。在 DNA 分子内某一点的扭曲会引起分子构形整体的变化。另一个典型例子是蛋白质的折叠。蛋白质是一个复杂的生物大分子，在溶液里显示出由它们的序列决定的三维结构。最简单的蛋白质是肌红蛋白（参见图 1-30）。实际上，分子生物学的一个重要问题是确定怎样从序列转变成结构，即从一维无活性的多肽链如何折叠成一个三维活性的蛋白质。一个给定系列的氨基酸组成的异质聚合物折叠成惟一的一种具有生物功能的三维几何形状的折叠过程中，其相互作用的竞争是极其复杂的，包括碳氢链上的亲水、疏水作用的竞争，局域和空间（非键-键）相互作用的竞争。其折叠过程中从大量的构形约化成一种自然的构形，反映着多肽链的构形熵对折叠过程起着重要的作用。后期折叠态由于周围动力学效应的辅助，促使它快速形成稳定的有序结构。因而，蛋白质自组织的有序化过程反映了相互作用竞争、熵效应和动力学相互作用效应的共同作用。我们可以说内能和熵的竞争以及动力学流动构成生物大分子折叠的物理效应，而遗传密码构成折叠蛋白质的生命内容。P.G.Wolynes 和 H.S.Chen 等用蛋白质折叠的能量漏斗结构对蛋白质折叠的物理过程中熵和能量角逐作了很好的解说。他们认为蛋白质折叠过程中就像如图 21-6 的能量位形中从上到下面最底能谷的一个扩散过程。能量面是以蛋白质分子不同构像的自由能作为自由度展开的地形。如图 21-6，横向尺度 S 表示位形熵的大小，纵向尺度 E 和 Q 分别表示能量和有序度的大小。其漏斗顶部对应于未折叠状态，位形熵 S 最大。漏斗底部的最底能谷处对应于自然的折叠态， Q 最大，随着蛋白质从一维多肽链折叠成三维的活性蛋白质，体系的位形熵逐渐变小，蛋白质分子不同组态的自由度降低。为了使系统自由能变小，必须通过内能的减少来补偿。最后折叠到惟一组态的紧致的自然蛋白质折叠态。在整个折叠过程中，由于相互作用竞争引起的能量失措（表现在图 21-6 中能量起伏引起的各种小能谷）导致热力学过程中蛋白质折叠并非一帆风顺。实际上，蛋白质这类软物质由于动力学相互作用效应的存在，却辅助它更快地折叠到自然的蛋白质结构。这是由于动力学效应的存在会尽可能抑制由于能量相互作用引起的失措。蛋白质折叠是一个相互作用竞争、动力学效应和熵效应一起参与的典型过程，是目前一个富有挑战的问题。

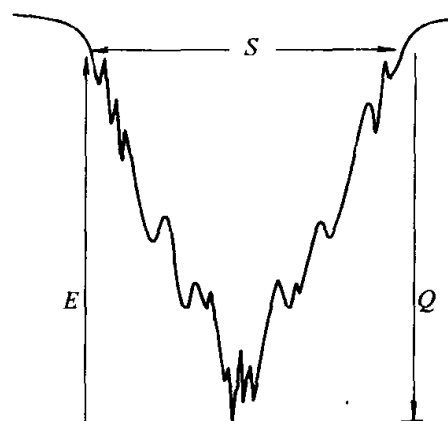


图 21-6 蛋白质折叠的能量漏斗构形
(引自 P.G.Wolynes)

(3) 熵驱动下的自组织 熵用来度量系统中无序的程度, Boltzmann 表示式给出熵的定义

$$S = k_B \ln W \quad (21-6)$$

这里 k_B 是玻耳兹曼常数, W 是与宏观状态对应的微观状态数。无序越大, 熵越大。根据热力学第二定律, 在这种意义下表明系统从无序相到有序相的转变必须要求熵的损失由更多的内能降低来补偿。依赖于上面公式关于熵对无序程度的定义, 这种说法是完全正确的。例如我们可以直观地定义各向同性的液体为无序, 晶体为有序。从各向同性的液体到晶体的自发相变要求冻结过程必须充分降低系统的内能以抵消熵的损失。这种有序相变是能量驱动的, 即所谓的能致相变。这在很多情况下常常发生。但如果我们非得依赖式 (21-6) 来对所有宏观有序作直观定义却是一种误解。通常被认为的许多有序相变实际上是属于熵致相变, 而不是能致相变。原因是这时体系的内能无贡献, 自由能降低的惟一要求是熵增加, 即微观无序度的增加, 而导致的结果是宏观有序度的增加。这种现象会引起许多与直觉相反的现象, 例如在熵驱动下的胶体晶化和相分离等。造成这种与直觉相反的原因在很大程度上是由于对宏观有序度的定义, 实际上, 微观无序度增加也有可能导致宏观有序的增加。

为了观察这种熵致相变, 我们考虑系统内能是温度的函数, 而不是密度的函数, 即在温度不变下内能不发生变化, 这时惟一的办法只有通过增加熵来实现自由能的降低, 硬球胶体模型就是一个很好的例子。我们以胶体聚集过程中熵作用为例说明熵致相变的重要性。表面看来, 胶体出现的丰富的相行为需要了解粒子间复杂的相互作用。不管怎样, 实验的条件告诉我们, 在胶体系统中似乎只有简单的胶体接触引起的排斥相互作用在起作用。因此只有熵才是胶体有序化 (如晶化) 的驱动力。熵作为有序的驱动力似乎与人们的直觉相悖。早在 1949 年, Onsager 指出, 一个各向同性的液晶系统随着液晶棒浓度的增加, 一定会出现一个由液晶棒沿某一方向取向有序的向列晶相的转变。他认为, 在浓度足够高时, 取向有序导致的熵损失一定会有向列晶相中液晶棒更多的自由体积增加引起的平移熵增量来补偿。从各向同性的无序相到向列晶相过程中, 取向熵和平移熵的贡献不断增加。实际上, 取向几乎平行的液晶棒所占有的自由空间肯定比无规取向的棒自由空间大, 从而导致液晶棒有更大的平移熵。对由硬球组成的胶体晶化过程也可以类似地理解。当浓度很高时 (如图 21-7), 胶体晶化导致位置有序的混合熵损失被更多因胶体聚集后小分子自由体积增加引起的平移熵增加所补偿。对于硬球组成的胶体晶化和液晶棒形成的有序向列晶转变, 都说明了熵是这类排斥粒子聚集和有序化的驱动力。进一步, 由胶体球和棒组成的胶体系统将会出现更

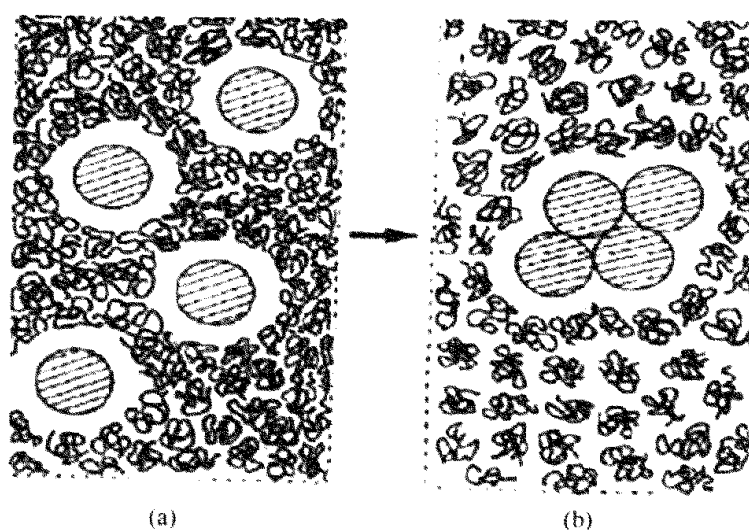


图 21-7

(a) 胶体粒子分离时由于排斥体积引起聚合物小分子浓度较高；(b) 粒子聚集后聚合物小分子浓度降低

(引自 Daniel J C 和 Audebert R, *Soft Matter Physics*, edited by Daoud M 和 Williams C E. Springer-verlag, 1999)

为丰富的相转变行为 (图 21-8)。自然, 这种相行为导致的最后状态同样要求熵最大。因此, 系统取向、位置和混合熵的损失一定会由最终有序相出现而引起更多的自由体积增加贡献的熵增量来补偿。由不同大小的硬球组成的

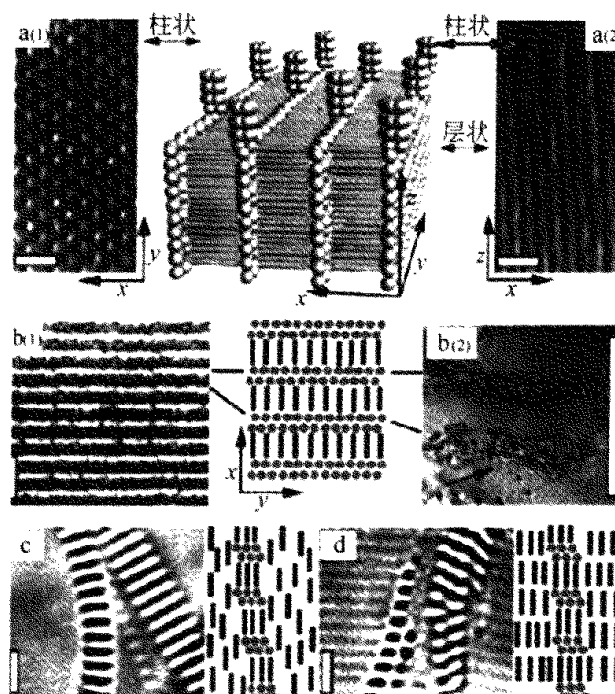


图 21-8 胶体球和胶体棒在溶液中由于熵致相变产生的微相周期结构

(引自 Adams M. et al. *Nature*, 1998, 393: 349)

混合物或者是硬棒和硬球组成的混合物中，自由体积效应是相当明显的。这对理解熵致相变是极其有意义的挑战。

我们可以来分析由两类半径分别为 r_l 和 r_s ($< r_l$) 的胶体球组成的一个简单体系。假定胶体分散在同一液体中，液体的惟一作用被假定是辅助胶球在溶液中作 Brownian 运动，而胶球之间惟一相互作用是接触时才有排斥，即系统看成是一个二元硬球混合物。于是，大球能否均匀分散在溶液中保持

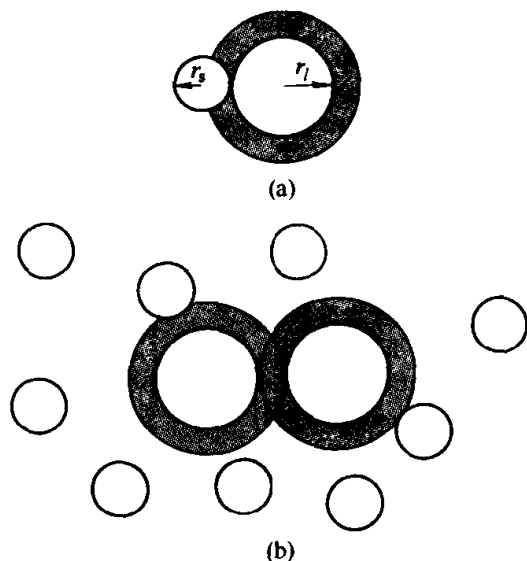


图 21-9 排斥体积引起的耗尽效应
(a) 半径为 r_s 的小球中心不能进入半径 r_l 的大球周围厚度为 r_s 的壳层；(b) 两大球重叠产生更多小球可达的体积

稳定成为一个人们感兴趣的问题。表面看来，系统混合熵的贡献要求大球均匀分布在大量小球组成的系统中是合理的。实际上，随着大小球尺寸相差足够大，在某些胶球浓度范围系统会相分离成由大球和小球组成的两相，即大球聚集或晶化。为了理解驱动这类相分离的物理机制，我们首先来考虑一个大球和一个小球接触的情形，如图 21-9 (a) 所示。由于硬球不可渗透，在大球周围厚度为 r_s 的区域对小球的质心不可达到。如果在溶液中有一些大球存在，这些大球周围厚度为 r_s 的壳层体积对小球不可达。但如果两个或更多大球接近到大球表面间距不到 $2r_s$ 时 [图 21-9 (b)]，这些排斥体积将会部分重叠，

这将增加小球可达的体积，因而导致体系总的熵增加。由于小球的数量远远超过大球，熵贡献主要来自于小球部分，如何调节大球在溶液中分布使得系统或者说小球的熵贡献最大是系统达到稳定的原因。于是，胶体溶液的熵致相变原理是简单的，通常增加溶液小球的自由体积，会导致系统平移熵的增加，但同时也降低了混合熵。在大的胶体粒子浓度较低时，由于混合熵的作用也会导致溶液相溶。随着胶体粒子浓度增加，为了保证溶液中小球有足够的平移熵，导致相分离发生，胶体出现聚集。即使粒子间相互作用是排斥的，同类粒子也会经受有效的吸引。这类现象在物理和化学中称为耗尽吸引。在目前的系统中，有效的耗尽吸引完全来自熵贡献。

最后，为了理解微观无序度的增加有可能导致宏观有序的增加，以进一步理解熵致相变，我们考虑一个由大小球组成的一维系统的简单微正则系综 (如图 21-10)，来估计可能的大小球排列方式。如果我们选择小球尺寸占一

个格点位置，而大球占 1.5 个格点间距，导致大球排斥其他小粒子占领邻接的栅格。如图小球的排列方式数有 $R = g! / s! (g - s)!$ ，这里 s 是小球数量， g 是小球可活动的栅格数。我们发现大球分离时或者大球离开墙壁时排斥体积就会增加。因而在图中从上到下不同位形下小球的可排列方式数从 21 降落到 1。这种可能排列数的折叠转变表明无规放置的较可能状态（微观状态数多）是大球靠在一起（宏观有序），如果大球再靠近墙壁将更可能。图 21-10 从下到上表明随着微观状态数（无序度）的增加，导致大球聚集，宏观有序增加。

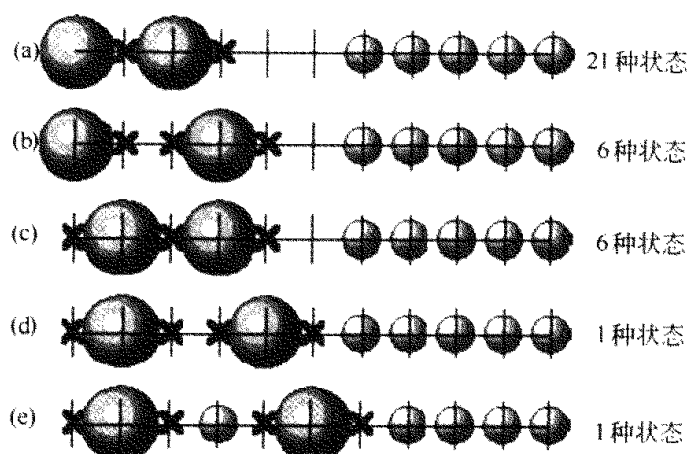


图 21-10 一维格子中熵驱动有序的例子

两个大球和五个小球形成的位形。围绕大球的排斥体积用×符号表示。

右边给出了大球不同的分布导致的小球排列方式数

(引自 Shinbrot T. Muzzio F.J. Nature. 2001, 410: 251)

(4) 非平衡动力学自组织 自然界中许多有序的结构是在非平衡自组织过程中实现的。通过外部影响或者系统本身的耗散（如化学反应体系和颗粒物质）作用驱动体系自组织形成各种丰富的结构。这类现象大多在软物质体系中得以实现。这里我们仅仅以流动驱动下的自组织现象为例。溶液中快速流动的粒子引起的流体动力学效应常常会抑制粒子作任何扩散的 Brownian 运动；而缓慢的流动也会使动力学惯性效应完全可以忽略。一个粒子运动引起的流动会影响附近粒子运动的改变，这种动力学相互作用会对结构流体产生重构的影响。如两嵌段或三嵌段共聚物在热力学平衡时可以形成宏观有取向的层状结构，如图 21-11。但是在相分离过程中通常都是通过淬火后快速有序，不会形成有某个特定取向的宏观有序结构。这种相分离过程导致的结构短程有序，但长程各向同性，非常类似于把水快速冷却到冻结点以下导致多晶态一样。为了使这类系统形成长程有序的结构，可通过施加对称破缺场获得，类似于在居里点以下通过对铁磁材料施加磁场会形成永久磁矩，或者

施加电场来产生高度取向的液晶材料等。我们知道，动力学相互作用通常会磨平系统的非均匀性，提高系统的对称性（变得更无序）。另一方面，由于系统中相互作用竞争引起的失措常常导致体系存在拓扑缺陷或无规的非均匀性，将不利于系统有序。但当适当大小的动力学相互作用出现时：一方面由于动力学相互作用磨平或抑制了体系中失措引起的拓扑缺陷，另一方面，动力学相互作用并不足够大以至于扰乱相互作用能量驱动的有序。其结果是动力学流动辅助相互作用能量驱动有序并克服了相互作用能量竞争引起的失措所带来的无序因素。

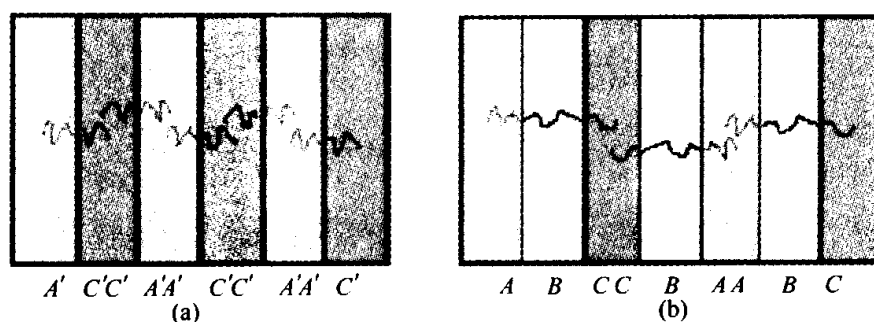


图 21-11 两嵌段共聚物 (a) 和三嵌段共聚物 (b) 形成的形貌。

成分不同的聚合物一般不相溶，所以 AA 和 CC 接触比 AC 界面接触产生更低的能量，导致两种情况下中心对称结构的形成

(引自 Gido S.P. Nature.1999, 398: 107)

在对称破缺（如剪切流场）作用下结构流体常常出现宏观有序的结构。通常对液体施加剪切力场需要花费能量，如果体系是简单液体，这部分能量转变成热能（粘致热效应），但对于结构流体，流动场驱动系统自组织形成亚稳的高能量但非常有序的结构。通过对两嵌段共聚物组成的体系两端界面施加剪切力就可以使它形成高度有序的层状结构。由于剪切频率的不同，嵌段共聚物出现不同取向的宏观排列方式，如图 21-12 所示。当剪切频率较低或较高时，出现如图 21-12 (b) 的平行层状排列；当剪切频率取中间值，取向方式变成图 21-12 (a) 的垂直层状相排列。这种有意义的结果表明如聚合物等软物质在超过一定的外部作用下会显示大的戏剧性效果，因而出现强大的非线性行为。这种在轻微不同的工作条件下，形成两种不同排列方法的手段为从同一类聚合物产生根本不同的宏观材料提供了可能。当前，这种流动驱动下的取向有序结构形成在分子科学和材料工程中是一种十分重要的手段。

剪切力也会诱导无规取向的液晶状态转变成向列晶相，甚至会产生全新的结构，例如当双层排列的近晶相结构受到一定大小的剪切流场作用，层状

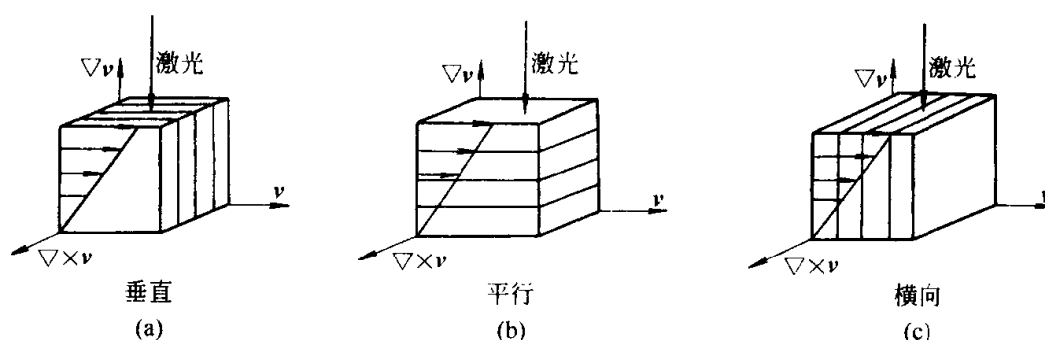


图 21-12 流动引起嵌段共聚物有序的排列。剪切频率不同分别产生平行和垂直取向的层状结构排列
(引自 Chen Z R et al. Science. 1997, 277: 1248)

结构将会重新自组织形成一个密堆积、排列规则的洋葱状结构。这种结构一旦形成后，可通过调节剪切力场大小来改变洋葱结构的层状数目和大小。一旦剪切流动停止作用，洋葱结构仍然保持相当长时间的稳定。这种流变的物理起源是相当复杂的，如何理解剪切力诱致这种转变是目前一个重要的挑战。

21.2.3 控制和设计自组织形貌的途径

自组装/自组织形成的多畴结构常常是无规取向的。而实际的功能材料要求结构具有非常好的有序或方向性。为了达到要求的取向结构，必须控制其宏观有序。通过外界驱动，如流动包括动态振荡和稳定的剪切流场是一种非常有效的方法。另一种重要的手段是通过空间取向和几何约束来改变微畴的取向。此外通过衬底诱发、聚合化、噪声、凝胶和化学反应等途径也能有效控制 and 设计自组织形貌。

如果我们要形成更小尺度（如纳米尺寸）下的结构，将需要有多种相互作用或动力学机制的竞争参与。例如，双亲分子在溶液中形成有限尺寸是由于双亲分子的亲水和疏水竞争作用；相分离系统仅仅有短程相互作用（界面张力）只能形成宏观的畴结构，长程排斥力的引入会抑制宏观相分离产生微观尺寸下的畴结构（参见第 21.6 节）。除了平衡系统中短程吸引和长程排斥的竞争，在动力学系统中也会观察到动力学机制竞争导致的微观畴结构，如在化学反应系统中，动力学快变量和慢变量之间的竞争会导致螺旋波结构。但竞争作用引起的失措会导致系统结构出现拓扑缺陷。如何消除失措引起的拓扑缺陷，从而形成高度有序的结构是自组装的关键。通过外界驱动、衬底诱发、动力学效应等的引入有可能消除拓扑缺陷。

下面，我们以衬底诱发和空间几何约束为例介绍自组织有序结构的形成。

(1) 衬底诱发 软物质体系放置在衬底上, 由于其流动性和衬底对软物质的相互作用诱发系统自组织形成有规则的结构。如果我们把两嵌段共聚物放置在平面衬底上, 当衬底分别浸润嵌段共聚物 (A-B) 的 A 和 B 子链, 将形成 A-BB-AA-BB-AA-BB-A 和 B-AA-BB-AA-BB-AA-B 的中心对称的层状排列, 类似于图 21-11 (a) 所示; 对三嵌段共聚物 (A-B-C), 当衬底分别浸润 A 和 C 子链, 将形成 A-B-CC-B-AA-B-CC-B-A 和 C-B-AA-B-CC-B-AA-B-C 的中心对称的层状排列, 类似于图 21-11 (b) 所示。这种纳米尺寸的周期结构排列技术已成为自组织/自组装方法制备应用器件, 如液晶显示, 频率调制器和光学开关的重要手段。如果在这种尺度下的材料带有极性或磁性, 将会显示许多诸如非线性光学效应等特性。因此如果构成共聚物的分子带有永久磁性, 也许会产生有永久磁性的块结构。但无论在图 21-11 中由两嵌段共聚物和三嵌段共聚物形成的层状结构里, 由于结构中心对称导致块结构磁性抵消, 如果我们通过平面衬底诱发两嵌段共聚物和三嵌段共聚物组成的混合物就会产生非中心对称的层状排列, 从而可以形成有磁性的自组装纳米结构。假如三嵌段共聚物 (A-B-C) 中的 B 分子链有极性或磁性, 通过调节嵌段共聚物的不同子链对衬底的界面特性, 将分别形成如图 21-13 所示有三种极性取向 (向上、向下和向左) 的纳米层状结构。

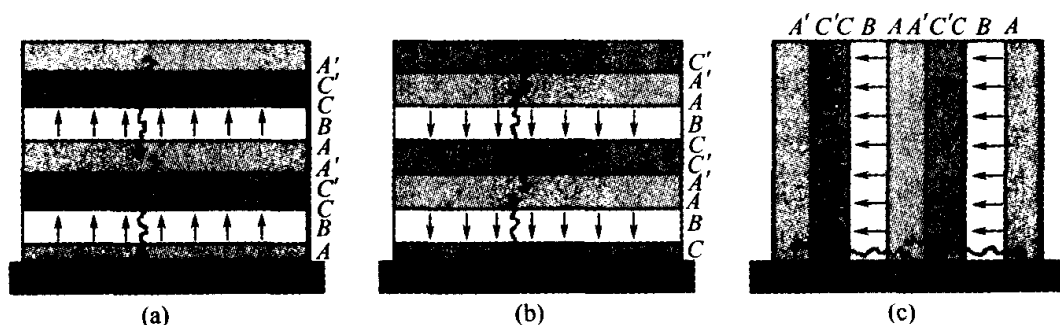


图 21-13 两嵌段共聚物 (AC) 和三嵌段共聚物 (ABC)

混合物放置在衬底上形成的形貌

- (a) 如果聚合物 A 子链和平面衬底的界面能较低, 系统自组织形成的具有极性向上的多层结构; (b) 如果聚合物 C 子链和平面衬底的界面能较低, 系统自组织形成的具有极性向下的多层结构; (c) 如果聚合物 B 子链和平面衬底的界面能较低, 系统自组织形成的具有极性水平向左的多层结构

(引自 Gido S.P. Nature. 1999, 398: 107 and Goldacker et al., Nature, 1999, 398: 137)

(2) 空间几何和取向约束 自组织/自组装常常受到空间几何和取向约束的影响。一般来说, 约束常常来自于固定的边界约束如多孔表面或人工纳米结构构成的媒体。尽管如此, 约束也可能来自系统本身的特性, 如浸润的液体膜束缚在固体表面和蒸气平衡。

两嵌段共聚物能自发地自组织形成有特定厚度的二维层状结构，每一层由同一种聚合物组成，可看成一种液体在 XY 平面内保持平移对称性，而聚合物链或多或少地沿着 Z 方向。自组织是由于构成共聚物的 A 组分和 B 组分不相溶产生的驱动。形成层状结构的目的是为了使 $A-B$ 接触面最小，即界面能最小以抵消聚合物熵的相应损失，因此，两嵌段共聚物最后自组织成 $A-B$ 、 $B-A$ 、 $A-B$ 交替堆积的多层结构。增加层厚度可以减少 $A-B$ 区域间的界面面积，从而减少界面能，但层厚度的增加迫使聚合物链拉伸导致熵减少。层的厚度 L_0 可以根据这两种相反趋势来确定，即 $A-B$ 界面能和聚合物拉伸成沿 Z 方向的弹性能（熵减少部分）达到平衡时确定。当嵌段共聚物系统放置在容器中，容器尺寸与系统尺寸相比拟时，系统形貌受到容器的扰动常常出现戏剧性的变化。通过边界面的扰动，会导致系统自由能中焓和熵的损失。一方面体相互作用，要求嵌段共聚物形成其内禀周期 L_0 ，另一方面，由于平板提供的表面相互作用，驱动共聚物按界面的作用分离。竞争相互作用会驱动系统出现至少两个不相容的结构。可能出现的模式将根据相互作用的竞争因素择优地出现。如果共聚物溶液放在一个容器底板上，与空气保持平衡时，液汽接触界面将会出现凹凸不平的峰结构。但如果我们放置对称的两嵌段共聚物构成的薄膜限制在宽度为 L 的两个平行板之间，凹凸不平的结构被抑制，约束墙与聚合物的相互作用会改变层状结构的厚度和取向。一般说来，当约束墙对构成嵌段共聚物的 A 和 B 聚合物是中性的，最低能量结构将使平衡图像是垂直于平板取向的多层结构，以便共聚物内禀周期 L_0 可以在任何平板间距离 L 下实现。当约束墙偏爱某一类聚合物时，共聚物形成的微畴尺寸和取向将变得复杂，层状结构的周期和取向相对于平板可以改变，其目的是以最大限度释放系统存在的失措。如果约束墙吸引某一类聚合物较强，层状结构的取向与平板平行，改变平板之间宽度 L 将改变层状结构的周期；如果约束墙吸引某一类聚合物较弱，当平板之间宽度 L 是层状结构的内禀周期 L_0 的整数倍，层状结构的取向与平板平行。如果平板之间宽度 L 不是层状结构内禀周期 L_0 的整数倍，平板提供的表面相互作用和嵌段共聚物体相互作用的竞争产生的失措使层状结构的取向有可能垂直于平板。

21.3 软物质的复相分离

21.3.1 软物质复相分离的热力学

前面我们介绍了软物质自组装/自组织引起的各种丰富的相行为，我们可以发现软物质如复杂流体通常是有两种以上的物质构成，而软物质各种丰富的自组织行为往往是在相分离过程中实现的。因此，复相分离已成为软物

质自组织一个极其重要的手段。为此，我们在下面以聚合物体系为例，对复相分离的热力学和动力学作简单介绍。

混合系统初始在高温时处在均匀混合的稳定区域，由于外界参数如压强或温度的突然变化，经一级相变淬火冷却到亚稳或不稳定相区域，开始进行动力学相分离，经无限长时间后，系统完全分离成具有各自浓度的两个稳定相，期间系统将进行一个非平衡的动力学相分离过程。与其他各种相互作用能量的竞争导致不同尺度下的相分离形状。生长过程中首先是相互作用能量如界面张力起作用，随着一定尺寸的畴结构形成，流体动力学效应开始起作用，辅助界面张力驱动的相分离成各种有序的畴结构。相分离过程以完全无序的方式生成有序的模式。此外，对复相分离所产生的不同尺度下畴结构的了解，将有助于我们理解其他软物质自组织形成的不同尺度结构以及相互作用竞争机制。

按照相分离生成图案的畴尺寸大小，相有序可分为三类：宏观、介观和微观相分离，如图 21-14。图 21-14 (a) 表示两种不相溶的聚合物由于界面张力（来自同类分子吸引、不同类分子排斥的短程作用）导致界面面积的减少，发生宏观相分离，最后形成宏观尺度下的两共存相。如果系统是由两种不相溶聚合物如 A 和 B 构成的嵌段共聚物，通过共价键约束抑制两种不相溶聚合物发生宏观相分离，最后形成大约 5~100nm 的微观尺度下的成分调

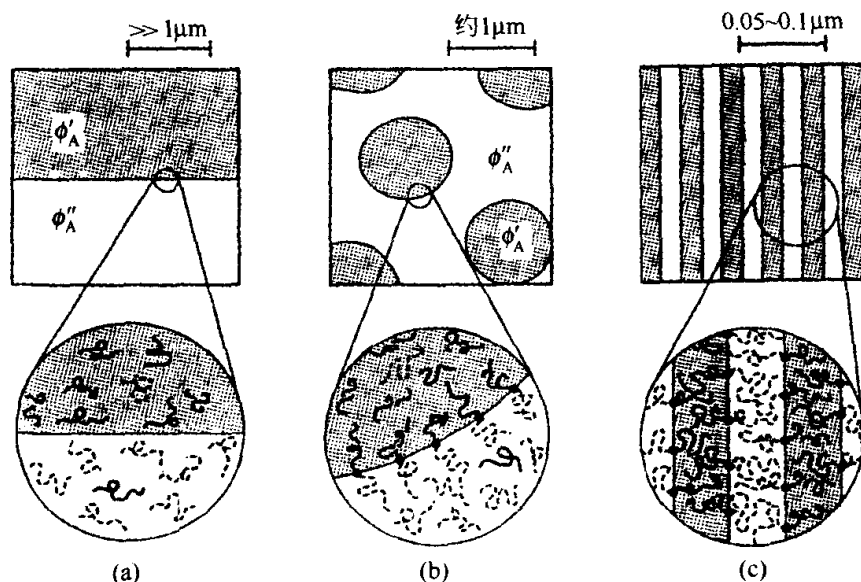


图 21-14 聚合物相行为的宏观、介观和微观相有序

- (a) 热力学不相溶的两种聚合物混合系统产生的宏观相分离；(b) 不相溶的两种聚合物加入两嵌段共聚物充作活化剂稳定了介观相有序；(c) 两嵌段共聚物由于共价键束缚产生微观如条纹相结构

(引自 Bates F S. Science. 1991, 251: 898)

制结构，在 A 和 B 体积分数相等情况下形成如图 21-14 (c) 所示。如果多分量混合物既包括 A 和 B 两种不同的聚合物，也有少量由 A 和 B 聚合物构成的两嵌段共聚物，于是共聚物作为一种表面活性剂分布在 A 和 B 聚合物界面，以减少两相界面张力，从而抑制了宏观相分离，形成介观尺寸的微乳液结构，如图 21-14 (b) 所示。

相分离热力学是研究相分离形成不同相共存的平衡热力学。下面我们以前述混合物为例讨论由于相互作用导致相分离形成不同尺寸下的畴结构。

(1) 宏观相分离 混合物是否相溶取决于混合过程中的自由能变化是否小于零，对于高分子聚合物混合，由于高分子聚合物的分子量很大，混合时熵的变化很小，而且混合过程焓变化常常大于零，因此，混合过程中自由能通常大于零。这表明大多数高分子混合物是不相溶的，高分子混合物往往出现宏观相分离。近 60 年前，Flory 和 Huggins 独立给出了高分子聚合物共混体系的自由能表达式，反映了这类系统中混合焓和混合熵的竞争。在无规混合近似下，二元混合系统的自由能由相互作用能和混合熵组成：

$$\frac{F}{k_B T} = \frac{\phi}{N_A} \lg \phi + \frac{(1-\phi)}{N_B} \lg(1-\phi) + \chi \phi(1-\phi) \quad (21-7)$$

这里 N_A 和 N_B 分别表示 A 类聚合物和 B 类聚合物链段数， ϕ 表示 A 类聚合物的体积分数。右边前两项表示混合熵，由于混合导致系统更无序。长链比短链的混合熵更小，即混合熵随着 N 增加而减少。右边第三项表示混合焓，可以大于零或小于零，取决于 AB 链相互作用强度的符号（参见 1.3.3 节）。

$$\chi \propto \frac{1}{k_B T} \left[\epsilon_{AB} - \frac{1}{2}(\epsilon_{AA} + \epsilon_{BB}) \right] \quad (21-8)$$

$\chi > 0$ 有助于系统发生相分离。对于聚合物， χ 与温度 T 有关，通常 χ 被假定为 $\chi = \alpha T^{-1} + \beta$ 。式 (21-7) 是一个忽略了空间组分涨落的平均场理论。当 $N = 1$ 约化为通常描述小分子混合物的热力学理论。随着外界参数如压强或温度的变化，如果自由能在整个浓度范围内从单势阱变成双势阱形式，系统会发生相分离，最后形成的两共存相可根据自由能最小确定，即由 A 和 B 两共存相对应的双势阱的公切线决定 A 和 B 两相平衡。如前面 18.3.4 节所述，我们可获得两相共存线：

$$\left. \frac{\partial F}{\partial \phi} \right|_{\phi_A} = \left. \frac{\partial F}{\partial \phi} \right|_{\phi_B} \quad (21-9)$$

如果我们用化学势定义 $\mu = \frac{\partial F}{\partial \phi}$ ，式 (21-9) 表示两相共存时的化学势相等 ($\mu_A = \mu_B$)。随着温度改变，两相的组分 ϕ_A 和 ϕ_B 也相应变化，于是不同温度下两相共存对应的组分值构成的曲线即为两相共存线。混合体系从均匀相淬火到两相共存线内，会发生相分离，如图 18-42 所示。注意高分子聚合物

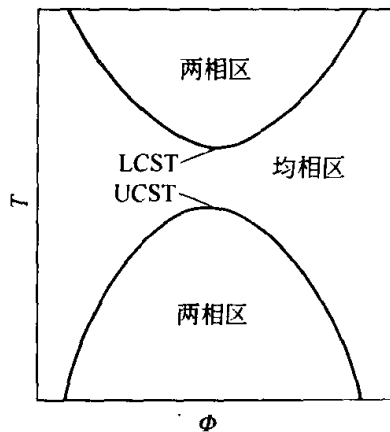


图 21-15 二元混合物相分离相图
由于低混合熵，高分子聚合物常常出现低临界溶液温度 (LCST) 相分离，即随着温度升高，相分离发生。而小分子混合物往往出现高临界溶液温度 (UCST) 相分离
当自由能二阶导数等于零，即拐点

长链组合的混合体系由于低混合熵，常常发生低临界溶液温度 (LCST) 相分离，即随着温度增加有可能导致相分离；如果链很短，有可能发生高临界溶液温度 (UCST) 相分离，就像小分子系统随着温度淬冷，相分离发生。图 21-15 上面曲线和下面曲线分别给出了低临界溶液温度 (LCST) 相分离和上临界溶液温度 (UCST) 相分离。

与 18.3.4 节有关小分子系统相分离分析类似，在聚合物中我们也发现在两相共存线内自由能的二阶导数会分别出现大于零和小于零，对应于混合相自由能为亚稳和不稳定。

$$\frac{\partial^2 F}{\partial \phi^2} = 0 \quad (21-10)$$

形成一系列曲线，构成了区分混合物热力学不稳定和亚稳态的失稳 (spinodal) 线。如图 18-42 虚线所示，在失稳线内，混合物不稳定，将失稳分解，出现自发相分离；而在两相共存线和失稳线间，为亚稳区。最后，临界点 (为 LCST 相分离的最低温度和 UCST 相分离的最高温度处) 可根据自由能的三阶导数等于零得到：

$$\frac{\partial^3 F}{\partial \phi^3} = 0 \quad (21-11)$$

组合式 (21-10) 和式 (21-11)，我们得到二元聚合物临界组分

$$\phi_c = \frac{N_A^{1/2}}{N_A^{1/2} + N_B^{1/2}} \quad (21-12)$$

和相应的相互作用强度

$$\chi_c = \frac{(N_A^{1/2} + N_B^{1/2})^2}{2N_A N_B} \quad (21-13)$$

对于对称聚合物混合 ($N_A = N_B$)， $(\chi N)_c = 2$ 和 $\phi_c = 0.5$ 。 χN 在聚合物相分离控制中是一个十分重要的物理量，反映了聚合物系统焓 ($\propto \chi$) 和熵 ($\propto 1/N$) 的竞争。

(2) 介观相分离 混合物界面或膜不仅在界面内可以伸缩，而且可以在

垂直界面方向发生形变。界面张力的降低常常会抑制相分离的生长。介观相如微乳状液由于活化剂的加入, 表面张力大大减小, 这时不在平面内形变的弯曲模式导致的弹性能量不可忽略。德国物理学家 W. Helfrich 最早引入弯曲弹性自由能用来考虑泡和膜模型,

$$F_c = \frac{1}{2} K \int ds \left(\frac{1}{R_1} + \frac{1}{R_2} - \frac{2}{R_0} \right)^2 + \bar{K} \int ds \frac{1}{R_1} \frac{1}{R_2} \quad (21-14)$$

这里平均曲率 $H = (1/R_1 + 1/R_2)/2$, 自发曲率 $C_0 = 1/R_0$ (反映双亲分子的极头堆积与碳氢尾巴体积的竞争), 高斯曲率 $(1/R_1) \times (1/R_2)$, 式 (21-14) 右边第二项与拓扑结构变化有关, 通常不考虑。上式表示使自由能最小的平均曲率 H 应为自发曲率 C_0 , 偏离自发曲率 C_0 的形变所需花费的能量由弯曲模量 K 决定。通过组合 Flory-Huggins 自由能和弯曲弹性能, 使用与宏观相分离相平衡同样的分析, 可获得介观尺度下的微乳状液形成 (介观相分离), 得出介观相分离的二相共存线和失稳线。

(3) 微观相分离 如果混合系统是 A 和 B 子链构成的嵌段共聚物 $A_f B_{1-f}$, 这时 A 和 B 链结合形成单体, 不能发生宏观相分离, 式 (21-9) 不能描述最后的相共存状态。不管怎样, 构成共聚物的 A 、 B 单体排斥相互作用强度 χ 试图保持 A 、 B 子链分离。在高温下, 由于熵效应起主导作用, 系统形成一个类似于液体的无序结构, 但在低温下由于 A 、 B 子链排斥作用变成主导地位, 导致一系列有序相结构, 随着组分浓度 f 改变, 结构会发生一系列转变如层状、圆柱和球状相等, 取决于相互作用强度 χ 和共聚物结构 (如链段数 N 和 A 、 B 组分浓度 f), 在给定温度下可获得一系列有序相相图。我们首先考虑 A 和 B 链段浓度相等 ($f = 1/2$) 的情况, 当 $\chi > 0$, 嵌段共聚物中 A 和 B 链段接触的减少将使系统内能减少, 但这种 A 、 B 链段分离过程只能局域出现 [如图 21-14 (c)], 而因此又引起链段拉伸, 导致系统位形熵的损失。嵌段共聚物的位形熵 S 也正比于 $1/N$, 因此 χN 描述了微相分离中内能和熵效应的竞争, 它是控制微相分离的重要参数。当 $\chi N \ll 10$, 熵 S 贡献起支配作用, 嵌段共聚物以均匀态无规地分布。增加 χ 或 N , 逐步让内能和熵贡献平衡, 当 $\chi N \approx 10$ 时, 内能和熵贡献达到平衡, 进一步增加参数 χN 导致系统出现一级相变到有序态, 这时熵偏爱但能量耗费的无序态转变成有序的周期层状结构。当 $\chi N \gg 10$, 能量贡献起支配作用, 嵌段共聚物形成高度有序的条纹相, 如图 21-16。在这种强分离极限下 ($\chi N \gg 10$), 有序结构的形成具有非常重要的实用意义。

如果 A 、 B 链段的浓度体积分数不相等, 结构将会发生相应的变化。随着 f 的改变, 有序微结构的形状和堆积对称性受到影响, 减少或者增加 f 使得共聚物两端堆积和链延伸的约束不对称, 导致不同的 f 下出现相应的

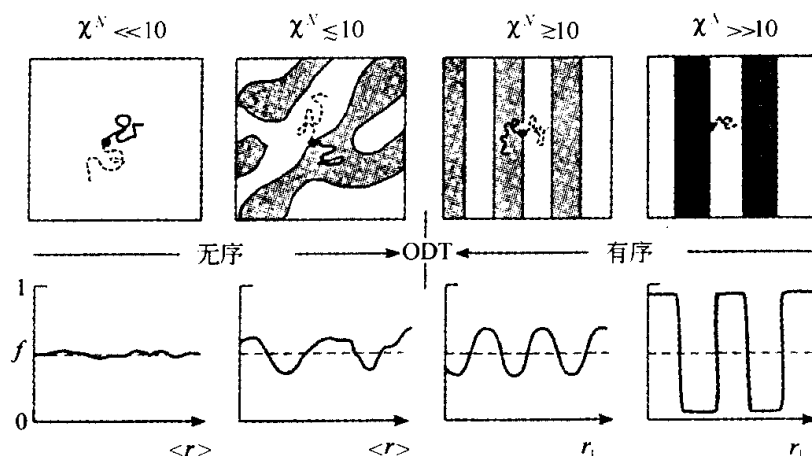


图 21-16 对称浓度下的嵌段共聚物 (PS-PI) 随着反映能量和熵竞争的参数 χN 改变而结构发生的变化

在 $\chi N \approx 10$, 系统熵 (约 $1/N$) 或能量 (约 χ) 发生小的变化将分别导致有序状态 ($\chi N \geq 10$) 和无序状态 ($\chi N < 10$)。随着体系从无序到有序转变 (ODT), 相应的 A 链段浓度从均匀改变成周期调制

(引自 Bates F.S. Science. 1991, 251: 898)

新有序相对称性, 参见图 3-32 所示。对于聚苯乙烯 (PS) 和聚异戊二烯

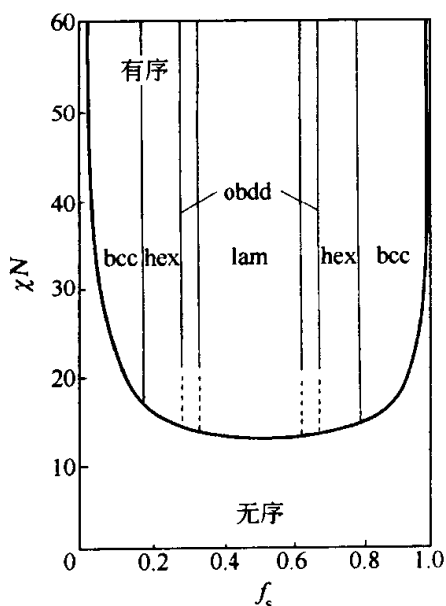


图 21-17 改变 (PS-PI) 嵌段共聚物 A 和 B 链段的相对浓度引起结构的变化相图

f_s 是聚苯乙烯的体积分数。图左边 bcc, hex, lam 分别表示聚苯乙烯以体心立方、六角堆积和层状结构有序地分布在聚异戊二烯的主体相中。图右边 bcc, hex, lam 分别表示聚异戊二烯以体心立方、六角堆积和层状结构有序地分布在聚苯乙烯的主体相中。而 obdd 表示聚苯乙烯 (PS) 和聚异戊二烯 (PI) 构成的双连续金刚石互联有序结构

(PI) 构成的嵌段共聚物, 当聚苯乙烯的体积分数 $f_s < 0.17$, 聚苯乙烯形成的微球 (PS spheres) 以体心立方结构有序地分布在聚异戊二烯的主体相中。增加成分 f 到 $0.17 < f_s < 0.28$, 产生六角堆积的圆柱状 (PS cylinders) 微结构。当 $0.28 < f_s < 0.34$ 出现聚苯乙烯 (PS) 和聚异戊二烯 (PI) 构成的双连续金刚石 (PS OBDD) 互联有序结构。在 $0.34 < f_s < 0.62$ 之间系统导致层状微结构 (PS, PI lamellae) 的出现, 进一步增加 f_s 出现逆有序结构, 如图 21-17 给出了改变 χN 和 f 导致 PS-PI 共聚物的相图。

为了讨论强分离极限下 ($\chi N \gg 10$) 微观条纹相分离形成的物理机制, 我们从竞争相互作用机制出发作进一步探讨。实际上, 软物质如

双亲分子或共聚物构成的系统常常出现微观尺度（几十个埃）下规则的畴形状如二维情形下的条纹和泡，三维情形下的片状结构和六角排列的球滴（微观相分离）。大量物理化学系统在平衡时也会出现类似却规则的畴状，似乎与微观相互作用的细节无关。这种周期性的结构特征显示建议用一种可能的普适机制来考虑。Seul 和 Andelman 采用唯象的 Ginzburg-Landau 理论对微观相分离形成作了细致的分析。至今人们通常认为短程的吸引与长程排斥的竞争会导致局域的微观相分离，例如：磁系统、铁电流体由于长程库仑和极性相互作用；复杂流体或软物质如双亲分子或共聚物的存在；传统的二元混合物由于化学反应存在等等。

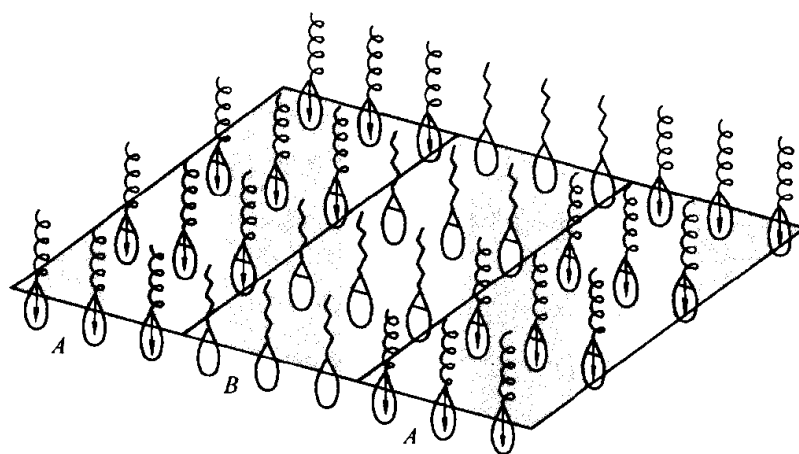


图 21-18 双亲分子（A 和 B）在油-水界面的示意图

（引自 Seul M. Andelman D. Science. 1995, 267: 476）

我们考虑由两种双亲分子所组成的二维系统（如图 21-18），其中一类双亲分子的体积分数为 ϕ ，极矩为 μ ，周期相的显示用调制结构序参数 $\phi(\mathbf{r})$ 来表示。系统的自由能 F 由短程吸引 F_ϕ 和长程极化排斥 F_d 两部分贡献组成

$$F = F_\phi + F_d \quad (21-15)$$

这里

$$F_\phi = F_0(\phi) + \frac{1}{2} \Lambda \int d^2 \mathbf{r} |\nabla \phi|^2 = F_0 + \sigma l \quad (21-16)$$

式 (21-16) 右边第一项为体贡献，为 ϕ 的 Ginzburg-Landau 级数展开形式，第二项是界面能（界面线张力 σ 乘界面长度 l ）。长程极化排斥部分 F_d 为

$$\begin{aligned} F_d &= -\frac{\mu^2}{2} \iint d^2 \mathbf{r} d^2 \mathbf{r}' \phi(\mathbf{r}) g(\mathbf{r}, \mathbf{r}') \phi(\mathbf{r}') \\ &= -\frac{\mu^2}{2} \int d^2 \mathbf{q} \phi_{\mathbf{q}} G(\mathbf{q}) \phi_{-\mathbf{q}} \end{aligned} \quad (21-17)$$

这里 $g(\mathbf{r}, \mathbf{r}') = |\mathbf{r} - \mathbf{r}'|^{-3}$ 表示长程极性相互作用。 ϕ_q 为 $\phi(\mathbf{r})$ 的 Fourier 变换, 即

$$\phi(\mathbf{r}) = \phi_0 + \sum_q \phi_q \cos(\mathbf{q} \cdot \mathbf{r}) \quad (21-18)$$

式 (21-15) 自由能 F 的 Fourier 变换为

$$F_q = \frac{1}{2} \int d^2 q [-\mu^2 G(q) + \Lambda q^2] \phi_q^2 \quad (21-19)$$

这里 $G(q)$ 为 $g(\mathbf{r}, \mathbf{r}') = g(|\mathbf{r} - \mathbf{r}'|)$ 的 Fourier 变换。稳定状态要求 $\frac{\partial F_q}{\partial q} = 0$:

$$q = \frac{\mu^2}{2\Lambda} G'(q) \neq 0 \quad (21-20)$$

这表示式 (21-18) 只保持一个调制频率, 从而出现一个新的特征长度。因此宏观相分离得到了抑制, 导致微观尺度下 A 和 B 周期调制的条纹相形成, $\phi_s = \phi_0 + \phi_q \cos qx$ ($q \neq 0$)。序参数的这种空间周期调制是由于非局域极性相互作用和局域线张力的竞争结果。此外, 导致宏观相分离抑制的长程相互作用也可以来自于其他因素, 详情参阅评论文章 M. Seul 和 D. Andelman Science. 1995, 267: 476。

21.3.2 软物质复相分离的动力学

(1) 相分离动力学的两种机制 相分离动力学一般分成两类, 如图 21-19 (类似现象也出现在固态相分离中, 参见 18.3 节)。

① 成核生长 (nucleation and growth)。一级相变冷却到亚稳区域, 少数相在多数相溶液中出现小滴。从均匀相到成核相, 最初小滴生长通过超饱和

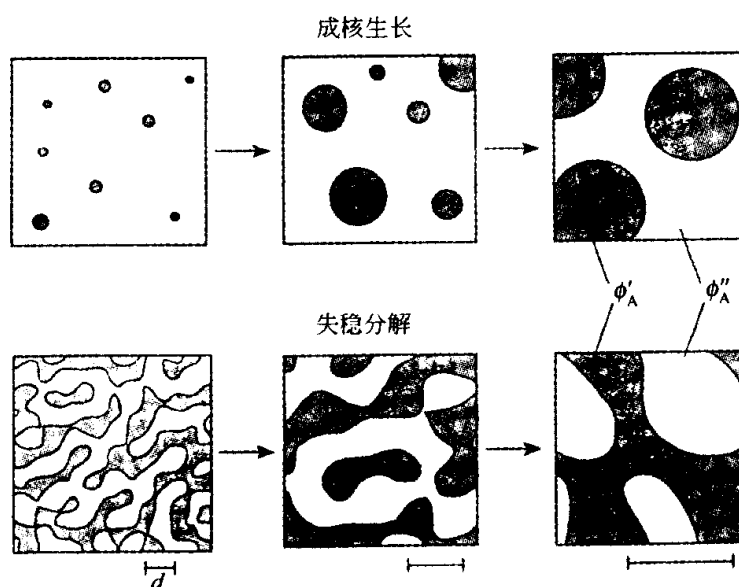


图 21-19 成核生长和失稳分解

(引自 Bates F.S. Science 1991, 251: 898)

溶液中自发出现。畴尺寸的进一步增加通过小滴弥合或者粗化,即通过小滴蒸发引起大滴生长。由于聚合物的低扩散和高粘性,生长的第二阶段即球状核生长融合成大核是一个极其缓慢的热激活过程。

② 失稳分解 (spinodal decomposition)。快速冷却 (淬火) 到不稳定区域,长波涨落不稳定性导致系统会自发相分离,形成无序的双连续两相交织而成的结构,如图 21-19。由于 $\partial^2 f / \partial \phi^2 < 0$, 无能量势垒,涨落大,混合物自发相分离,这类相分离手段具有非常重要的应用价值。下面我们介绍失稳相分离动力学的一般理论。

(2) 失稳分解动力学理论 我们假定样品 i 的浓度和质量流分别为 $\phi_i(\mathbf{r}, t)$ 和 $j_i(\mathbf{r}, t)$, 质量守恒的连续性方程要求

$$\frac{\partial \phi_i(\mathbf{r}, t)}{\partial t} = -\nabla \cdot \mathbf{j}_i(\mathbf{r}, t) \quad (21-21)$$

对不可压缩的二元 A/B 混合物 $\phi_A + \phi_B = 1$ 。只存在一个独立的运动方程,因而可消去注标 i 。而 $\mathbf{j}(\mathbf{r}, t) = \phi(\mathbf{r}, t) \mathbf{V}(\mathbf{r}, t)$ 是由于化学势 μ 的不均匀引起,即

$$\mathbf{j}(\mathbf{r}, t) = -M(t) \nabla \mu(\mathbf{r}, t) \quad (21-22)$$

这里 $\mu(\mathbf{r}, t) = \frac{\delta F(\mathbf{r}, t)}{\delta \phi(\mathbf{r}, t)}$ 实际上是系统相分离的驱动力。

对于混合系统,自由能由体贡献 $f[\phi(\mathbf{r})]$ 和界面张力的贡献两部分组成,即

$$F\{\phi(\mathbf{r})\} = \int_v d\mathbf{r} [f[\phi(\mathbf{r})] + \frac{1}{2} \Lambda |\nabla \phi(\mathbf{r})|^2] \quad (21-23)$$

此式称为 Cahn-Hilliard 自由能,式 (21-23) 右边第二项与混合物界面有关,描述产生两不同的均匀畴 (A 和 B 相) 界面需要的能量, $\Lambda > 0$ 。自由能密度 $f[\phi(\mathbf{r})]$ 是局域的均匀混合的自由能,在二相区域具有双阱结构。组合式 (21-21) ~ 式 (21-23), 我们得到与时间有关的 Cahn-Hilliard-Cook 非线性扩散方程

$$\frac{\partial \phi(\mathbf{r}, t)}{\partial t} = M \nabla^2 \frac{\delta F\{\phi(\mathbf{r}, t)\}}{\delta \phi(\mathbf{r}, t)} + \eta(\mathbf{r}, t) \quad (21-24)$$

这里 $\eta(\mathbf{r}, t)$ 是热噪声项 (Cook 项)。 F 通常可写成 Ginzburg-Landau (GL) 自由能形式,

$$F = \int d\mathbf{r} \left[A\phi^2 + B\phi^4 + \frac{1}{2} \Lambda |\nabla \phi|^2 \right] \quad (21-25)$$

故式 (21-24) 也称为具有守恒序参量的含时 Ginzburg-Landau 方程。自由能形式 (21-25) 中惟一的相分离驱动力来自界面张力 (同类分子吸引, 不同类分子排斥), 因而导致宏观相分离发生, 最后演化成宏观尺度下的两个共

存相。如果计及长程排斥互作用, 宏观相分离将会被抑制, 这时自由能变成

$$F'\{\phi(\mathbf{r}, t)\} = F\{\phi(\mathbf{r}, t)\} + \frac{c}{2M} \iint \phi(\mathbf{r}, t) G(\mathbf{r}, \mathbf{r}') \phi(\mathbf{r}', t) d\mathbf{r} d\mathbf{r}' \quad (21-26)$$

其中 F 是标准的 Cahn-Hilliard 自由能函数 (见公式 21-23), 右边第二项是长程排斥项, $G(\mathbf{r}, \mathbf{r}')$ 是 Laplace 方程的格林函数 $\nabla^2 G(\mathbf{r}, \mathbf{r}') = -\delta(\mathbf{r} - \mathbf{r}')$, c 是长程排斥互作用的强度。于是我们获得修正的 Cahn-Hilliard 扩散方程

$$\frac{\partial \phi(\mathbf{r}, t)}{\partial t} = M \nabla^2 \frac{\delta F'\{\phi(\mathbf{r}, t)\}}{\delta \phi} \quad (21-27)$$

$$\text{即} \quad \frac{\partial \phi(\mathbf{r}, t)}{\partial t} = M \nabla^2 \frac{\delta F\{\phi(\mathbf{r}, t)\}}{\delta \phi(\mathbf{r}, t)} - c\phi(\mathbf{r}, t) \quad (21-28)$$

式 (21-28) 右边第二项的存在将抑制宏观相分离, 在某一时刻后出现畴不再生长, 即钉扎现象。

注意方程 (21-28) 没有考虑动力学相互作用, 因此最后得到的微观图案是由于不同力程下短程吸引和长程排斥作用竞争的结果, 原则上由能量最小决定; 在动力学相互作用存在下的非平衡系统中, 微观图案并不是由能量最小决定。相互作用能量常常决定着相分离的趋势, 即起始阶段。随着时间的演化, 动力学相互作用开始起作用, 试图释放两种对抗相互作用产生的矛盾 (失措), 即抑制体系失措引起的拓扑缺陷, 辅助体系迅速进入有序结构。

21.4 结语: 自组装的基本问题 and 研究前沿

软物质是近年来兴起的凝聚态物理与化学 (如超分子化学包括合成化学、聚合物和胶体科学)、材料科学 (如自组装材料制备) 交叉的学科新生长点。该方向应用性强、本身富有学术挑战, 对新材料的制备、新结构的实现以及自组装纳米材料控制和设计具有十分重要的意义, 也为材料工业可持续发展提供一定的物理基础。软物质物理研究已形成它特有的特色: ①研究对象主要来自化学和生命中的有机物质; ②通过化学自组装, 合成和相分离等研究手段实现高度有序的新功能和新结构材料; ③侧面已从凝聚态物理传统的相互作用能量机制变成动力学机制和熵效应起重要作用, 其贡献对软物质有序结构的形成带来新的机遇; ④研究范围集中在探索亚微米和纳米尺度下有序微结构的物理特性; ⑤由于自组装带来的材料可设计和可控性, 重在人工微结构新材料和新工艺制备技术的探索。当前, 自组装已成为所谓从小到大合成纳米和亚微米材料的基石, 特别是生物和各种有用的日常物品制备的必须手段。这使得软物质自组装/自组织研究成为日益重要的国际研究热点。在物理学领域和化学领域, 软物质自组织和超分子化学已分别成为学科研究的新生长点。

软物质的研究面临新的挑战包括：研究对象越趋复杂；学科从单一走向多学科交叉；理论和实验的结合更为迫切（新材料的实现和预测需要物理学工作者来探索其物理机制）；大量的问题有待解决，特别是软物质体系中动力学效应和熵效应如何产生各种尺度下的复杂有序结构？如何来影响生长动力学等等。为此，揭示各种有序微结构产生的物理机制和动力学引起的新生长规律，从而来控制 and 设计新材料、新器件和新工艺具有非常重要的科学意义和实际应用价值。

当前，软物质物理和超分子化学中许多问题缺少清楚的物理图像和机制。软物质自组织中存在一些非常重要的问题：软物质的相行为和受约条件下的纳米结构制备大多依赖于经验调节，有必要建立完整的物理机制；在从小到大的自组装过程中以及复相分离中动力学效应和熵效应很少考虑，而这两个因素对软物质自组织的后期演化形貌和生长动力学起着至关重要的作用，特别是人们通常认为动力学流动、熵效应和涨落这些因素是不利于系统出现有序结构的，而恰恰在软物质中我们往往观察到由于动力学效应、熵效应和外界涨落的引入，软物质系统在纳米和亚微米尺度下观察到高度有序的结构形成；由于构成软物质的超分子结构更为复杂，所形成的图像不仅具有空间的位置序，而且具有取向序和其他对称性，因此需要把现有理论推广到多分量序参数情形。

软物质研究的对象非常广泛，并且十分复杂，是一项物理学，化学、生命科学和材料科学交叉的学科领域（见图 21-20）。因此，面临科学工作者的是一项既有机遇又富有挑战的课题，其相应的基本问题 and 研究前沿也显得丰富多彩。其中，自组织现象特别丰富，自组装技术特别重要，熵效应和动力学效应特别明显。目前，国际上有关软物质自组装/自组织前沿问题研究的主要研究进展及其科学意义包括以下几方面。

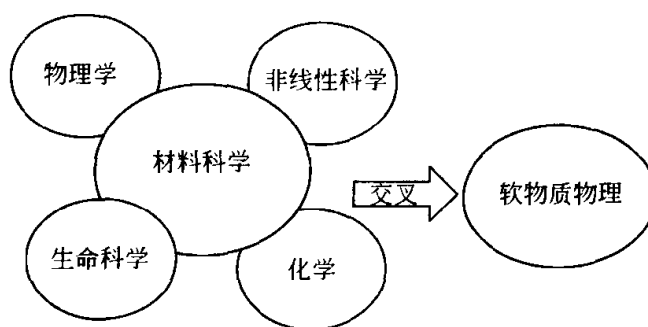


图 21-20 软物质涉及的多学科交叉性

① 从简单的有序结构到复杂的结构（如海绵状或有序的复杂液晶相）相变过程的统计物理研究。超分子自组装和分子聚集形成各种有序程度不同

结构（如胶体、微乳液、液晶等）的过程中，随着分子浓度的增加有序结构发生变化的相变机理仍是今后软物质研究的基本问题。

② 悬浮胶粒、微乳液或分子聚集体之间相互作用导致进一步的有序化过程，生物大分子的折叠以及晶化过程是今后软物质研究的主要方向。目前的主要进展如包括：微乳液、胶粒等分散在各向异性介质（如向列液晶）聚集成链；耗散介质中外力驱动下结构有序化研究；改变溶液组分，在临界涨落附近将大大提高蛋白质结晶成核速度；超分子组成的高分子结构由于空间螺旋性将会自组装形成更大的螺旋结构等等。

③ 软物质形态变化熵作用的理解。胶体等颗粒在耗散介质如聚合物溶液和液晶溶液中聚集过程熵作用的影响。

④ 各种不稳定性导致软物质形貌变化是软物质物理在生物技术和化学工业中具有重要应用背景的课题。包括高分子熔融和超分子溶液等复杂液体的相分离和相行为控制等。相互作用竞争（失措）反映不稳定模式的存在，这类系统的临界涨落如普适律还有待实验的进一步研究。

⑤ 衬底和表面模板诱发软物质形貌变化是今后工艺、器件（如光子晶体）设计和新材料、新结构发现的重要途径。主要进展如高分子聚合物和超分子混合溶液因衬底花样导致的微观相分离结构和采用溶胶-凝胶手段形成巨多孔晶体等。

由于软物质本身的特点，软物质形貌的可控制性和新材料的可设计性将使材料科学工作者不断地去研究、开发和实现新材料，同时需要物理工作者不断地去探索新材料形成过程中形貌的变化和新有序结构产生的内在机制。如何控制和设计软物质形貌的变化在材料制备和生产中具有重要的应用前景。因此，研究软物质自组装/自组织过程中复杂相互作用的竞争、熵力和外界能源的驱动对自组织有序结构的形成具有重要的学术和实际意义，将有助于我们理解软物质在不同尺度下的形貌（图案）生成，结构稳定性和非平衡动力学演化和生长速度控制等机理。

最后，我们必须指出，软物质的自组织不仅对软物质材料制备十分有用，对于硬物质（如半导体、磁性材料）结构制备也非常有用。利用软物质的自组装/自组织制备硬物质纳米和微米结构成为近年来物理、化学和材料科学合作攻关的重要研究方向，特别是金属和半导体纳米材料如纳米颗粒的制备。目的是应用它们作为光学和电学特性的材料。量子尺寸效应是这些孤立纳米结构显示新物理现象的一个代表性例子。这种自组织纳米颗粒也许会带来更令人兴奋的实用价值，原因是由于不同纳米颗粒间的光、磁和电相互作用也许会产生意想不到的物理特性。我们知道，传统的物质微加工试图从体材料产生微器件（从大到小）的技术在制备纳米尺寸的结构并不是很有

效,即便是采用最现代的平版印刷技术。而从分子单体自组装形成分子聚集的纳米结构(从小到大的技术)克服了这一困难。物质的自组装/自组织广泛存在于自然界,是构成生命和日常生活的基本前提。了解物质的自组装/自组织,我们相信将来必然成为产生许多新的未预见的合成材料(无论是软物质还是硬物质)提供基本的思路和策略。

参 考 文 献

- 1 Gelbart W. M. Ben-Shaul A. Roux D. Micells. Membranes, Microemulsions, and Monolayers. Springer-Verlag, 1994
- 2 Israelachvili J. N. Intermolecular and Surface Force. New York: Academic Press, 1992
- 3 Lubensky T. C. Solid state Commun. 1997, 102: 187
- 4 Sarfan S. A. Statistical Thermodynamics of Surfaces, Interfaces and Membranes. Addison-Wesley Publishing Company, MA. 1994
- 5 Goldacker T., Abetz V., Stadler R., Erukhimovich T., Leibler L. Nature. 1999, 398: 137
- 6 Shinbrot T. Muzzio F. J. Nature. 2001, 410: 251
- 7 Chen Z. R. et al., Science. 1997, 277: 1248
- 8 Seul M. Andelman D. Science. 1995, 267: 476
- 9 冯端,冯步云.放眼晶态之外:漫谈凝聚态物质之二.长沙:湖南教育出版社,1994
- 10 Zhou H. J., Zhang Y., Ou-yang Z. C. Phys. Rev. Lett. 1999, 82: 4560
- 11 Witten T. A. Rev. Mod. Phys. 1999, 71: S367
- 12 de Gennes P. G. Scaling Concepts in Polymer Physics. Ithaca and London: Cornell University Press, 1979
- 13 Kellogg G. J. et al. Phys. Rev. Lett. 1996, 76: 2503
- 14 Frenkel D. Physica A 1999, 263: 26
- 15 Turner M. S. Phys. Rev. Lett. 1992, 69: 1788
- 16 Lekkerkerker H N W, Stroobants A. Nature. 1998, 393: 305
- 17 de Gennes P. G. Badoz J. Fragile Objects: Soft Matter, Hard Science, and the Thrill of Discovery, Copernicus, New York: Springer-verlag, 1996 (中译本:卢定伟等译.冯端校.软物质与硬科学.长沙:湖南教育出版社.)

第4篇 展 望

第22章 材料表征及其进展

撰稿人 叶恒强

现代材料科学在很大程度上依赖于对材料性能与其成分及显微组织关系的理解。因此，对材料性能的各种测试技术，对材料组织从宏观到微观不同层次的表征技术构成了材料科学与工程的一个重要部分，也是联系材料设计与制造工艺直到获得具有满意使用性能的材料之间的桥梁。

从新材料的研制发展中，可以看出检测表征技术所起的作用。例如，高温超导氧化物的发现揭开了超导材料研究的新篇章。为了拓展存在高温超导性的无机化合物的范围，对已发现高温超导氧化物的结构表征是必须的。X射线结构分析给出钇钡铜氧化物的单胞点阵常数，电子显微像与电子衍射分析给出钇钡铜氧化物的孪晶结构和铋系氧化物的调制结构，中子衍射对氧原子空位的测定给出比X射线、电子衍射更多的信息。这些技术对推进高温超导氧化物的实用化和机理研究，都起了重要作用。又如碳60的发现早在20世纪80年代中，但证实其富勒球结构却经历了数年用光谱技术揭示其键合特性。而富勒球结构的直观原子尺度显示至今仍是科学家的目标。在公认的材料科学与工程四大要素——合成与制备、组成与结构、性质、使用性能——之中，性能检测与表征技术通过组成与结构在化学、物理学和环境科学基础上与其他三大要素相互联系和反馈，同样构成材料科学与工程的关键共性技术。

22.1 性能检测

材料的性能是指材料在给定的一套条件下，当某一条件发生变化时材料所产生的响应。对材料性能的测定需要建立一种广义力与广义位移的关系，在关系为线性时，材料性能由线性常数表征，如弹性模量（联系力与位移）、热容（联系熵与温度）等。在关系偏离线性时，材料的性能需要由高阶的常数来描述，如光学晶体在强光照射下的极化率（联系光场电振幅与介质极化强度）。材料的传输性能与材料内部参数的梯度所产生的流有关，由广义力

与所产生的流之间的比例常数来表征，如热导率、电导率、扩散系数等。

新材料的发展还提出一些难以用单项性能进行评价的材料特征。如金属材料的可焊性，材料的可加工性，高温熔盐气氛下的抗腐蚀性能等。这些特性涉及到的因素比较多，具体材料在使用场合提出的要求侧重也不一样。这都给材料性能检测技术提出了更高的要求。这就需要考虑一个测试体系，创造一个测试环境，以求所测试得到性能尽量真实反映使用时的状况。有一些材料性能可能更为复杂。如对使用于生物体内的材料所进行的在活性生物体对其性能影响的测试；对使用于航空航天器材料的振动、环境性能测试等。

无论是材料的单项性能，还是综合性能，其测试技术都随着材料的发展而相应革新。但总体看来，材料性能检测技术还是比较稳定，相应标准也较规范。有鉴于此，再考虑到材料性能的多样与复杂，这里不拟对材料检测技术作进一步的介绍。

22.2 显微组织分析原理

材料的组织形态、结构与成分可以用材料与电磁波辐射、电子、离子、中子的相互作用来进行检验，这些相互作用的信息构成有关分析技术的物理基础。利用光波在材料表面反射所得信息的光学显微术，也属于电磁波辐射范围。这部分技术已十分成熟，本文不作介绍。尽管分析技术发展很快，但其原理仍大多在这些相互作用信息的范畴之内，易于理解 and 选择。如基于原理的深化发展出来的广延 X 射线吸收边精细结构、电子能量损失谱的吸收边精细结构分析技术；基于技术进步在原子探针基础上发展出来的原子定位的三维原子探针等。也有些分析技术是创新性的，如利用近程隧道效应电流和作用力发展出来的近场扫描显微术，为材料表面表征及纳米结构制备开拓了崭新的领域。

下面将材料显微分析技术按辐照源及探针与材料相互作用信息分类作简要介绍。

22.2.1 电子与样品相互作用

如果材料被一个细的、约束较好的聚焦电子束照射，并且入射束有足够的束流以产生显微分析所需的信号，如图 22-1 所示，电子与样品相互作用产生的信息可分别叙述如下。

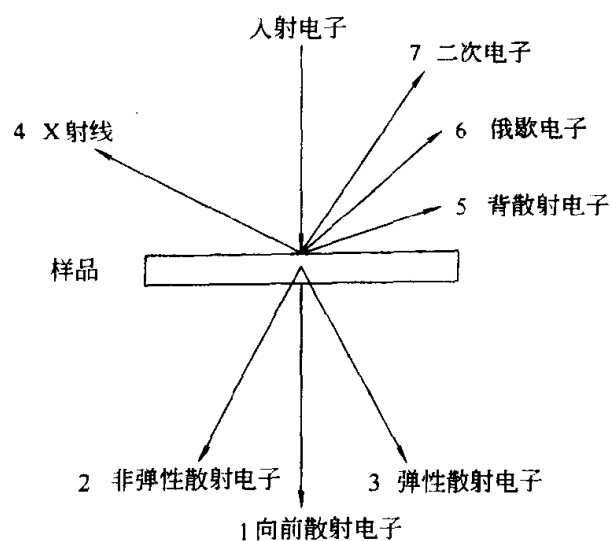


图 22-1 电子与材料相互作用原理示意图

(1) 向前散射电子 即透射电子, 这是电子穿透样品的部分。这些电子携带着被样品吸收、衍射的信息, 用于透射电镜的明场像和透射扫描电镜的扫描图像。以揭示样品内部微观结构的形貌特征。单一地透射电子成像的分辨率在亚微米水平。

(2) 非弹性散射电子 这些电子在穿过样品时损失了部分能量, 方向也有微小变化。用于电子能量损失谱, 提供成分和化学信息。也能用于特殊成像或衍射模式。

(3) 弹性散射电子 在晶体材料中, 电子按布拉格定律, 被具有不同取向及面间距的周期排列原子平面散射到不同角度, 产生电子衍射图, 提供晶体结构、对称性、取向和样品厚度等信息。

(4) X 射线 入射电子在样品原子激发内层电子后外层电子跃迁至内层时发出的光子。不同原子被激发的 X 射线是特征的, 其谱线波长及强度可用于电子照射区域的化学成分的定性及定量分析。X 射线产生的效率很低, 一万个人射电子才产生一个光子。最常用的探测器是固体能量色散 X 射线谱仪。一般情况下, 能探测 Na 以上元素。在特殊装置中, 可探测 B、C、N、O 等轻元素。其绝对检测量、准确度和空间分辨率与元素、样品、入射电子能量多种因素有关。

(5) 背散射电子 入射电子在样品中经散射后再从上表面射出来的电子。扫描电子显微术、透射扫描电子显微术收集这些电子成像可反映样品表面不同取向、不同平均原子量的区域差别。

(6) 俄歇电子 俄歇电子和 X 射线一样由原子内壳层电子被入射电子电离而产生。不同的是跃迁电子结合能变化时产生的多余能量不是通过发射 X 射线而是发射俄歇电子。俄歇电子的能量与入射电子无关, 仅与产生它的原子和电子壳层有关, 从而可表征原子及电子结构。这个信息仅来自样品表面几个原子层, 因此扫描俄歇电子谱成为表征材料表面化学成分, 特别是元素偏聚的有力工具。

(7) 二次电子 由样品中原子外壳层释放出来, 在扫描电子显微术中主要用它成像, 反映样品上表面的形貌特征。

(8) 荷集电流 高能入射电子进入半导体, 激励价带电子进入导带, 在价带中留下空穴 (具有正电荷电子性质), 产生电子-空穴对。在正常条件下, 两者在很短时间内就能复合 (小于 $1\mu\text{s}$)。如果在样品上附加电场, 电子和空穴在外电场作用下分别向相反方向迁移, 产生了局部导电性, 称电子束感生电导。回路中产生电流称荷集电流。半导体的缺陷会改变载流子的局部复合率。将荷集电流放大成像提供了直接观察半导体中单个晶体缺陷的方法。

(9) 阴极发光 上述电子-空穴对复合时释放的能量, 部分以可见光或红外线从试样辐射出来, 称阴极发光。可收集成像, 空间分辨率约 $1\mu\text{m}$ 。

(10) 热波与声波 电子束照射到样品上时, 试样局部产生周期性的表面热。这是热波源。一部分热波能量总是转换成同频率的声波。声波传播较远, 且易于检测。基于此原理发展的热波显微术, 包括光声、电声、离子声显微术等可检测晶界、薄膜与涂层界面、电子器件的表面与亚表面、半导体掺杂浓度分布等。

要说明的是, 上述携带样品信息的射线, 虽然在样品中有多种会同时产生, 但由于仪器设备的功能所限, 并不是所有的讯号都能被收集。各种分析仪器设备各有其专长和特点, 其适用的对象和场合也有所区别。这是材料工作者在选择表征技术时要注意的。

22.2.2 X射线与物质的相互作用

X射线是另一类探测源。X射线与物质的相互作用包括下述效应。图22-2给出简单的示意图。

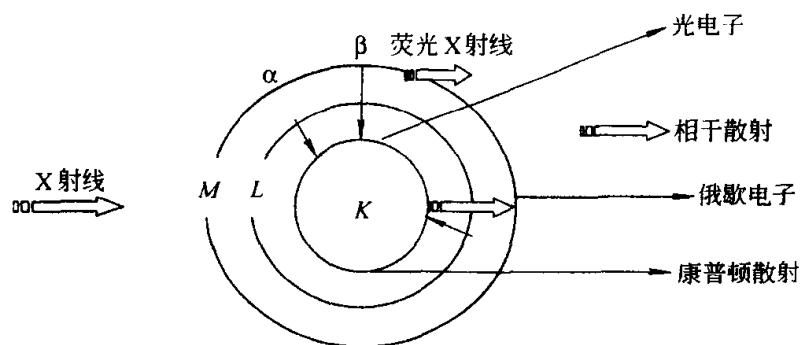


图 22-2 X射线与物质相互作用示意图

(1) X射线分析所提供的信息

① 光电子谱。光子的所有能量消耗于结合电子的打出, 可用于测定原子和分子轨道的结合能。

② 荧光。被激发最内层 K 或 L 壳层电子的离化原子极不稳定, 从能量较高轨道来的电子将级联式填充空位, 并发射二次荧光光子。X射线荧光光谱术可用于快速定性分析材料中所含的元素, 考虑影响荧光强度因素后还可作定量分析。

③ 俄歇电子。当 K 壳层电子吸引入射 X 射线离开原子, 一个外壳层电子就会落入 K 壳层填充空位, 引起发射一个 K_α 或 K_β X 射线光子。这个光子在离开原子之前, 又被价电子吸收并飞离原子。第二次离开原子的电子即为俄歇电子。俄歇电子的动能与使 K 电子离化的 X 射线光子能量无关。可产生 K 空位的 X 射线都能引发具有固定动能的俄歇电子。利用俄歇电子不

仅可作样品表面原子元素种类及组成比例的化学分析,还可测量化学键能及其间的微细能量差。

④ 相干散射 X 射线。这是引起 X 射线衍射现象的弹性散射。X 射线被不同取向及面间距的周期排列的原子平面按布拉格定律散射到不同方向,产生 X 射线衍射谱。它构成 X 射线衍射物相分析和结构分析的基础。

⑤ 吸收。材料对 X 射线的吸收取决于组成材料原子类型及材料密度。这在 X 射线强度分析时必须考虑。在 X 射线吸收谱上,曲线的陡然不连续处为吸收边,表明此时的吸收机制不同于两侧的吸收机制。吸收边会给出有关原子结构、原子环境等诸多信息。吸收边的精细结构可准确给出原子中电子能量的信息,而电子能量又随原子周围键环境的不同而稍有变化。这样通过一些运算有可能反推出原子配位和键环境。

(2) 同步辐射 X 射线分析 X 射线分析的一个特殊分支是同步辐射 X 射线分析。当以相对论能量运动的带电粒子被强制沿一曲线轨迹运动时,会发出电磁辐射。其辐射强度正比于粒子的能量,反比于其质量,因此电子和正电子是最有效的发射粒子。所有环形粒子加速器中都产生这类辐射,称同步辐射。

同步 X 射线辐射具有许多优点,在材料表征中有特殊的应用。

① 使 X 射线吸收光谱法广泛可用。X 射线吸收光谱法能得到诸如非晶、多元材料、合金等复杂体系中选定原子种类的局部结构的信息,由于材料研制趋向多元、复合化而日显重要。同步辐射源发出的 X 射线通量比一般的源大 3~4 个数量级,获取吸收边谱数据的时间从几周缩小到几分钟。同步辐射的准直性好,使能量分辨高达 10^{-4} 。另外,连续同步辐射谱宽,使大多数元素的 L 系和 K 系吸收边都可记录到。

② 就结构分析而言,对于大多数晶体结构,同步辐射源优点不突出。但对于团簇系统,点阵参数精确测定、原位表征固态化学转变等,能量分散 X 射线衍射和劳厄衍射都可在很宽的谱范围提供极大量信息。

③ 真实表面和界面的衍射和散射。同步辐射的高能量和强度,使之用于各实用环境中掩埋在致密外表层下的界面结构进行非破坏性的检测。这种技术正越来越多用于表征衬底和异质外延层之间的点阵共格性。

④ X 射线形貌术利用了衍射强度对晶体点阵的局部位移及取向变化非常敏感这一特点,能够将晶体缺陷直接成像或成动力衍衬像。用同步辐射源提供的白光辐射和窄带单色形貌术在谱范围宽、高光子通量、束发散窄等方面比其他实验室辐射源提供的形貌术有无可比拟的优越性。

22.2.3 离子束分析

当带有能量的离子撞击样品表面时,入射离子与价电子或内壳层电子相

互作用，并能激发原子或使原子离化。在非常接近靶核时，就会发生库仑相互作用或核相互作用。利用这些原理可得到关于样品成分、结构和作用过程的一些信息。表 22-1 列出了应用这些作用的可能性及对应检测表征技术。

表 22-1 离子束分析的作用原理及技术

作用对象	过 程	入射离子	试样原子	应 用
电子	弹性散射 非弹性散射	能量损失 能量损失	激发 离化 光发射 X 射线发射 二次电子	光谱分析 质子 X 射线荧光分析 俄歇电子分析
原子核	轰击 沿晶向 库仑弹性 库仑非弹性 核弹性 核非弹性 反应	散射 小角散射 散射 能量损失 能量损失 散射 能量损失 散射 吸收	被溅射 反冲 库仑激发 反冲 激发 γ 射线 嬗变	离子探针质谱分析 离子沟道效应分析 卢瑟福背散射谱分析 弹性反冲探测 卢瑟福背散射谱分析 共振核反应分析 核反应分析

离子束分析是 20 世纪 50~60 年代核物理鼎盛发展时代的相应产物。经过近半个世纪的发展应用，离子束分析技术的主体已很成熟。

22.2.4 中子衍射与小角度中子散射

中子主要被原子核散射，绝大多数原子散射中子的能力在 2~3 倍之内。而且其散射可以被认为是点散射，其散射因子与角度基本无关。这与 X 射线的散射因子角分布大不相同。这决定了中子衍射适合于研究轻原子在试样中的分布。另一方面，中子的磁矩和不成对的电子自旋的相互作用也造成散射，与自旋在样品中如何排列有关。所以中子衍射可提供磁结构的信息。

前述结构分析用的 X 射线、电子、中子衍射都是探测具有周期结构在纳米以下尺寸的点阵单胞、调制结构、超晶格、长周期结构等。对于那些密度变化起伏在亚微米尺度的组织结构，就需要用到小角度散射所提供的信息。散射几何定律告诉我们，引起入射波矢 K_0 向 K_g 方向散射的样品中密度调制波的波矢 g 起点在入射波矢的端点，终点在散射波的端点。记波矢 K_0 及 K_g 的模为 $2\pi/\lambda$ ，则 $g = K_0 - K_g$ 。若散射角为 2θ ， g 的模长为 $2\pi\theta/\lambda$ 。即密度调制波波长是波矢的倒数。对于波长为纳米以下尺度的 X 射线，其小角散射探测的调制结构尺寸在数十纳米。中子波长为纳米量级，其小角散射所能探测的调制结构尺度可达微米。就表征亚微米尺度的材料不均匀性

而言,电子显微术可提供直观的图像。但小角度散射可提供平均结构信息,在特定场合是很有用的。

22.3 显微结构表征

显微结构表征包括观察组织的形貌、确定其原子排列方式和分析化学组分。分析方法可按观察形貌的显微镜,测定结构的衍射仪及分析成分的各种谱仪进行分类。

22.3.1 形貌观察

光学显微术和金相学是在微米尺度观察材料组织结构的最常用的方法。

扫描式和透射式电子显微镜把观察显微组织的尺度推进到纳米层次。扫描电镜已成为材料断口形貌分析的主要工具,由于分辨率改进到纳米尺寸,近年来已经作为质量监控设备安装在大规模集成电路器件的生产线上。透射电镜加上X射线能谱仪,电子能量损失谱仪和高亮度、高相干性的场发射枪电子源,使它成为能在纳米尺度给出材料组织形貌、结构和成分的综合仪器,功能强大。

场离子显微镜利用针尖端表面原子层轮廓边缘的电场不同,借助惰性气体离子轰击荧光屏可以得到针尖正面原子排布的投影像,达到0.2nm的分辨率。可以直观显示晶界或位错露头处原子排列及气体原子在表面的吸附行为。

20世纪80年代中期发展起来的扫描隧道显微镜,借助一根针尖与表面之间的隧道效应电流的调控,在表面作X,Y方向扫描的同时,在保持隧道效应电流恒定的电路控制下,针尖将依表面原子的起伏而在Z方向游动。这种移动经电信号放大并由计算机进行图像处理,可以得到表面原子分布的图像。其纵、横向分辨率分别达到0.01nm及0.2nm,为材料表面表征及纳米结构制备技术开拓了崭新的领域。在此技术有关的利用近程作用力而发展出来的原子力显微镜也日益扩展其应用。此外基于扫描探针近场技术发展的显微镜还有:研究磁性样品表面磁畴和磁场分布的磁力显微镜;观察微电路上电特性的静电力显微镜;近场光学显微镜和光学扫描隧道显微镜;测量样品表面微区温度变化的扫描热显微镜;可用于探测金属与半导体次表面电子性质的弹道电子发射显微镜等。

22.3.2 结构测定

材料的结构测定仍以X射线衍射为主。这一技术包括粉末照相相分析,高温、常温、低温衍射仪,背反射谱和透射劳厄照相,测定单晶结构的四联衍射化,织构的极图测定等。由于X射线在晶体中衍射基本上是运动学衍射,因此衍射束的振幅与晶体单胞的结构因子有线性关系。这样从X射线

衍射图的几何分布可确定晶体材料的点阵类型和单胞大小，从衍射强度可反推出晶体单胞内各原子的位置。在计算机的帮助下，只要提供试样的尺寸及完整性满足一定要求，X 射线单晶衍射仪就可以打印出测定晶体样品的晶体结构详尽资料。但 X 射线不能在电磁场作用下会聚，所以通过狭缝或准直的方法使 X 射线束变小同时也就削弱了强度。要分析微米大小的单晶材料需要更强的 X 射线源，例如同步辐射 X 射线源。

由于电子与物质相互作用比 X 射线强 4 个数量级，而且电子束又可以会聚得很小，所以电子衍射特别适用于测定微细晶体或材料的亚微米尺度结构。在透射电镜中进行电子衍射试验，除了有结合形貌观察以鉴别微细物相的有利条件外，还可以作结构的直接观察，或称高分辨像。将晶体置于电子透镜的前焦面，通常透射电镜样品就是放在这个位置，则电子经晶体的衍射图就在电子透镜的后焦面上形成，而且进一步在像平面上呈现晶体的点阵像。现代电子透镜的分辨率已达 0.2nm 甚至更高，完全可以在有利的取向将晶体的投影原子柱之间的距离清楚分开。因此只要将晶体试样制备得足够薄，使电子穿过晶体是一种运动学相互作用，这样在 X 射线中通常需要从衍射强度反推原子排列的过程可以自动由电子透镜完成，将晶体沿电子束投影方向的原子排列成像于像平面。下面以新型高温结构陶瓷 Ti_3SiC_2 的高分辨结构像为例作一说明。碳化物陶瓷熔点、硬度都很高，碳化钛、碳化硅是两种典型的高温结构材料和耐磨材料。但它们的可加工性很差，要将它们成型为合用的部件要较长的工艺过程。它们的复相陶瓷对韧性有所改善。近年发现碳硅钛三元陶瓷有室温可切削性，而且其电子结构也很有特点，有望用于功能性材料。 Ti_3SiC_2 的结构可描述为三类原子分别以平面六角密堆排布，面与面之间按密堆积结构的三个球形成的凹陷堆垛。由于一个六角密堆面上的凹陷有两种，如果将形成凹陷的原子平面记为 A 的话，则两种凹陷位置分别记为 B, C。图 22-3 (a)、(b) 的结构模型图给出了三类原子的分布，而且标注了各原子平面在密堆垛中的记号。从图可以看出，Ti-C 之间是 ABCABC 堆垛，即原来 TiC 面心立方结构。Ti-Si 之间则是 ACA 堆垛，相当于原来 SiC 的六角密堆垛，不过这里的 C 原子由 Ti 原子置换了。由于沿 $[110]$ 方向投影看过去，Ti-C 原子间距仅为 0.16nm ，在 0.2nm 分辨率的电镜中不可分辨，加上碳原子的散射能力弱，故碳原子在图 22-3 (c) 的高分辨像中不能单独成像。在高分辨像中钛、硅原子均可单独成像，在图中与白亮点对应。如以 Si 原子面为 A 位置，则 Ti 原子在 Si 原子面上下两侧的排布分别为 BCA，呈镜面对称，在图中已用线条勾勒出来。高分辨原子像对研究晶体结构而言不仅有直观性的优点，而且对于局域的结构变异或缺陷的揭示，更显示它的长处。

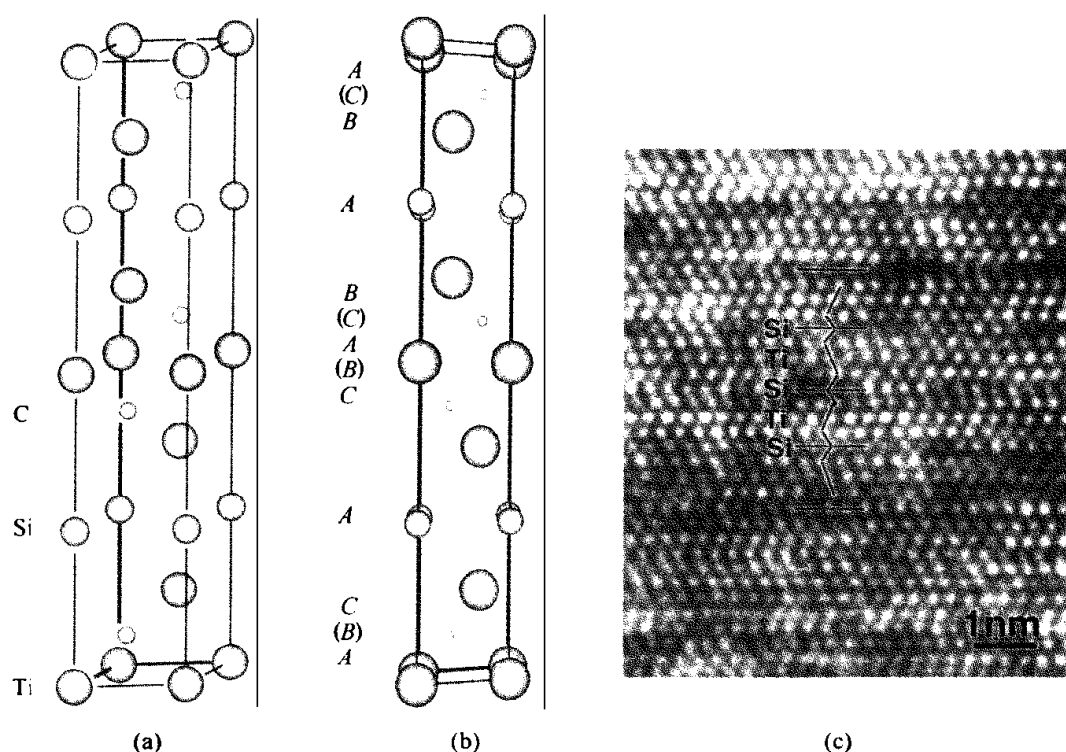


图 22-3 Ti-Si-C 三元系化合物晶体结构的高分辨像

(a) Ti_3SiC_2 单胞原子排列；(b) 沿 $[110]$ 取向的结构投影图，旁边

A、B、C 标注密堆积平面的位置；(c) $[110]$ 取向高分辨像，

元素符号标注原子种类，折线表示原子面堆积顺序

中子衍射也是结构分析的手段之一。中子受物质中原子核散射，轻重原子核对中子的散射能力差别不大。中子衍射有利于测定轻原子的位置，是上述两种主要结构分析方法的有力补充。

22.3.3 化学组分分析

材料的化学成分分析除传统的分析化学技术外，还包括质谱、紫外、可见光、红外光谱分析，气、液相色谱，光发射与吸收谱，X 射线荧光分析谱，俄歇与 X 射线光电子谱，二次离子质谱，拉曼谱，电子探针，原子探针（与场离子显微镜联用），激光探针等等。适用于各种不同情况的分析方法，在这里不能一一叙述。

另有一类谱仪是基于材料受激发的发射谱，专为研究晶体缺陷附近的原子排列状态而设计的，如核磁共振谱仪、电子自旋共振谱仪、穆斯堡尔谱仪、正电子湮没等等。

扫描探针显微镜的主要功能虽然是形貌观察，但控制探针的隧道电流、近程力等，除与探针跟试样之间的距离有关外，还与试样表面的化学元素、电子态、自旋态等有关。所以这类仪器从原理上说，也有探测表面化学状态

的潜力。有些扫描探针显微镜也已开发出这种功能。

就各种方法的适用层次而言，一方面要知道需探测组织的尺度，另一方面需要知道各种分析方法自身具备的能力。图 22-4 是几种典型的显微组织表

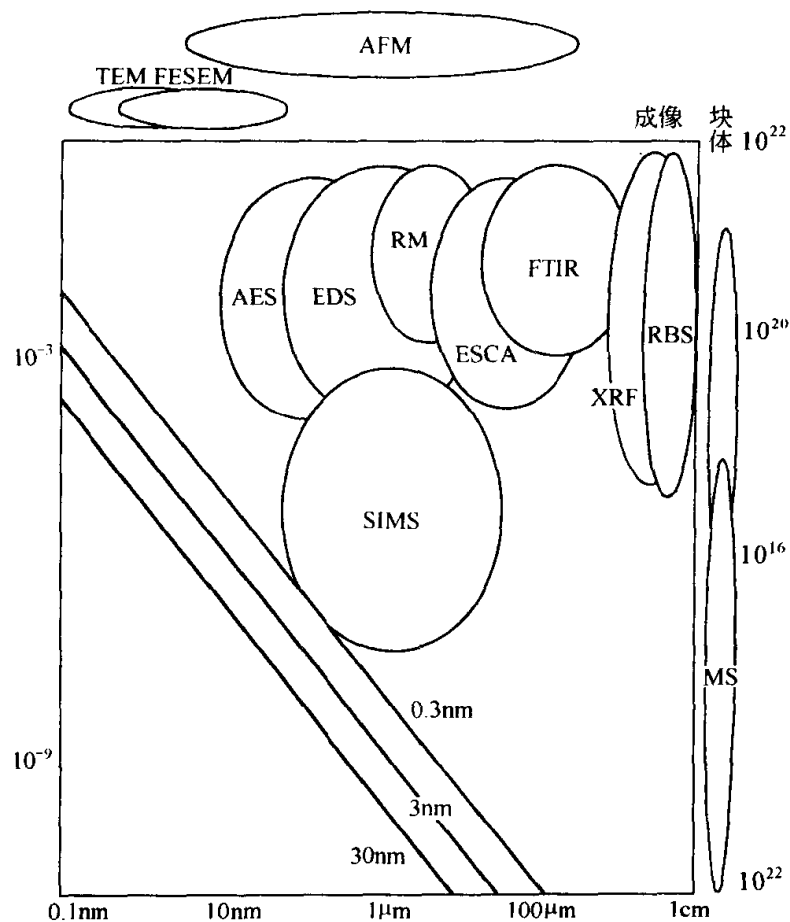


图 22-4 各种分析手段的分辨率与灵敏度

纵坐标为检测范围(左)/每立方厘米原子数(右)，横坐标是分析束斑大小。

图上方表示成像模式的分辨率，右方表示对块体而言。图内三根斜线分别

表示对于 0.3nm, 3nm, 30nm 的样品深度进行成分分析的物理极限

征设备分析空间分辨极限和灵敏度的综合示意图，图中各分析技术英文缩写含意如下：

- RM——拉曼谱；
- TEM——透射电子显微术；
- FESEM——场发射扫描电子显微术；
- AFM——原子力显微术；
- AES——俄歇电子能谱；原子发射光谱；
- EDS——X 射线微区分析能谱；
- FTIR——傅里叶变换红外光谱；

ESCA——化学分析电子能谱；

RBS——卢瑟福背散射；

XRF——X 射线荧光光谱；

SIMS——二次离子质谱；

MS——质谱。

22.4 从材料出发的综合分析

为了说明从材料研制的需要及从测试技术提供的可能两方面的结合，以纳米碳管和氮化镓棒这类一维材料为例，介绍所涉及的多种性能测试技术及显微分析手段。同时也说明材料向低维方向延伸对表征技术的促进。按表征的要求分类，这些测试技术有以下几种。

① 形貌观察。一维管、棒生长得完整与否，杂质多少，缺陷分布，在电子显微镜均可清楚显示。但有些特殊形貌，如纳米碳管的手性特征，则尚未能直接成像。

② 结构测定。碳对 X 射线散射能力很弱，所以纳米碳管多用电子衍射分析。电子衍射对碳管的手性特征分析有其独到之处。

③ 化学成分。当碳管缀饰入元素或化合物，需要在非常特定的区域作鉴定时，场发射枪透射电子显微镜所提供的纳米尺度电子束适合于此类分析。

④ 光学性质。氮化镓是正在研制中的发光材料，当其二维尺寸受限后的发光特性是大家感兴趣的课题。

⑤ 电学性质。纳米碳管的场发射特性近年来颇受关注，它作为量子线也提到研究日程。纳米碳管依其螺旋度及管径，可以表现出金属性或半导体性的导电特性。所以相关的电测量也日益精细化。

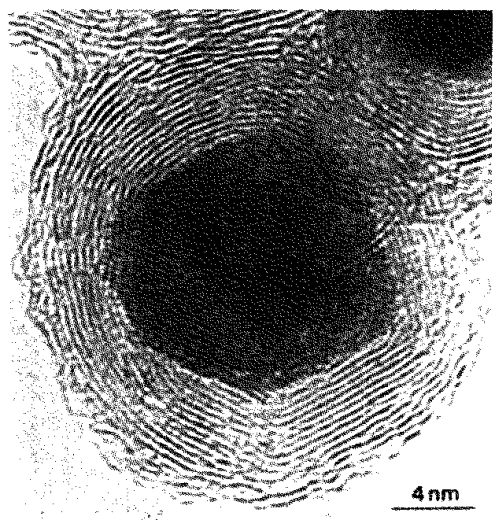


图 22-5 氧化镓纳米粒子包覆着氮化硼膜的高分辨像

⑥ 力学性能。石墨纤维的高模量和高强度已经很吸引人，而纳米碳管的力学性能更为诱人。已经在电子显微镜观察的范畴内提出通过一维管或棒受力弯曲的程度来测量其基本力学参数的电子秤概念。

⑦ 化学性能。如稳定性，特别是在特殊环境中的稳定性，如太空。还有如吸放氢特性，也很吸引人。

图 22-5 是一种纳米复合粒子的高分辨像。中心的二维点阵像给出这是

ZrO_2 ，相应的微衍射图也可以给出同样的信息。包覆在粒子外层的膜有明显的层状结构，但因为只有一个参数，难以作判断。图 22-6 (a) 给出从粒子中心得出的电子能量损失谱，可以确定它是氧化锆。在 200~450eV 区可以察觉到 B 和 N 的存在。由于粒子是近似球形的，入射电子在进入及离开氧化锆时都必然令穿过表面覆盖层，所以表层为 BN 的可能性很大。将电子束只集中在覆盖层采集电子能量损失谱，如图 22-6 (b) 所示，可以清楚看到 BN 的两个元素峰。透射电镜将电子束会聚到纳米大小，在试样的纳米区域提供结构、形貌、组分的综合信息，这个例子是很好的说明。

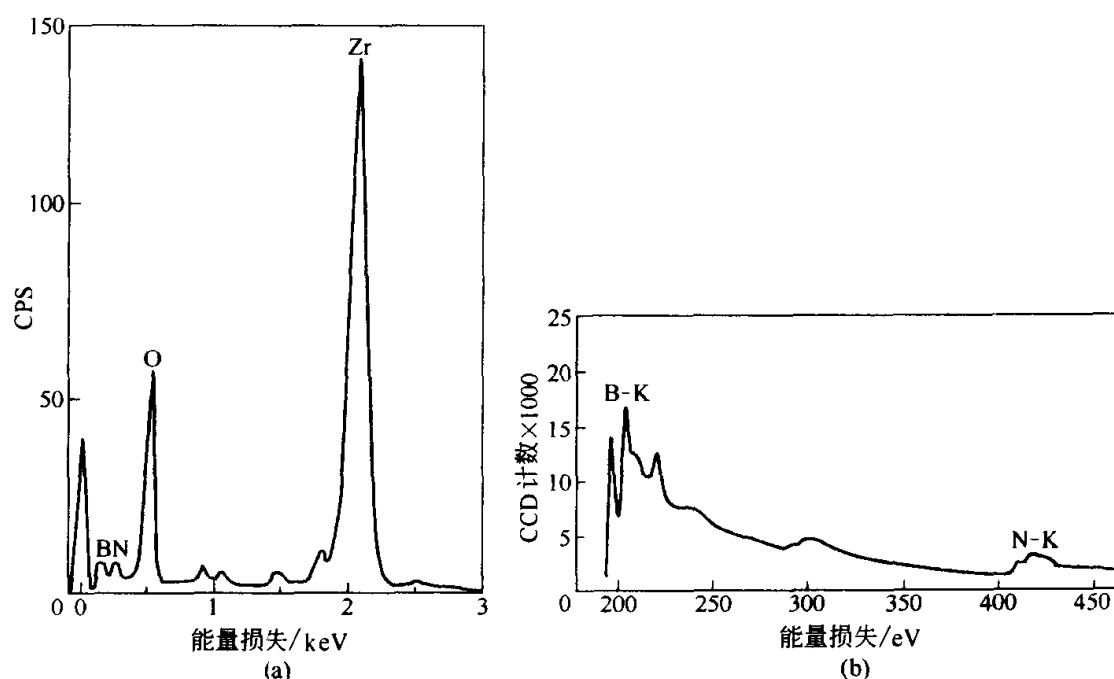


图 22-6 氧化锆-氮化硼复合粒子的电子能量损失谱

(a) 电子束辐照整个复合粒子所得的电子能量损失谱；

(b) 电子束只辐照表面膜的电子能量损失谱

22.5 材料性能与组织表征技术发展展望

22.5.1 分析技术的交叉与综合

分析表征技术各有专长。技术发展的第一趋向总是将技术的特长指标步步推向极致。如分析成分的技术把定量化做得更精确，把在材料内单一成分的检测极限一再降低等。对于新材料的研制，如能在一个设备中收集到多种信息，能从多方面反映材料的性能及组织特征，是材料工作者所希望的。因此分析技术的发展有下述两种趋势。

一是发展综合分析技术。以透射电子显微镜为例，电子束经过样品后，由电子透镜在后焦面上形成衍射图，提供样品物相晶体结构与缺陷的信息；

透镜还进一步在像平面显现样品的放大像，在有利情况下，甚或可以显示原子排列的直观像。通过收集相关特征 X 射线和电子能量损失谱，可以得到试样的化学组分信息。20 世纪末期装备有场发射枪电子源的透射电子显微镜已逐渐成为商品透射电镜的主流，这种电子源能提供束斑尺寸在纳米量级、亮度又足够进行成分分析的入射电子束，使样品在纳米范围内进行形貌、结构、成分的综合分析成为可能，有力地促进了材料界面、纳米结构、微电子器件的表征研究。

二是性能测试与显微组织表征同步进行。为验证金属多列原子作为量子线的电导性能，可以将金的针尖与平面试样接触后再慢慢拉开，从透射电镜中可以清楚看到随着拉开的过程，联接针尖与平面的原子列从数列一列一列地减少到单列直至断开。这样从多列到单列的原子线模型材料是实现了，但还要同时测量金多列原子的电导性质。单壁纳米碳管也有作为量子线的潜力。为了考察这一点，已有人设计出这样的实验，将定向生长在某基体的长度参差不齐的单壁纳米碳管倒置过来，逐渐浸入汞液中，在电子显微镜可看到最长的一根管与汞液相接触。这时接连基体与汞液的电路测量到一个相当于量子电导的电流。当这根最长的管继续浸入汞液时，电流保持不变，意味着碳管在汞液的浸入长度对电流没有影响。仅当第二根次长的碳管与汞液接触时，电路电流跳跃到另一个平台。这能确定了单壁纳米碳管作为电导量子线的性质。

22.5.2 分析结果的定量化与可视化

计算机的快速发展在多方面促进材料检测与评价分析技术。

首先是数据处理能力的提高使分析定量化。自然所有分析数据的处理都能从计算能力的提高中得到好处。例如一条曲线，如果需要的话，可以画出其微分或积分曲线；在冲击或硬度试验中，过去只能简单记录结果的冲击值或硬度值，现在则可从摆锤或顶头接触试样的一刻开始记录受力与位移曲线，从中分析出其他力学量，如模量、屈服强度、加工硬化率等等。分析定量化还可以从透射电子显微镜的高分辨像定量分析作为例子进行说明。首先是像点几何位置的定量化。在有利条件下，高分辨像的像点排列可被认为是样品中投影方向原子柱的排列。对于完整晶体，人们从原子像中看原子排列的周期性、对称性就可以了。然而对于局部缺陷区域，如界面、层错、位错、空位等等，如果想了解原子排列位置的畸变，仅凭肉眼只能有定性的印象。现在可以首先将每个像点求出他们的“重心”，作为其几何位置。进一步将完整区的像点位置在几何上平均为一个网格，然后将这个网格与局域缺陷区的取样网格对照，将缺陷区的像点位移一一测量出来。这就是缺陷区的位移场。如果有足够的力学参数，还可从进一步推导缺陷区的应力场等等。

其次是像点强度的定量化。高分辨像每个像点强度与真实投影原子柱的原子序之间的关系目前还没有严格的定量关系。但在一些特定情况下，可以作一些类比。例如先将像点的强度定量化，分成若干个强度等级。然后设计模型，根据可能的原子排布进行像模拟。从模拟图中的像点强度也进行等级分类。实验图与计算图强度等级如能从基体（或绝大多数）原子中找出一个可比较的认定值，即作为参照物。进而把其他少量原子的像点强度区分出来。这种能对像中像点的原子类别作辨认的定量分析法还在发展中。

其次是分析结果的可视化。在人们与客观世界交流信息的手段中，视觉具有接受信息量大、直观性强的特点。如果分析结果只是一堆数字，虽然很精确，但失之枯燥。现在计算机的介入将这些数据作处理，变成图表、曲线甚至生动的图像。扫描隧道显微镜的表面拓扑像的原始数据是扫描针尖在 X 、 Y 方向扫描过程中沿 Z 方向上下位移的记录。初步的可视化是一根根沿 X 或 Y 方向的起伏线，再把这族线沿 Y 或 X 方向叠起来。进一步可以将它们处理成表面拓扑像，如我们常看到的 Si 表面重构图，把一个个原子形象地用圆球表示出来。如果按时间顺序把各幅画面作动画播放，就能把表面重构、化学反应的过程清楚地显示出来。

22.5.3 低维材料制备与分析测试结合

在传统意义上，材料的分析测试都是在制备与加工后，截取部分试样进行检查。由于纳米科技日趋成熟，许多新材料研制是在微米-纳米尺度进行的。这与材料显微组织表征的尺度正好相近。这就产生了双重效应：一是在制备过程中进行检测表征，另一是利用显微组织表征设备作为纳米结构的制备手段。下面以透射电子显微镜和扫描隧道显微镜为例作一说明。

透射电镜兼有显微组织表征中的形貌、结构和成分分析的综合功能。为了制备与检查功能材料的需要，可以在透射电子显微镜的镜筒一侧，并行放置离子注入设备、聚焦离子束设备和分子束外延生长设备。这些样品制备工具可以与镜子分立，通过真空管道作样品传输，也可以与镜筒接通后将离子束引入镜筒，直接对正受电镜观察的试样进行加工。另一方面，电子束在半导体器件光刻技术中本身就是加工源。在场发射枪源及电子聚光镜的帮助下可将电子束会聚到纳米大小，远远小于光刻技术的束斑尺寸，可以制备量子点和量子线的模型材料。这方面的结合已经得到一些有趣的结果。例如有人无定形 SiO_2 的薄膜上，用纳米束斑的电子束照射，使相应大小的区域的二氧化硅分解出晶体颗粒的硅。电子束有规律的移动并停住辐照，可在无定形的二氧化硅基体中出现规则排列的硅量子点。硅量子点的晶体特征可以从高分辨像一览无遗，对应的电子能量损失谱中的化学位移更进一步说明硅晶体与二氧化硅中硅原子的峰位的差别。这是在电镜中制备量子点的事例之

一。利用超高真空透射电镜的样品室蒸发金属使其气体冷凝后在透射电镜支持膜上淀积着金属的微细颗粒，只有纳米大小。这也非常方便透射电镜的原位表征。对于面心立方结构的金属，这种纳米颗粒常以五重(111)孪晶的形式出现。一般来说，由于每两个(111)面之间的夹角为 70.5° ，五重孪晶转圈围起来留余下 7° 多的缝隙。这是在宏观尺度的面心立方矿物常常看到的形态。但在纳米尺度，小晶体似乎不喜欢这条缝隙，出现五重孪晶完全闭合的五次对称构形。这时其中一块孪晶畴的(111)面夹角变为 78° 。在纳米硅颗粒的情况，还出现五块孪晶畴的(111)面间夹角都均匀地变为 72° 的情况。据此，一些人推断在那些变了形的孪晶畴中实际上已成为亚稳的体心四方结构。并依据类似于钢中奥氏体与马氏体的对应取向关系对这种亚稳变体的形成机制作出了解释。在透射电镜中配合聚集离子束在GaAs、Si、SiO₂薄膜中注入Ga，可以用Ga布成有规则的图案，供研究有关的物理特性。

扫描隧道显微镜除了是分辨率极高的表面分析仪器外，还成为操纵原子，在原子尺度加工纳米器件的有力工具。就其极致而言，搬动并放置单个原子无疑很吸引人，并且在一定条件下也被演示过。但更具现实意义的是用扫描探针制备量子点、量子线，并随之表征它。这是一段时期以来较活跃的领域。

参 考 文 献

- 1 Characterization of materials, Part I and II, Vol. ed. Eric Lifshin, Weinheim; New York; Basel; Cambridge; VCH. 1992. (中译本, 材料的特征检测, 北京: 科学出版社, 1998)
- 2 Joy D. C., Romig A.D., Goldstein J. I., Principles of analytical electron microscopy, Eds. New York: Plenum Press, 1996
- 3 Cullity B. D, Addison-wesley, Elements of X-ray diffraction. 4th ed. Reading, M.A., Addison-Wesley. 1978
- 4 Bowen D. K, The application of synchrotron to problems in materials science, Ed. SERC Daresbury Lab. Report DL/SCI/R19. 1983
- 5 Wiesendanger R, Guntherodt. H. J, Scanning tunneling microscopy II -further application and related scanning techniques. In Springer Series in Surface Sciences. Vol. 28. Berlin: Springer; 1992
- 6 Briggs, D. Seah M. P, Practical surface analysis by Auger and X-ray photoelectron spectroscopy. 2nd ed. New York: John Wiley, 1991

第 23 章 材料设计及其进展

撰稿人 熊家炯 朱嘉麟

23.1 什么是材料设计

材料设计 (materials design), 是指通过理论与计算预报新材料的组分、结构与性能。或者说, 通过理论设计来“订做”具有特定性能的新材料。这当然说的是人们所追求的长远目标, 并非目前就能充分实现的。尽管如此, 由于凝聚态物理学、量子化学等相关基础学科的深入发展, 以及计算机能力的空前提高, 使得新材料研制过程中理论和计算的作用越来越大, 直至变得不可缺少。

1995 年美国国家科学研究委员会 (National Research Council, 缩写为 NRC) 邀请众多专家经过调查分析, 编写了《材料科学的计算与理论技术》^[1] 这一专门报告, 其中这样说: “materials by design” (设计材料) 一词正在变为现实, 它意味着在材料研制与应用过程中理论的分量不断增长, 研究者今天已处在应用理论和计算来“设计”材料的初期阶段。

1999 年美国能源部发表一篇题为“计算材料科学: 一场科学革命将成为现实”的白皮书^[2]。此文谈到: 由于计算机能力的不断提高, 材料科学正处在另一场科学革命的边缘。科学家可以利用太拉 (10^{12}) 级以上的计算机通过模拟运算来指导先进材料的发展, 并进一步阐明材料是如何形成的, 了解它们在变化的条件下产生何种作用, 以及怎样才能获得优化而达到更好的使用效能。该文作者特别强调理论计算对于发展先进材料的指导意义。他们援引 Encyclopedia of Advanced Materials (4 卷, 近 3000 页) 对“先进材料”所下的定义: 先进材料是指那样一些材料, 对这些材料的系统合成和结构控制已进行过先行的理论研究, 使得它们具有经过精细剪裁而获得的性能组合, 以满足应用的需求^①。

2000 年美国政府发布了“国家纳米技术倡议——导致下一次工业革命”的文件^[3], 简称 NNI。这是一个引起世人瞩目的计划。它首先冲击的是纳米材料学和纳米制造学。在 NNI 中, 突出地强调了理论、模型和模拟

① 英文原文为: “those where first consideration is given to the systematic synthesis and control of the structure of material in order to provide a precisely tailored set of properties for demanding applications.”

(theory、modeling and simulation, 缩写 TM&S) 对于发展纳米材料、器件和系统起着不可或缺的、决定性的作用。因为对于纳米尺度下的物质材料, 许多情况下难以直接进行实验测量, 或者由于测量方法和手段对被测对象带来了无法避免的干扰, 以致测量结果显得杂乱无章。此时, 只有借助第一原理的理论分析, 建立模型和发展模拟计算来加以阐明、区别和论证, 从而提供研究所需的数据, 指导材料、器件和制造工艺的发展。NNI 认为, 发展 TM&S 在最近 5~10 年要达到以下目标: 减少设计新材料所需的时间; 加速从新材料向纳米器件的发展; 增加新器件运行的可靠性和预见性; 以及设计和优化新的纳米技术工艺。

从上述内容看出, 关于材料设计, 或者说设计材料这类研究工作, 迄今在国际上还没有统一的提法或流行术语。日本学者 1985 年就提出了“材料设计学”一词^[4]。俄国学者把材料设计包括在“材料学”(М а т е р и а л о в е д е н и е) 中。美国学者在《90 年代材料科学与工程》^[5]报告中称这类工作作为材料“计算机分析与模型化”(computer-based analysis and modeling)。在文献[2, 3]中, 另一些美国学者又称之为“计算材料科学”, 或者统称为 TM&S, 即理论、模型和模拟。我国 1986 年开始实施“863”计划时, 对新材料领域提出了探索不同层次微观理论指导下的材料设计这一要求, 并从那时起, 在“863”新材料领域便设立了“材料微观结构设计 with 性能预测”研究专题。由此可见, 各国学者论及“材料设计”时的用语虽然有所差别, 但其基本含意则是共同而且明确的。

材料设计可按研究对象所涉及的空间尺度和时间尺度而划分为不同层次, 划分的办法也因不同的出发点而有所差别。大致而言, 可划分为 3 个层次: 微观设计层次, 空间尺度在约 1nm 量级, 是所谓原子、电子层次的设计; 连续模型层次, 典型尺度在约 1 μ m 量级, 这时材料被看成连续介质, 不考虑其中单个原子、分子的行为; 工程设计层次, 尺度对应于宏观材料, 涉及大块材料的加工过程和使用性能的设计研究。这 3 个层次的研究对象、方法和任务是不同的。本章主要介绍材料的微观设计, 稍微涉及同连续模型设计的衔接问题, 而没有涉及材料的工程设计层次。

23.2 材料设计的发展概况

23.2.1 前期研究的回顾

“材料设计”设想始于 20 世纪 50 年代。前苏联在 50 年代初期, 开展关于合金设计以及无机化合物的计算机预报等早期工作。那时前苏联卫星上天, 说明其使用的材料是先进的。前苏联学者于 1962 年在理论上提出人工半导体超晶格概念, 不过当时他们没有提出如何在技术上加以实现的建议。

后来到 1969 年, Easki (江崎) 和 Tsu (朱兆祥) 才正式从理论和实践结合上提出了通过改变组分或掺杂来获得人工半导体超晶格。20 世纪 80 年代中期日本材料界提出了用三大材料在分子原子水平上混合, 构成杂化材料的设想。1985 年日本出版了《新材料开发与材料设计学》一书, 首次提出了“材料设计学”这一专门方向, 书中介绍了早期的研究与应用情况, 并在大学材料系开设材料设计课程。1988 年由日本科学技术厅组织功能性梯度材料 (functionally gradient materials) 的研究任务, 提出将设计-合成-评估三者紧密结合起来, 按预定要求做出材料。为此已连续组织有关这个课题的国际学术讨论会。

1989 年美国由若干个专业委员会在调查分析美国 8 个工业部门 (航天、汽车、生物材料、化学、电子学、能源、金属和通信等) 对材料的需求之后, 编写出版了《90 年代的材料科学与工程》报告^[5], 对材料的计算机分析与模型化作了比较充分的论述。该报告认为, 现代理论和计算机技术的进步, 使得材料科学与工程的性质正在发生变化。计算机分析与模型化的进展, 将使材料科学从定性描述逐渐进入定量科学的阶段。

正如文献 [5] 所预见的那样, 在以往的 10 年中, 材料设计或材料的计算机分析与模型化日益受到重视, 究其原因主要有以下几点。

① 固体物理、量子化学、统计力学、计算数学等相关学科在理论概念和方法上有很大发展, 为材料微观结构设计提供了理论基础。

② 现代计算机的速度、容量和易操作性空前提高。几年前在数学计算、数据分析中还认为无法解决的问题, 现在已有可能加以解决; 而且计算机能力还将进一步发展和提高。

③ 科学测试仪器的进步, 提高了定量测量的水平, 提供了丰富的实验数据, 为理论设计提供了条件。同时反过来, 在这种情况下更需要借助计算机技术, 将理论与实验沟通起来。

④ 材料研究和制备过程的复杂性增加, 许多复杂的物理、化学过程需要用计算机进行模拟和计算, 这样可以部分地或全部地替代既耗资又费时的复杂实验过程, 节省人力物力。更有甚者, 有些实验在现实条件下是难以实施的或无法实施的, 但理论和模拟计算却可以在无实物消耗的情况下通过理论分析提供信息。

⑤ 以原子、分子为起始物进行材料合成, 并在微观尺度上控制其结构, 是现代先进材料合成技术的重要发展方向, 例如分子束外延、纳米粒子组合、胶体化学方法等。对于这类研究对象, 材料微观设计显然大有用武之地。

23.2.2 当前面临的挑战

1969 年江崎和朱兆祥提出了由两种不同半导体薄层构成超晶格、量子

阱的概念。后来借助分子束外延等制备技术,使这一设想得以实现。现在已被证明,这一概念是20世纪80年代以来凝聚态物理学和材料科学中最有价值的概念之一,它开辟了人工设计低维材料,并利用其量子特性制备出各类功能器件的先河。20年来从量子阱到量子线,量子点直到当今的纳米结构,一直是材料科学中最富有生命力的前沿领域之一。

事实表明,现代材料科学研究必须深入到微观层次。今天不论是对材料物性的了解,还是对材料性能的表征,都要求深入到分子、原子以及电子层次。许多先进材料的加工制备,也已进入所谓“原子学过程”(atomistic processes)由单个原子聚合成簇,或形成零维、一维、二维、三维材料,都属于所谓“原子级工程”(atomically engineering)。又如今天各类关键技术所使用的材料中,发生在材料表面和界面的现象,都要求从原子水平和化学键水平进行了解,才能阐明其本质。在21世纪方兴未艾的纳米科技中起关键作用的材料,是由为数很少的原子所组成,其中增加(或减少)一个原子(或电子)都将使其状态发生明显变化,例如单电子器件就是如此。这是对当今材料科学理论和计算的挑战。

可见,原子水平上的材料设计必然受到高度重视,它在现代材料科学技术的发展中将起举足轻重的作用。

前面已提到1995年美国NRC发表了专门报告,他们专门对需要从原子水平上进行材料研究的13个领域进行了调查,提出了材料科学中理论和计算工作所面临的挑战与机遇。他们认为,如果说20世纪70年代以前,还做不到对实际材料的性能作出理论预言的话,那么,今天由于理论和计算能力的进步,已经可能实现从理论上预测新材料。同时,考虑到计算机能力将由于并行化而提高 $10^2 \sim 10^3$ 倍,这对于材料科学中的理论与计算将带来新的机遇。因此,必须抓住机遇,大力加强理论上预报材料的工作。有关这个报告的重点内容将在23.6节中介绍。

23.3 材料设计的范围与层次

1985年日本三岛良绩在《新材料开发与材料设计学》^[4]一书中为材料设计的工作范围提出了一个轮廓,如图23-1所示。此图可作为一种参考,它表明从材料制备到材料性能、再到使用,都属于材料设计的工作范围。其中包括了组成、结构和特性的微观设计在内。《90年代材料科学与工程》报告^[5]中,美国学者提出现代材料科学研究应由4个要素组成:即固有性质(properties)、结构与组分、使用性能(performance)及合成与加工。这4者是相互联系的整体。报告认为材料设计在这4个要素中都起重要作用。这里把properties与performance分开,前者指材料的固有性质,后者则同材料应

用联系, 包括寿命、速度、能量效率、安全、价格等, 它们是衡量 performance 的因素。

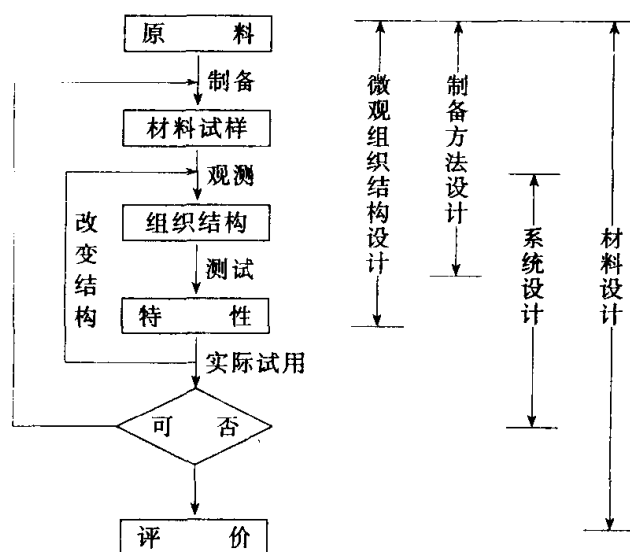


图 23-1 材料设计的范围

该报告谈到, 材料设计在材料加工制备过程中也应起重要作用。尤其是以原子、分子为起始物, 采用化学和物理学方法进行材料合成, 要求在微观尺度上控制其结构时, 离不开理论的指导。如果能配合使用实时传感器和无损检测器的情况下, 对微观结构进行优化与控制, 则可实现在理论设计指导下的智能加工。

尽管材料设计贯穿在材料从制备、测试、性能到使用的整个过程, 其核心部分仍是在物理、化学原理基础上, 对材料性能-结构关系进行理论计算与分析。当然, 除了性能-结构关系之外, 还有加工-结构、加工-性能以及加工-使用性能之间的关系。等等, 也都属于材料设计的工作范围。应根据具体的设计对象、设计目标来确定具体的工作范围及其重点。

前面已说过, 材料设计可按研究对象的尺度大致分为微观、连续模型和工程设计等 3 个层次。在《材料科学的计算与理论技术》^[1] 报告中, 作者把微观层次又分为几个范围, 并同连续模型层次连接起来。

由于在不同空间尺度范围内发生的物理现象都对应一定的时间尺度范围, 且在不同空间尺度及相应的时间尺度范围内进行材料设计所用的理论方法和模型是不相同的。该报告作者将理论方法、空间尺度及相应的时间尺度三者的对应关系表示在图 23-2 中。

该报告作者把这张图比作理论描述的“食物链”, 意图是由前一级时间/空间尺度范围的计算所输出结果, 可作为下一级(较大)时间/空间尺度范围进行计算的输入。从图 23-3 可以看出, 在不同时间/空间尺度范围内所用

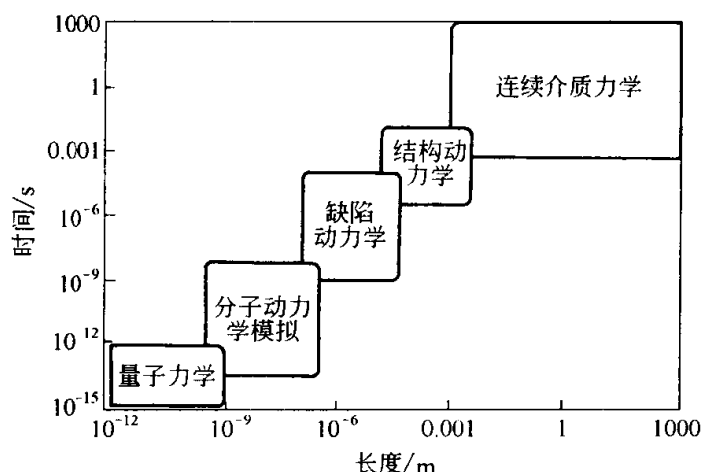


图 23-2 理论方法与空间、时间尺度对应图

理论方法包括了从量子力学计算到分子动力学模拟，然后是缺陷动力学、结构动力学，再向连续介质力学方法过渡。

必须指出，关于材料设计工作范围的确定和层次的划分，并没有固定的、通用的“范式”（paradigm）。应根据所设计材料的种类、设计要求，以及当前能达到的“可设计”程度来确定具体的设计目标。本节所介绍的内容可作为一般参考。下文中读者还会看到某些具体设计事例。

23.4 材料设计的主要途径

材料设计作为现代材料科学的一个分支学科，尚处在草创阶段。材料的品种繁多，性能各异，不同材料的复杂程度不同，理论设计的可实现程度也差别很大。所以正如上节末尾提到的，目前关于材料设计的范式还没有正式建立起来，还需要有一个探索和积累的过程。概括地讲，设计方法主要是在经验规律基础上进行归纳或从第一性原理出发进行计算（演绎），更多的则是两者相互结合与补充。常见材料设计的途径大致可分为如下几类。

23.4.1 材料知识库和数据库技术

数据库是随着计算机技术的发展而出现的一门新兴技术。材料知识库和数据库就是以存取材料知识和性能数据为主要内容的数值数据库。

计算机化了的材料知识和性能数据库具有一系列优点：如存储信息量大、存取速度快、查询方便、使用灵活；具有多种功能，如单位转换及图形表达等；已获得广泛应用，并可以与 CAD（计算机辅助设计）、CAM（计算机辅助制造）配套使用，也可与人工智能技术相结合，构成材料性能预测或材料设计专家系统等。与早期数据的自由管理方式和文件管理方式相比，计算机的材料知识库和性能数据库还具有数据优化、数据独立性、数据一致性，以及

数据共享及数据保护等优点。

在数据库系统中,还应用一个负责数据库管理和维护的软件系统,称为数据库管理系统,它对数据库中数据的各种操作提供一种共用的方法,负责数据库的建立、操纵、管理和维护。数据库管理系统又分为层次型、网络型和关系型3种,其中关系型数据库管理系统的出现,促进了数据库的小型化和普及,使得在微观上配置数据库系统成为可能。

除了数据管理软件外,数据的收集、整理和评价是建立数据库的关键。一个材料数据库通常应包括材料的性能数据、材料的组分、处理过程、材料的试验条件以及材料的应用和评价等。

当前,国际上的材料数据库正朝着智能化和网络化的方向发展。智能化是使材料数据库发展成为专家系统,网络化是将分散的、彼此独立的数据库相连而成为一个完整系统。

利用大型知识库和数据库辅助材料设计的一个典型例子是日本三岛良绩和岩田修一等建立的计算机辅助合金设计 (computer-aid alloy design, 简称CAAD系统)。这一项目得到日本政府的支持,其任务是为未来的可控热核反应炉所用材料进行设计和选择。在大型计算机中储存了各种与合金设计有关的信息,其中包括各种元素的基本物理化学数据、合金相图、合金物性的各种经验方程式,各类合金体系的实验数据,各种合金的性能、用途以及有关文献目录等等。它们以元素的含量(百分数)为坐标,构筑以70多种元素的含量为坐标的多维空间,将上述各种信息记录在该多维空间中,然后按下列步骤实现计算机辅助合金设计,框图如图23-3所示。

① 输入对材料性能的要求。即要给出定量数据、性能的上限(P_U)和下限(P_L),以及使用条件 C :

$$C = (c_1, c_2, \dots, c_k) \quad P_U = (a_1, a_2, \dots, a_r)$$

$$P_L = (b_1, b_2, \dots, b_r)$$

② 检索材料信息,寻找符合要求的资料。材料 M 由一系列变量(元素组成、热处理条件等)表征。

$$M = (m_1, m_2, \dots, m_p)$$

检索合乎要求 I 的材料,构成集合

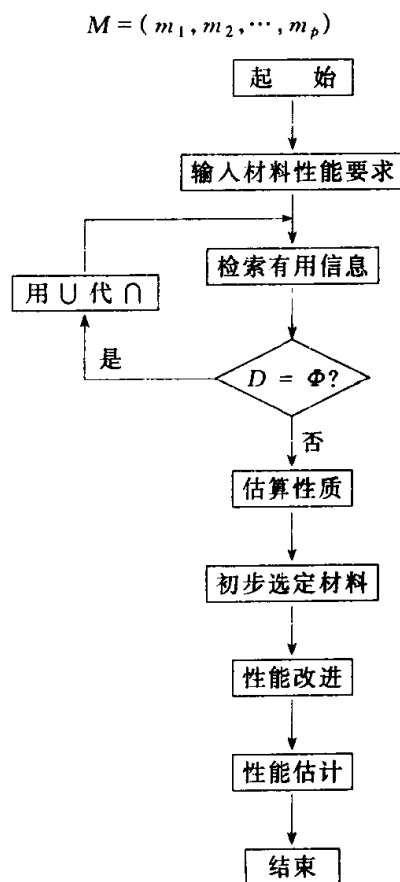


图 23-3 合金设计程序

$S_i (i=1,2,\dots,r)$, 根据对材料的全面要求, 求各 S_i 的交集。

$$D = S_1 \cap S_2 \cap \dots \cap S_r$$

若 D 不是空集, 说明现有资料已可提供全面满足要求的材料, 即可输出有关信息给用户。若 D 是空集, 说明必须寻找新材料, 此时将 M_k 有关的性质分为两段, 即仅由成分决定的性质 (如核性质、密度、热容等) 和与显微结构有关的性质, 对满足前一类性质的材料的集合 S'_i 仍取交集

$$D_s = S'_1 \cap S'_2 \cap \dots \cap S'_r$$

对满足其余性质的材料的集合 S''_i , 则取并集

$$D_t = D_s \cup S''_1 \cup S''_2 \cup \dots \cup S''_r$$

③ 计算所选材料的性质。对前一步骤选得的 D_t 集合, 根据理论、经验式、内插、外推等各种方法估算其性质。

④ 在计算性质的基础上寻找指标高的未知材料, 将预报点规定为初步选定材料。

⑤ 应用演绎法、归纳法和数据库中的资料, 试图改善初步选定材料的性能, 以推荐最终选定的材料。

⑥ 计算最终选定材料的性能。

CAAD 系统的内容还在充实完善, 他们的初步成功显示了这类合金设计是有希望的。

23.4.2 材料设计专家系统

这是一种具有相当数量的与材料有关的各种背景知识, 并能运用这些知识解决材料设计中有关问题的计算机程序系统。针对特定项目的需要, 在一定范围和一定程度上它能为某些特定性能材料的制备提供指导, 以帮助研究人员进行新材料的开发。专家系统的研究始于 20 世纪 60 年代中期, 近年来应用范围越来越广。最简单 (原始) 的专家系统包括一个知识库和一个推理系统。专家系统还可以连接 (或包括) 数据库、模式识别、人工神经网络以及各种运算模块。这些模块的综合运用可以有效地解决设计中一些有关的问题。

最理想的专家系统, 是从基本理论出发, 通过计算和逻辑推理, 预测未知材料的性能和制备方法。但由于制约材料结构和性能关系的因素极其复杂, 在可以预见的将来, 这种完全演绎式的专家系统还难以实现。目前的专家系统是以经验知识和理论知识相结合 (即归纳与演绎相结合) 为基础的。

材料设计专家系统大致有以下三类。

① 以知识检索、简单计算和推理为基础的专家系统。由于材料科学研究需要的知识面广, 有关资料极其庞杂, 任何一位专家都不可能记住全部有关资料, 所以单靠个人就会丧失许多灵活运用这些资料的机会。利用计算机

则可以弥补这个缺陷。

② 以模式识别和人工神经网络为基础的专家系统。模式识别和人工神经网络是处理受多种因子影响的复杂数据集、用于总结半经验规律的有力工具。材料设计中两个核心问题是结构-性能关系和制备工艺-性能关系。这两类关系都受多种因素的制约，故可用模式识别或人工神经网络（或二者结合）从已知实验数据集中总结出数学模型，并据此预报未知材料的性能和达到此性能的优化配方、优化工艺等。

③ 以计算机模拟和运算为基础的材料设计专家系统。材料研究的核心问题之一是材料的结构和性能关系。在对材料的物理、化学性能已经了解的前提下，有可能对材料的结构与性能关系进行计算机模拟或用相关的理论进行运算，以预报材料性能和制备方案。

除上述 3 类外，近年来还发展了以材料智能加工为目标的材料设计专家系统。材料智能加工是材料设计研究的新发展，其目标是通过在位传感器在材料制造过程中采集信息，并输入智能控制系统以实现控制决策，使制备中的材料能循着最佳途径成为性能优良、稳定以及成品率高的材料。材料智能加工研究始于 20 世纪 80 年代中期，已在大直径砷化镓单晶制备、碳纤维增强碳素复合材料制备以及粉末热压和喷射成型等方面得到应用并取得良好效果。

23.4.3 材料设计中的计算机模拟

利用计算机对真实的材料系统进行模拟“实验”，提供模拟结果，指导新材料研究，是材料设计的有效途径之一。自从电子计算机出现以来，人类的计算能力差不多每 5 年提高一个数量级，过去 50 年间计算机浮点运算速度大约提高了 10^8 倍，加上算法的进步，人类计算能力提高了大约 10^{10} 倍。目前，大规模并行运行计算机的发展，还会带来计算能力的进一步提高。因此，对材料科学来说，计算机模拟，特别是采用太拉级 (10^{12}) 计算机的模拟将导致革命性的发展，其目标是发展能预报实际材料性能的模拟计算，而且能对大范围内不同空间和时间尺度上发生的物理现象进行模拟，从而极大地促进材料科学的发展。美国能源部提出的材料战略模拟计划 (Strategic Simulation Initiative, 缩写 SSI)^[2]，以及美国西北大学高性能钢研究组 (Steel Research Group, 缩写 SRG) 开展的计算材料设计^[6]，可以作为这方面的代表性工作。我们将在 23.6 节中加以介绍。本节先对计算机模拟作一般性介绍。

材料设计中计算机模拟对象遍及从材料研制到使用的全过程，包括合成、结构、性能及使用等。人们把涉及复杂材料某一过程、某一层次上物理现象的基本性质准确地转化为一定的数学模型，该模型一方面可以由计算机

求解，另一方面可以描述或者预测某些可观测的材料性能。这便是计算机模拟的基本做法。

材料科学中的计算机模拟，现在被认为与材料理论和实验享有同等重要性。在许多难以进行或无法进行实验的情况下，模拟则尤为突出地重要。所以模拟计算和实验工作两者应是完美的结合。

计算机模拟按模拟尺度可分为3类。

① 原子尺度模拟。所用的主要方法是分子动力学方法和蒙特-卡洛方法。分子动力学方法应用极为普遍，它在单个原子的牛顿运动方程组合的基础上，根据给定的原子间相互作用力和外场，计算多粒子系统的结构和动力学过程。蒙特-卡洛方法是用人工抽样试验来估计数值数学问题并求解的一种方法，它在计算系统的时间平均性质上比分子动力学方法更为有效。早期的原子尺度模拟研究及其成果，可参看文献[7]。

② 宏观尺度模拟。这是将材料看成连续介质，其宏观物理现象由代表质量、动量、能量平衡原理的偏微分方程所表述。对特定的材料则通过本构方程（constitutive equations）来反映。基于连续介质模型的模拟计算有两个根本出发点：一是要建立和反映材料宏观特性可测量之间数学的关系式，二是要运用材料微观性质的知识，探索它们同宏观性质之间的联系。早期的连续模型计算研究，可参阅文献[8]，其他例子例如：在功能梯度材料研制过程中用模拟方法计算热应力分布，为寻找合理的结构提供依据。在用液态合金经急冷而制得非晶态合金宽带材料时，采用计算机模拟计算液体合金快冷时的传热传质过程，有助于设计合理的设备和工艺。基于热力学方法的模拟计算，可以预测某些材料的相变过程及相变产物的显微结构，等等。

③ 介观尺度模拟。这是介于原子尺度和连续介质之间的模拟，也是目前发展得尚不成熟，困难最大的模拟。如何从原子尺度的模拟逐步推向连续介质模拟，或者反过来，如何从连续介质模拟逐步往下同介观尺度及原子尺度模拟挂起钩来，正是计算机模拟研究所要解决的问题。上面讲到美国能源部的SSI计划提出了所谓的“粗粒化”（coarse graining）方法（参见第23.6节），其目的就在于推进介观尺度的模拟。

现在回过头来再对原子尺度的模拟作一些补充。下一节将讲到，要对一种材料的电子结构进行第一性原理的计算，无论是用能带理论还是用量子化学方法，都是费时费工的，而且可靠的结果目前主要限于基态，对于激发态迄今计算方法不多，且精度有限。因此，如果能将材料中粒子之间的相互作用势用适当的“有效势”来取代，并在此基础上进行计算机模拟，这样虽然会损失一些细节效应，但仍大体上反映出由相互作用势所决定的电子结构，以及由电子结构所决定的材料性质，这无疑是一种很好的想法。基于这种

“有效势”的计算机模拟已得到广泛应用。

计算机模拟中原子间“有效势”模型的选择，显然因材料的类型而异，早期主要采用纯经验的拟合势，近年来越来越多地考虑从第一性原理计算导出合适的“有效势”，包括采用集团近似的量子化学计算，从几何位形的变化来求出相应的有效势模型。这样做不仅是为计算机模拟提供了更为可靠的依据，而且还可以反过来论证基本第一性原理的电子结构计算方法同有效势之间的联系。

有了适当的原子间相互作用有效势之后，便可以针对要研究的问题进行计算机模拟。模拟技术除分子动力学方法、蒙特-卡洛方法外，还有能量极小值法。后者最为简单，它是在有效势作用下改变原子分布的几何位形，从而求出对应于能量为极小值的原子位形。分子动力学方法是计算量较大的模拟方法，它是对离散的时步来求解牛顿运动方程，因而在模拟中包括了动能，特别有利于研究物质在高温下的行为，如扩散、相变等问题。在图 23-2 中，我们看到分子动力学占有自己的相应位置，它在量子力学适用范围与缺陷动力学、结构动力学范围之间起连接作用。蒙特-卡洛方法实际上是一种统计力学的计算技术，对结构引入某种随机变化，而根据能量判据加以取舍，这对于处理无序系统特别有利。

计算机模拟方法有些情况下比第一性原理的电子结构计算省事，而且可以获得与时间有关的物理量和热力学量的信息，这是量子力学所不能得到的。利用计算机模拟，不仅可以模拟固体，也可以模拟液体在不同温度下的行为，特别在研究庞大复杂的系统方面具有良好的发展前景。

23.4.4 基于第一性原理的计算

物质材料是由许多相互接近的原子排列而成，排列可以是周期的，也可以是非周期的。从原子尺度看，材料的构建单元是电子和核，它们之间通过库仑力相互作用，它们运动服从量子力学薛定谔方程。70 多年前建立的量子力学，打开了了解物质原子组成物的行为特性及其物理规律的大门。反过来，可以利用这些知识来设计和控制物质材料的结构和合成。从原则上说，任何材料的性能和结构应能在量子力学方程和已知物理定律基础上通过计算而得出。人们常常称之为第一性原理的计算。不过，实际上并非这么简单，求解薛定谔方程时数学上的复杂性来自电子的费米性质，以及难以描述众多粒子之间的相互作用，因此除了最简单的一些情形下可以精确求解之外，稍复杂的情况就不好办了。材料中电子和核的数目达到 $10^{24}/\text{cm}^3$ 的数量级，直接利用计算机进行模拟也是不可能的。

虽然如此，第一性原理的计算仍在不断取得进展。1998 年 W.Kohn 被授予诺贝尔奖，因为他在发展第一性原理计算方法——密度泛函理论方面作

出了重大贡献。这是一个信号，表明第一性原理计算已是一个成功的领域。

与此同时，现在计算机能力的提高，已经接近达到可以对包含数千个原子的大系统求解量子力学方程，以阐明发生在其中的复杂现象。在此基础上，人们已有可能寻找向介观尺度进行模拟计算的途径与方法了。

关于从第一性原理出发进行材料设计的进展，读者可参阅文献[9]，它是为纪念发展了赝势法和总能量方法的著名学者 Cohen Marvin L 而出版的。该书由 30 多名学者撰写，介绍了用量子理论（第一性原理）阐明和测报“实际材料”（real materials）性能所取得的进展。从该书书名上便突出了一个概念：即量子理论用于“实际”材料而非“理想”材料。书中说，量子力学建立于 20 世纪 20 年代，但对于固体的了解仅在过去 30 年才开始，原因是固体中存在复杂的电子-离子、电子-电子相互作用。目前，固体量子理论的发展在利用计算机的条件下已经用来探索和预测尚未合成的新材料。Cohen 教授所发展的第一性原理的方法，近年来在预测新材料性能方面有两个突出的成功事例：一是预报存在 Si 的高压金属相及其超导性，二是预报 C_3N_4 超硬材料。

还应指出，材料设计中的计算机模拟和基于第一性原理的计算，两者是相辅相成，互相促进的。介观尺度以及宏观尺度的模拟所用数学模型中的唯象参数最好能从第一性原理的计算中获得。第一性原理的计算能达到某种尺度以后，便希望通过模拟计算得以继续推向更大的尺度。材料界年轻的学者们应培养自己善于思考，如何为从原子到大块材料之间多种尺度上的设计计算架起桥梁。

第一性原理的计算方法很多，将在下节中选择主要的几种方法作简单介绍。此外，量子化学方法也是基于第一性原理的计算方法，它和固体量子理论是相互补充的，在材料设计中有同样重要的作用。关于量子化学与材料设计可参阅文献[17]的第 11 章。

23.5 基于第一性原理的主要计算方法

本节主要介绍近年来在实际材料的理论预报中起了重要作用的一些第一性原理的计算方法。这里只简单地介绍这些方法的基本思路和框架，以及大体能解决哪类问题，并不作详细推导和论证。若需要深入具体的了解，请读者参阅文献[9]。

材料的许多基本物理性质是由其电子结构决定的。要确定它们的电子结构，原则上需采用基于第一性原理的计算方法。第一性原理的出发点便是求解多粒子系统的量子力学薛定谔方程。这一系统的非相对论形式的哈密顿量可写成：

$$H = \sum_p -\frac{\hbar^2}{2M_p} \nabla_p^2 + \frac{1}{8\pi\epsilon_0} \sum_{p \neq q} \frac{Z^2 e^2}{|\mathbf{R}_p - \mathbf{R}_q|} + \sum_i -\frac{\hbar^2}{2m_0} \nabla_i^2 + \frac{1}{8\pi\epsilon_0} \sum_{i \neq j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} - \frac{1}{4\pi\epsilon_0} \sum_{i,p} \frac{Ze^2}{|\mathbf{r}_i - \mathbf{R}_p|} \quad (23-1)$$

式中, \mathbf{R}_p 、 \mathbf{R}_q 为原子核的位矢; \mathbf{r}_i 、 \mathbf{r}_j 为电子的位矢; M_p 、 m_0 分别为原子核和电子的质量。式 (23-1) 中包括离子和电子的动能项, 也包括离子之间、电子之间和离子-电子之间的相互作用项。这样复杂的两种粒子多体系统, 必须采用合理的简化和近似才能处理, 否则寸步难行。

由于核的质量比电子质量大很多, 因而, 电子的响应速度极快, 不妨将离子视为静止的, 这就是有名的玻恩-奥本海默 (Born-Oppenheimer) 绝热近似, 从而可将离子的运动与电子的运动分开来处理。

经过上述简化, 式 (23-1) 中前面两项可以舍去, 式中最后一项, 即电子与离子相互作用项, 可以用晶格势场 $\sum_i V(\mathbf{r}_i)$ 来代替。于是得到电子系统的哈密顿量的简化形式

$$\hat{H} = - \sum_i \nabla_i^2 + \sum_i V(\mathbf{r}_i) + \frac{1}{2} \sum_{i,j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (23-2)$$

这里已采用原子单位, 即 $e^2 = 1$, $\hbar = 1$, $2m_0 = 1$ 。

式 (23-2) 所对应的薛定谔方程实际上仍然很难求解, 困难在于存在电子-电子之间和电子-核 (离子) 之间的库仑相互作用项。系统的状态应该在库仑相互作用能和动能两方面取得均衡, 使总能量最小。进一步可以通过哈特里-福克 (Hartree-Fock) 自洽场近似将多电子的薛定谔方程简化为单电子的有效势方程, 在哈特里-福克近似中, 引入了电子与电子的交换能。它考虑了费米全同粒子的交换反对称性, 即系统总波函数相对于任意交换一对电子应是反对称的。电子系统的真实总能量与哈特里-福克总能量的差值称为关联能。交换能和关联能处理起来较为棘手。下面介绍的局域密度泛函理论和准粒子 GW 方法能够较好地考虑交换能和关联能, 而 Car-Parrinello 方法为研究有限温度下离子和电子的动力学特性提供了一个有力的工具。

23.5.1 密度泛函理论^[10,11]

20 世纪 60 年代, Hohenberg、Kohn 和 Sham (沈吕九) 提出了密度泛函理论, 简称 DFT, 这个理论不但建立了将多电子问题化为单电子方程的理论基础, 同时也给出了单电子有效势如何计算的理论依据。DFT 是研究多粒子系统基态的有效方法, 它的基本要点如下。

(1) 处在外势场 $V(\mathbf{r})$ 中的相互作用的多电子系统, 电子密度分布函数 $\rho(\mathbf{r})$ 是决定该系统基态物理性质的基本变量。

(2) 系统的能量泛函可写作

$$E[\rho'(\mathbf{r})] = \int V(\mathbf{r})\rho'(\mathbf{r})d\mathbf{r} + T[\rho'(\mathbf{r})] + \frac{e^2}{2} \int \frac{\rho'(\mathbf{r})\rho'(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + E_{xc}[\rho'(\mathbf{r})] \quad (23-3)$$

式中右边第一项为电子在外势场的势能；第二项为动能；第三项为电子间库仑作用能；第四项为交换-关联能泛函。DFT 证明，当 $\rho'(\mathbf{r})$ 为基态的电子密度分布 $\rho(\mathbf{r})$ 时，能量泛函 $E[\rho'(\mathbf{r})]$ 达到极小值，且等于基态能量。

(3) 将系统的电子密度分布写为 $\rho(\mathbf{r}) = \sum_{i=1}^N |\phi_i(\mathbf{r})|^2$ ，其中 $\phi_i(\mathbf{r})$ 为单电子波函数，将 ϕ_i 代入式 (23-3) 求变分极小值，可导出 Kohn-Sham 方程

$$\{-\nabla^2 + V_{KS}[\rho(\mathbf{r})]\}\phi_i(\mathbf{r}) = E_i\phi_i(\mathbf{r}) \quad (23-4)$$

其中

$$V_{KS} = V(\mathbf{r}) + \int d\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})} \quad (23-5)$$

这里的问题是交换-关联能量泛函 $E_{xc}[\rho]$ 到底取什么形式，这是非常重要的。在具体计算中，常用所谓局域密度近似，简称 LDA (local density approximation)。它是一个简单可行而又富有实效的近似。在密度泛函理论中，LDA 占有重要地位。LDA 的基本想法是，可利用均匀电子气的密度函数 $\rho(\mathbf{r})$ 来得到非均匀电子气的交换-关联泛函。在 LDA 框架下，可以给出 $E_{xc}[\rho]$ 的具体形式，然后对式 (23-4) 和式 (23-5) 进行自洽计算。

总之，密度泛函理论认为，固体的基态性质是由其电子密度惟一确定的。在局域密度近似 LDA 下，可从求解一组单粒子在有效势场中运动的方程而得到此电子密度分布，在此基础上计算固体的有关特性，它比哈特利-福克自洽近似更为严格、更为精确，因此，LDA 是研究固体能带、表面、界面、超晶格材料和其他低维材料的强有力工具。对于许多半导体和一些金属的基态物理性质，如晶格常数、结合能、晶格力学性质都能给出与实验值符合得相当好的结果。对大部分半导体和金属也能给出与实验符合良好的价带。但也遇到了一些困难，LDA 对金属的 d 带宽度以及对半导体禁带宽度的计算结果与实验值相差 35% ~ 50%。特别是用 LDA 计算 Ge 的能隙时，得出能隙 $E_g^{LDA} < 0$ ，从而错认为 Ge 是半金属。LDA 计算给出不合理结果的例子还有一些，说明这个方法还存在缺点，有待修正和发展。

23.5.2 准粒子方程，GW 近似^[12,13]

局域密度泛函理论取得了相当大的成功，但也遇到了一些困难。LDA 计算半导体能隙偏小，而计算金属价带宽度偏大。Kohn-Sham 方程的描述虽是严格的，但多粒子系统的全部复杂性仍然包含在交换-关联能泛函 $E_{xc}[\rho]$

中,而它是未知的。再者, LDA 计算可确定系统基态的能量、波函数和有关物理量算符期待值等,但一般认为,局域密度泛函理论不能给出系统激发态的正确信息。

为了克服这些困难,人们利用理论物理中的一些成果,并把它们移置到凝聚态物理学中来。早在 20 世纪 50 年代, Landau 在研究费米液体中引入了准粒子概念。20 世纪 60 年代 Hedin 提出了从多粒子系统格林函数出发,计算各种复杂多体效应对准粒子能量的贡献的方法,称为自能方法。于是, Hybertsen 和 Louie 借助准粒子概念和由单粒子格林函数求自能的方法,提出了准粒子近似。

在准粒子近似中,半导体能隙可理解为相互作用电子气中准粒子元激发所要求的能量,系统的低激发态看成是由独立的准粒子元激发组成的电子气。准粒子的能量 E_{nk} 和波函数 ϕ_{nk} 可由与局域密度泛函理论中相似的单粒子方程

$$[T + V_{\text{ext}} + V_{\text{coul}}]\phi_{nk}(\mathbf{r}) + \int d\mathbf{r}' \Sigma(\mathbf{r}, \mathbf{r}', E_{nk})\phi_{nk}(\mathbf{r}') = E_{nk}\phi_{nk}(\mathbf{r}) \quad (23-6)$$

确定这里 T 为动能项, V_{ext} 、 V_{coul} 分别为外势场和库仑项。同 Kohn-Sham 单粒子方程相比,式 (23-6) 中已将交换-关联项写成了新的形式,并引入了自能算符 Σ , 它与能量 E_{nk} 有关,代表了电子间交联-关联等各项相互作用。

求解方程 (23-6) 的主要问题在于寻找自能算符 Σ 的近似。GW 近似就是为此而提出的。这个近似认为:在最低一级近似下,自能算符可用单粒子格林函数 G 和动力学屏蔽库仑作用 W 表示,故称 GW 近似,即

$$\Sigma(\mathbf{r}, \mathbf{r}', E) = \frac{1}{2} \int d\omega e^{i\delta} G(\mathbf{r}, \mathbf{r}', E - \omega) W(\mathbf{r}, \mathbf{r}', \omega) \quad (23-7)$$

式中, δ 是正的无限小量。

在这里只说明一个思路和框架,实际应用时还有不少具体细节需要处理。原则上准粒子方程 (23-6) 需要自洽求解。所幸的是,在实际计算中,局域近似的 Kohn-Sham 单粒子方程的波函数 ϕ_{nk}^{LDA} 与准粒子 GW 近似的波函数 ϕ_{nk}^{qp} 十分接近,因此,可以将 ϕ_{nk}^{LDA} 作为零级近似,用它来构造单粒子格林函数,将自能用迭代式展开,用 ϕ_{nk}^{LDA} 代替自洽求解准粒子波函数 ϕ_{nk}^{qp} , 这样可以得到准粒子能量

$$E_{nk}^{\text{qp}} = E_{nk}^{\text{LDA}} + \langle \phi_{nk}^{\text{LDA}} | \Sigma - V_{\text{xc}}^{\text{LDA}} | \phi_{nk}^{\text{LDA}} \rangle \quad (23-8)$$

准粒子近似已被成功地应用于半导体能隙计算。以往 LDA 的计算值 E_g^{LDA} 普遍比实验值小。采用准粒子 GW 近似后,对于金刚石、Si、Ge、

GaAs、LiCl、GaN 等晶体，其计算值 E_g^{op} 与实验值符合得很好。有人用 GW 近似研究半导体的准粒子能带，以及 MnO 的准粒子能谱，得到了与实验比较接近的结果。现在已经清楚，LDA 计算中存在问题的症结，是由于将多电子系统相互作用简单归结为局域的交流-关联势引起的；而在 GW 近似中，用自能 Σ 代替局域的交流-关联势能够更完善地反映非均匀系统的多体效应，并在此基础上计算准粒子能量。所以近 10 年来 GW 近似取得了相当大的成功。

23.5.3 Car-Parrinello 方法^[14,15]

在原子水平的计算机模拟计算中，分子动力学 (molecular dynamics, 简记为 MD) 是十分有效的方法。其特点是利用原子间相互作用势 (经验的或从理论导出的)，模拟计算系统的平衡和非平衡态的物理性质，包括原子团簇，非晶态物质和流体等。对于惰性元素组成的系统，原子间势可用 Lennard-Jones 势。对于金属体系，近年提出了原子嵌入媒质的模型势。对于共价晶体，可用 Stillinger-Weber 势。对于半导体则要求引入三体相互作用势。在固体电子结构计算中，局域密度泛函理论 LDA 取得了很大成功。有鉴于此，如何将 LDA 和 MD 这两种方法结合起来，是人们所追求的。1985 年 Car 和 Parrinello 成功地将两种方法有机地联系起来了。

Car-Parrinello 方法最重要的一点是在真实的物理系统中引入一个虚拟的电子动力学系统，这样组成的新系统的势能 E 是离子和电子自由度的一个总泛函。虚拟系统的广义经典拉格朗日量

$$L = \sum_i^{\text{occ}} \int d\mathbf{r} \mu_i |\dot{\psi}_i(\mathbf{r})|^2 + \frac{1}{2} \sum_I M_I \dot{\mathbf{R}}_I^2 - E[\{\psi_i\}, \{\mathbf{R}_I\}] \\ + \sum_{i,j} \Lambda_{i,j} \left(\int d\mathbf{r} \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) - \delta_{ij} \right) \quad (23-9)$$

式中， L 为二套参量 $\{\psi_i\}$ 和 $\{\mathbf{R}_I\}$ 的泛函； μ_i 是具有适当量纲的任意参量。从式 (23-9) 看出， μ_i 相当于电子的“质量”，实际上它起着调节电子运动时间标度的作用。为简单起见，可以令 μ_i 与电子波函数 $\{\psi_i\}$ 无关，即 $\mu_i \equiv \mu$ 。于是式 (23-9) 的拉格朗日量中，第一项和第二项分别是电子和离子的动能。 $E[\{\psi_i\}, \{\mathbf{R}_I\}]$ 是电子和离子耦合虚拟系统的势能，

$$E[\{\psi_i\}, \{\mathbf{R}_I\}] = \sum_i^{\text{occ}} \int d\mathbf{r} \psi_i^*(\mathbf{r}) \left(-\frac{1}{2} \nabla^2 \right) \psi_i(\mathbf{r}) \\ + \int d\mathbf{r} V_{\text{ext}}(\mathbf{r}) \rho(\mathbf{r}) + \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \frac{\rho(\mathbf{r}) \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \\ + E_{\text{xc}}[\rho(\mathbf{r})] + \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \quad (23-10)$$

式 (23-9) 中最后一项的拉格朗日乘子 Λ_{ij} 是为了保证 $\{\psi_i\}$ 的正交性约束条

件而引入的。

从式 (23-9) 容易得到相应的 Euler 方程

$$\mu\ddot{\psi}_i = -\frac{\delta E}{\delta \psi_i^*} + \sum_j \Lambda_{ij}\psi_j \quad (23-11)$$

$$M_I\ddot{R}_I = -\frac{\delta E}{\delta R_I} \quad (23-12)$$

这里式 (23-11) 相当于式 (23-4) Kohn-Sham 方程, 式 (23-12) 相当于离子运动的经典动力学方程。

以上便是这个理论方法的框架。在这个框架中, 电子参量 $\{\psi_i\}$ 的动力学是一种想像的过程, 仅作为一个工具来实现所谓动力学模拟退火 (dynamical simulated annealing)。当 $\{\dot{\psi}_i\}$ 、 $\{\dot{R}_I\}$ 等变小, 相当于系统的温度 T 下降, 而直至温度 $T \rightarrow 0$ 时, 系统达到 E 极小的平衡态。在平衡态 $\dot{\psi}_i = 0$, 就会过渡到 Kohn-Sham 方程。

为了检验这个方案的有效性, 他们对晶态硅的静态和动力学性质作了计算, 并取得成功。此方法问世 10 多年来, 已被证明是研究凝聚态系统结构特性、电子特性和动力学特性的有力工具。它不仅能得到系统基态的结构和电子性质, 而且可用以研究有限温度下系统的离子和电子特性, 这是人们特别感兴趣的。

以上介绍了三种代表性的计算方法, 可分别应用于研究系统基态、激发态和动力学过程。这些方法的进展, 表明固体理论已能用于阐明和预报实际材料的性能。当然固体量子理论还有其他一些方法, 如紧束缚 (TB) 方法、赝势方法、有效质量理论及量子 Monte-Carlo 方法等, 它们在对实际材料的理论计算中都有重要应用, 这里不作介绍。

值得指出, 所谓“第一性原理”计算, 指的是从第一性原理出发, 针对实际材料体系和所研究的问题进行数值计算, 在具体处理时, 仍然需要合理的近似。因此提出恰当的物理模型是非常重要的, 计算材料科学界常说材料的“Modeling”指的也是这个意思。我国材料界学者也常常提到: 我们不能仅仅去“拼”计算时间、“拼”经费, 而应该同时注重创造性地发展理论模型和计算方法。

23.6 国外研究动态与展望

本节着重介绍以下三方面的动态及其展望。

23.6.1 1995 年美国 NRC 的专门报告^[1]

1995 年美国 Naval Research Laboratory (NRL) 为制定长期战略计划, 请求美国 NRC 组织专门小组, 对“材料科学的计算与理论技术”进行调查

并提出了报告。报告主要希望回答以下问题：即随着计算机能力将可能提高 2~3 个数量级，以及理论方法的进步，在未来 10 年中对材料科学和对 NRL 会带来什么影响与机遇？该报告的重点是考察原子水平上的材料研究前景。专门小组挑选了 13 个需要从原子水平进行材料研究并在 20 世纪末面临挑战的领域进行调查。这 13 个领域是：新材料，半导体，光学性质，表面与界面，人工膜生长，材料动力学行为的纳米工程，化学动力学（表面化学、腐蚀、爆炸），爆燃（detonation）的流体动力学，材料强度、缺陷及高温材料，复合材料、聚合物及陶瓷，合金相图，磁性材料以及强相互作用系统。报告中分别对这些领域的研究进展与前景作了分析，然后就 NRL 的材料研究发展问题，分析了 9 点机遇，并提出相应建议。可以认为，这个报告比较全面地反映了当时这个方面的研究动态，至今仍有重要参考价值。

纵观近 10 多年来国外在材料设计研究方面的进展，可以作出以下综述。

(1) 新材料及其理论方法 20 世纪 80 年代中期以来，出现了许多新材料，包括高 T_c 铜氧化物超导体、富勒烯及其衍生物、纳米材料、超硬材料、人工低维量子结构材料等。其中有些材料已从基础研究对象转化为实际应用对象，如量子阱激光器、高温超导器件、巨磁阻磁头等。虽然，大多数新材料都是非预期地被发现，但应当看到，计算和理论对预报和阐明这些新材料的性能起了重要作用。如果说 20 世纪 80 年代中期以前，还难以从第一性原理来计算实际材料的性能，那么，近 10 年来情况已大为改变。一些事实表明，理论和方法的进步已能对实际材料性能进行理论预测。正如前面所说，研究者已处在通过理论计算来“设计”材料的初期阶段。

重要的理论方法有以下几个。

① 局域密度近似 (LDA)。可计算几百个原子的系统，得到材料基态的结构、振动和力学性质。其成功的事例包括：揭示了高 T_c 铜氧化物超导体的正常态性质；预言了高压下材料新相及性质 (Si 的简六方相超导性)；预报了超硬材料 C_3N_4 ；确定了一些表面、界面及团簇的结构与性质；计算了富勒烯及其衍生物的结构和性质。

② GW 准粒子近似。它是目前能成功地计算固体中电子激发能的有效方法，已用于计算包括半导体、绝缘体、简单金属、表面、界面、团簇、高压下金属、 C_{60} 等富勒烯衍生物的光学性质，能定量阐明光谱测量结果。

③ 第一性原理的分子动力学方法，即 Car-Parrinello 方法。它是目前用于动力学和热力学性质计算的有效方法。随着计算机能力的提高，预期这一方法在研究材料界面动力学、化学动力学、缺陷动力学等过程中将起更大的作用。

其他方法，如新势法、紧束缚 (TB) 总能量法、量子 Monte-Carlo 等

方法也都有进展。在 23.5 节中曾简要地介绍了上述 3 种方法。

随着计算机能力的提高,可以提出如下展望:将上述理论方法尽可能移置到大规模并行计算机上,以扩大应用范围。例如,可望 LDA 方法计算的 Supercells 含 1000 个原子上;用量子分子动力学方法计算自由能,可发展到研究熔点及其相变;不仅将现有方法过渡到并行运算,而且要发展更有效的新算法,例如发展所谓 Real-Space Methods 等,目标是希望能计算几万个个原子的系统;此外要发展处理多电子效应的更好的理论方法。目前,虽然 LDA、GW 方法取得成功,但对强关联的磁性材料,这种从头算起 (ab initio) 的方法仍不好用。问题在于如何处理大系统中的电子交换-关联效应,因此还需要发展更好的理论与方法。与此同时,从另一个极端看,还需要改进 TB 分子动力学方法,以便考察那些在几百、几千个原子的系统计算中不能加以模拟的现象。

(2) 表面与界面的研究概况 当代关键技术中一个核心问题,是如何描述不同材料怎样在原子和化学水平上结合而成固体表面的。所有合成的、通过原子水平调制的结构,都包含不同化学介质层之间的界面。这些原子水平调制的界面结构表现出完全不同于体材料的光、电、磁和力学性质。计算材料物理的任务在于,揭示发生在表面、界面上各种现象的物理内涵,补充实验不能直接测量的东西。不止于此,材料动力学行为的纳米工程学,也已提上日程,也就是说很多重要技术中的界面问题,要求在纳米尺度上和相应的时间尺度上考察原子水平的集体动力学行为。例如要求设计特定技术所要求的固-固、固-液界面材料,包括控制其摩擦、磨损、优化润滑和动态接触过程等。化学动力学包括表面化学、腐蚀等,也已形成另一类挑战性的任务,即如何利用第一性原理的计量化学来了解固体表面化学微观机制,包括氧化、腐蚀、催化等。对于原子团簇的生长过程,目前以 LDA 为基础的量子分子动力学方法只能计算几十个原子的系统。腐蚀是一个包含表面化学、相变、裂纹传播、力学性质变化等多个材料转变的复杂过程。这些问题都寄希望于计算机能力的提高和算法的进步,才能进一步作出理论预报。

从以上叙述中可以看到,计算材料科学的一个重要任务,是进一步发展凝聚相界面有关的动力学过程的理论方法。而且,这种过程是原子水平上的、属于材料的纳米物理范畴。经典的连续介质物理在当今技术进入纳米尺度物理领域后,失去了适用性。当然这里有一个微观层次与连续介质层次在理论计算方法上的衔接问题。这需从两个方面努力来加以解决。一方面连续理论要致力于发展包括某些微观过程的所谓细观力学;另一方面,从 ab initio 的量子力学计算中提取出原子水平的参数,并使之延伸到连续理论所用的本构模型 (constitution models) 中去。看来这种努力还需要一定时间,

并要求发展相应的计算手段。

(3) 薄膜、复合材料的设计问题 人工薄膜生长过程,也是由原子水平上的物理学所主宰。生长过程中发生的许多现象涉及非平衡过程中所遇到的问题。目前最有力的理论手段是分子动力学模拟和蒙特-卡洛模拟。关于这类材料的设计问题,要求理论和计算比实验省时而且省钱,能了解微观机制,并能设计和预见新材料。前面已说过(参见图 23-2),可将材料行为按空间尺度分为四个计算范围,相应的理论方法分别为量子力学、经典分子动力学、多原子缺陷动力学和连续介质力学(其中非平衡缺陷和原子的效应通过经验的本构定律来反映)。现阶段这几个范围各有自己的局限,相互之间存在“间隙”,一时还无法统一。寄希望于硬件和算法的进步,使它们各自扩大其空间和时间范围,并发生交叠。其中重要的环节是,使分子动力学模拟达到规模并行化,使之能计算几千万个原子的系统。可望在三维情况下,模拟计算的尺度从纳米到 $0.1\mu\text{m}$,时间尺度小于几十 ns。不过,这种分子动力学的现实性,还取决于从更小的尺度内由计算获得原子间相互作用势的详情。而后者赖于 *ab initio* 的第一性原理计算。由这种分子动力学模拟得到的信息,被输送到下一步缺陷动力学范畴去,然后再过渡到连续力学。

关于复合材料、聚合物和陶瓷的设计问题,理论上首先应设计好作为构建单元(building blocks)的介观实体(entities)。这种介观实体由几十个原子至几千个原子组成,它们的结构和特性是多变的,有很强的可装配性。现在加工技术已能将它们制造出来,然后由它们组装成宏观材料。所组装成的材料的最后特性,并非其组分特性的平均,而可能得到新的最佳特性。不过,复合材料的设计由于其复杂性,目前仍有赖于经验模型。在原子水平计算的介观实体的电子结构,目前可做到原子数达 1000,但加入基体(matrix)后,特别是有应力场时,情况变得复杂。当原子间势足够简单时,用分子动力学模拟,可以做到包含数百万个原子的系统,并开始用于模拟陶瓷的应力与裂纹生长之间的相互作用。这类材料原子水平的设计研究是富有挑战性的。

(4) 从理论上预报和设计材料 总的来说,平衡态下不少性能预报与实验结果符合得好,包括力学、输运性质、热学关系及光谱特性等。挑战性的任务是,非平衡态下的性能预报和各种材料转变过程的预报。更具挑战性的任务是,根据加工历史预报材料使用效能,这要求新的方法。尤其是当材料涉及表面、界面反应时这类预报更为困难。随着先进材料的最优化使用,这种预报的必要性正在增加。

(5) 当前材料设计的机遇与展望 这里将美国学者在文献[1]中所谈的观点介绍如下。该报告认为由于材料科学计算与理论的进步,考虑到计算机

能力的进一步提高,对于美国 NRL 来说,面临着以下几个方面的机遇:

① 软件并行化将有利于现有理论方法的相互结合,并可能使软件的发展得到商业支持;

② 处理复杂问题的能力增强,从而使理论计算与实验配合的可能性大为提高;

③ 材料计算的精度可能提高到热化学的精度;

④ 处理电子关联效应的理论方法可望取得进步,这对于发展算法和计算机程序具有重要意义;

⑤ 材料动力学特性研究,可以覆盖从原子尺度直到介观尺度的范围;

⑥ 计算材料强度的软件可能大为改善;

⑦ 可望实现各种材料的线性和非线性光学性质的计算;

⑧ 可望以大为提高的精度实现多类材料的相图及相变点附近的动力学性质的计算;

⑨ 从电子结构计算中可以获得原子间相互作用的唯象势。

以上几点机遇可以这样分类:前 3 点是对整个材料研究而言;第 4 点涉及理论方法对于发展计算机算法与程序的推动作用;而后面各点则是针对一些具体研究对象所作的预测。

这里所说的机遇,实际上就是对可预见的将来在发展材料科学理论与计算方面所提出的展望。该报告作者建议美国 NRL 在制定长期战略规划时充分注意抓住这些机遇,把材料理论与计算放在重要地位上。由于 NRL 是世界著名的研究机构,它在材料方面的研究水平之高,是世界公认的;加上这个报告撰写小组成员在材料界具有的权威性,使得这一报告公布以来,在材料科学界受到了广泛重视。笔者认为,这篇报告以较为丰富的内容和明确的观点反映了本领域近年国际发展水平与对今后的展望。

实际上,发达国家的著名研究单位,以及与材料研究关系密切的著名公司,都为争夺材料科学研究的领先地位,正在不遗余力地加大对材料理论与计算方面的人力与资金投入。以美国 NRL 为例,他们的凝聚态物理部有 10 余名科学家从事材料的计算和理论研究,化学部也有科学家从事这方面研究。这些人员均是从事物理学或化学研究出身,他们研究课题的共同点是应用量子力学。研究领域包括材料电子结构计算、生长现象、强关联系统、铁电体、团簇、缺陷、输运性质、多体理论、大系统的模拟计算算法研究等。化学部的研究领域有化学气相沉积、表面反应、爆炸模拟、摩擦模拟、密度泛函方法的计算等。

由此可以看出,美国在计算材料科学方面处于领先地位。10 年以前在美国已有一些公司开发材料计算软件,例如 Biosym/MIS 公司,较早在世界

上推销原子水平上的有关材料光、电、磁、热等性能的计算软件。近年,美国 Motorola 公司专门成立了 Predictive Engineering Lab, 下设“计算材料科学实验室”, 他们购置了 Biosym/MSI 的软件, 并另图发展。他们认为, 到 2006 年左右, 半导体工业技术中起关键控制作用的将是材料中的原子过程, 因此, 他们强调要在原子学 (atomistic) 水平上开展理论分析和模拟计算。美国惠普公司也宣称, 到 2010 年, 现有的微处理器, 无论从经济上或从实际应用上都将不再被采用, 因此, 要加紧在理论和计算指导下, 设计和开发新的材料和器件。日本东北大学与日立公司联合成立了计算机辅助材料设计部, 几年前就在世界范围内聘请研究人员, 开展了有特色的工作。上述这些公司或单位都在寻求国际合作, 并对邀请我国科研人员进行合作很感兴趣。

23.6.2 美国能源部的 SSI 计划

1999 年美国能源部发表了关于材料部件的战略模拟计划 (Strategic Simulation Initiative)^[2]。该计划认为, 太拉级 (10^{12}) 计算机的出现, 为材料科学带来了空前的机遇: 它将大大加快科学发现的速率, 缩短发展周期, 优化加工过程。计划提出, 新建立的“计算材料科学中心”要加快建设国家级的设施, 组成多学科人员的队伍, 以迎接材料模拟计算的革命性大发展。

SSI 的构想大致如下:

当前的计算机能力已达到可以对包括几千个原子的大系统求解量子力学方程, 在这种原子学尺度第一性原理计算的基础上, 已有可能精确地开展介观尺度的模拟计算, 介观尺度上有些宏观材料特性, 例如强度, 可能由此被确定。介观尺度上的模拟结果将被输入到连续介质模型中, 从而进行材料的宏观工程设计。所以, SSI 的目标是发展可靠地预报实际材料性能的能力。为达到这一目标, 必须做到能可靠地模拟不同空间和时间尺度上的物理现象, 发展所谓“统一的多尺度模拟” (unified multiscale simulation)。他们将材料模型按尺度分为以下 4 个层次 (见下表):

尺 度	量子/分子	原子学的 (atomistic) /nm	介 观	宏 观
空间尺度 /m	$10^{-11} \sim 10^{-8}$	$10^{-9} \sim 10^{-6}$	$10^{-6} \sim 10^{-3}$	$>10^{-3}$
时间尺度 /s	$10^{-16} \sim 10^{-12}$	$10^{-13} \sim 10^{-10}$	$10^{-10} \sim 10^{-6}$	$>10^{-6}$

关键的任务是, 模拟发生在不同空间和时间尺度上的现象和过程, 揭示这些层次之间的联系, 探明主宰材料特性的复杂规律性。SSI 的重点将是模拟实际材料的力学性质和磁学性质。但当前人们对材料底层结构、缺陷或单个晶粒是如何决定力学和磁学性质的了解并不完全, 同样, 对化学组成及热历史或其他加工过程历史是如何影响晶粒本身的结构也不完全清楚。现有描述材料物理过程的数学模型基本上都是带有唯象参数的经验性模型 (电子和

原子尺度除外)。目前要使这些参数整个地由第一性原理计算得出,是一个重要的挑战性任务,一时恐怕难以做到。SSI 建议用计算机模拟来发展所谓“粗粒化”(coarse graining)的数学模型和参数,以免除模型和参数的经验性。“粗粒化”的含义是对大系统进行适当的平均化(averaging)或集团化(aggregating),各集团块中的力场将精确地用新的“重整化的”(renormalized)相互作用模型来描述。如何达到“粗粒化”,现在并没有可以遵循的通用范式。一般说来,从较小粒度的数学模型进入下一个较大粒度的数学模型是靠直观推断并且独立无关的。尽管它们与较小尺度的数学模型之间缺少严格的联系桥梁,但只要这些模型能充分抓住根本的物理过程,就能取得相当好的结果。当然,这样做需要花费很多计算机时间和很贵的数据库资源。

SSI 给出了适用于材料力学性能模拟计算的“粗粒化”概念示意图见图 23-4。

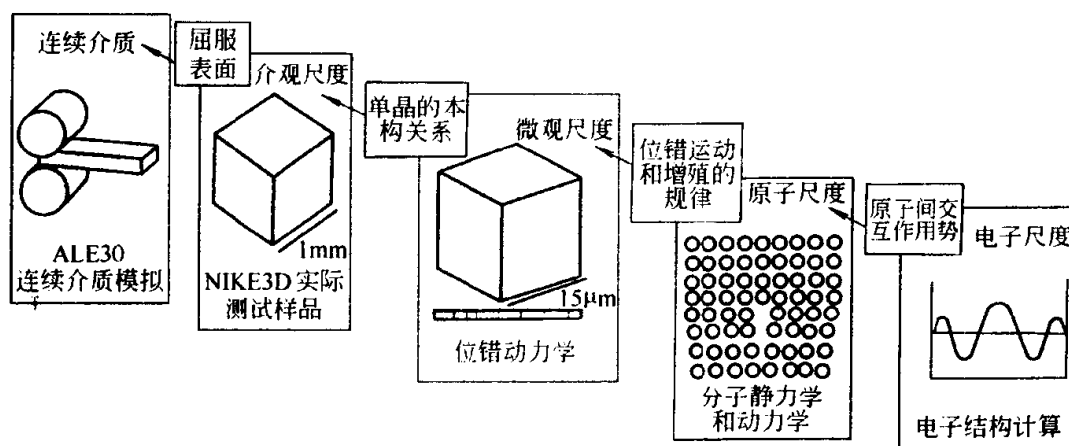


图 23-4 SSI 用于材料力学性能多尺度模拟的“粗粒化”概念示意图

“粗粒化”概念有助于建立和发展统一的多尺度模拟计算,它使各粒度之间的联系超越于定性的、唯象的性质,并把具有坚实基础的小粒度(电子水平)的计算逐步推向更大的尺度。

模拟计算量通常正比于系统的自由度数量,相互作用的复杂程度,以及模拟步骤的多少。因此,关键的任务就在于数学模型的确定。模型既要有相当的复杂性,以便可靠地反映实际的物理过程,又要不过于复杂,以致难以实现模拟计算。当今,太拉级计算机提供了可能性,使数学模型能达到相当高的复杂程度。例如,数学模型应尽可能考虑以下几方面的因素。

实际材料总有许多“缺陷”,诸如杂质、空洞、位错(包括界面不匹配位错)或其他毛病,晶粒边界,畴壁(磁性和铁电材料),第二相沉淀物,多孔性等等。正是这些缺陷的结构、它们的相互作用以及它们的热力学过程决定着实际材料的宏观性质和使用效能。有些与技术相关的材料中自然存在

或人工引入的缺陷，具有特定的合乎希望的特性，例如超导涡旋的钉扎、硬度和强度等。其他一些由缺陷引起的与技术相关的特性主要局限于使用方面，例如抗腐蚀性，某些光电子学性质，抗断裂性等。对于复杂材料，常有必要了解实际材料在远离平衡态，在高温、大应力以及与时间有关的外力下导致不可逆过程的特性。例如晶粒生长，重结晶，应力发展，弹性复形，裂纹成核和生长等。

扩展了的缺陷，例如位错和晶粒边界最直接影响材料的力学性质，包括延展性和抗断裂性。扩展缺陷与点缺陷的相互作用决定着固-熔凝结的基本机制。畴壁的结构和热力学以及它们同晶格缺陷的相互作用，对于确定永磁体的质量十分重要。许多与缺陷相互作用有关的最基本的东西现在尚未搞清楚，虽然已有一些定性的描述。原因是这些缺陷都太“大”，不属于数学模型已十分成熟的微观范围；它们又太“小”，也不属于数学模型已相当成熟的宏观范围。

一般地说，由于对材料中非均匀性 (inhomogeneities) 的性质从分子水平上了解尚不很充分，因而使介观尺度的模拟受到了局限。非均匀性指材料底层的界面过程以及决定微观结构演变的微观机制。现有唯象理论只能用于高度简化的微观模型，对许多重要现象根本不能提供信息，这些现象有赖于用原子学水平的模拟来确定，太拉级计算有可能担当用微观模型揭示这些关键过程和机制的任务，由此得到的知识和参数将成为介观模拟数学模型的输入。

SSI 声称：虽然现在还不完全了解实际材料的微观结构如何决定材料特性和使用效能，但太拉计算机具有能提供填补这个科学间隙所要求的能力。

为了建立完善的模拟计算体系，SSI 提出了必须事先具备的至少 5 种不同的模拟模式：

- ① 发现复杂特性的新关系、新范式；
- ② 提供检验数学模型结构的基准点；
- ③ 直接为数学模型计算输入参数；
- ④ 与实验比较以确认模型的适用性；

⑤ 预报复杂材料难以进行观测的特性。把这 5 种模拟模式组合起来，将使模拟计算成为强大的、革命性的工具，无疑将加速材料科学的进步。

SSI 注重硬件设施的建设和软件、程序的发展。他们注意到算法的进步至关重要。例如第一性原理的 DFT 方法，其计算量原来是与 N^3 成比例 (N 为粒子数)，现在新的算法希望做到计算量与 N 成线性比例。对于含上千电子的 *ab initio* 平面波电子结构计算程序已改进到 $N^2 \lg N$ 的比例关系。分子动力模拟的计算量也从 N^2 比例进步到线性比例 (即与 N 成正比)。

SSI 考虑发展一些专门程序, 它们是:

- ① 平面波 DFT 赝势程序, 直接模拟有关的动力学过程;
- ② 多次散射程序, 用以研究磁性材料的非平衡过程和有限温度下的性质;
- ③ 百万原子的分子动力学程序;
- ④ Pots 模型介观尺度程序, 用以模拟晶粒生长、杂质效应以及它们同裂纹的相互作用等;
- ⑤ 微观水平的三千万有限元的模拟程序。

作为例子, SSI 将多尺度模拟计算应用于以下 9 个课题:

- ① First Principles Studies of Finite Temperature Magnetism
- ② Melting a Diamond Crystal with Tight Binding Molecular Dynamics
- ③ Atomistic Simulations of Martensitic Transformations
- ④ Computational Simulations of Crack Propagation
- ⑤ Grain Sliding for Superplastic Deformation of Aluminum
- ⑥ Interfacial Behavior of Symmetric Diblock Copolymers
- ⑦ Atomistic Simulations of Radiation Damage in Metals
- ⑧ Simulation of Three-Dimensional Solidification: Dendritic Structures
- ⑨ Materials Process Modeling for Microstructure Optimization

上述美国能源部公布的白皮书, 只叙述了 SSI 的梗概, 并未发表很多技术细节, 但它仍为读者提供不少参考和启示。

23.6.3 美国西北大学 SRG 的工作^[6]

美国西北大学 Steel Research Group (SRG) 从 1985 年至今在 G.B.Olson 教授主持下一直开展材料的计算设计^[6]。

SRG 是由大学—工业界—政府三者合作组织的, 目标是探索材料设计的普遍方法、工具及建立其数据库, 并以高性能钢作为试验设计对象。SRG 的学者们相信, 当今科学知识基础已达到相当充实的程度, 使得传统的经验驱动的材料发展方式将被更为有效的理论驱动和以计算设计为基础的发展方式所取代。

SRG 的材料设计方法是以现代材料科学与工程中联接 4 个要素 (four elements) 的逻辑结构作为参照的, 4 个要素即制备、结构、性能和使用效能。SRG 材料设计的流程方块图如图 23-5 所示。

SRG 的研究者先充分收集用户的需求和材料发展前景的资料, 编写材料特性的设计目标和规范, 同时收集各类材料计算模型与数据。对照流程方块图先系统分析材料使用效能 (performance) 和材料特性 (properties) 的关系, 进而分析结构-特性关系和制备-结构关系。然后充分利用计算模型和数

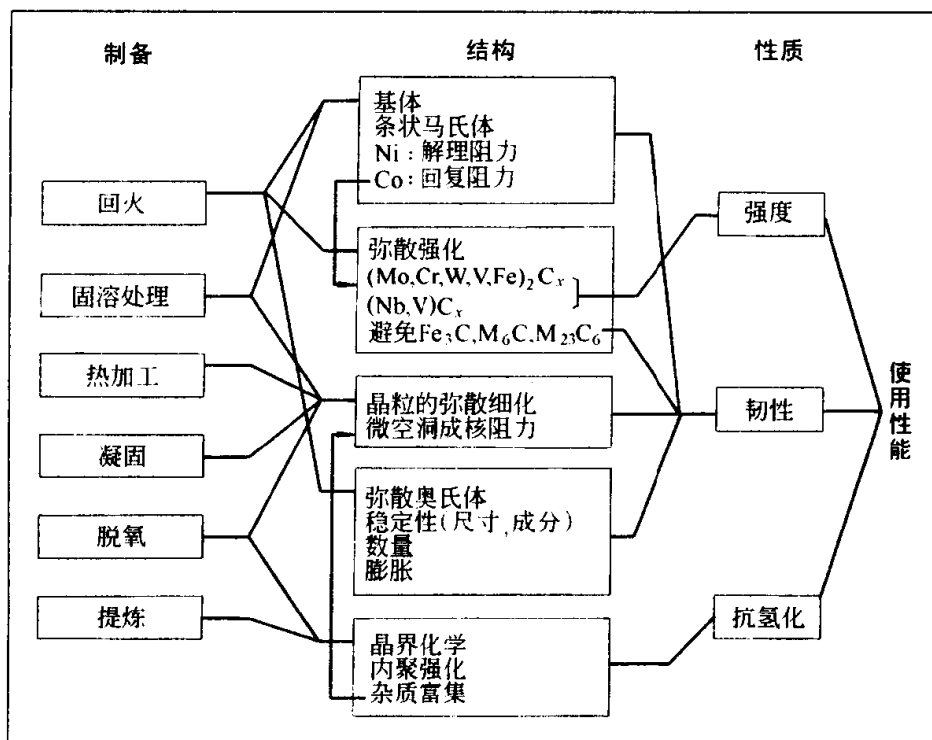


图 23-5 SRG 设计材料的流程方块图

数据库，寻找恰当的组分和制备方案，该方案应能使这些组分转变为符合设计要求的合金。过去 10 年，SRG 利用这样的框架发展出具有空前优良特性的新合金，有的已用于飞机、航空母舰、动力发动机和其他方面。

当然在这些设计过程中必须有附加的计算模型和经验数据，以补充知识上不足。而这些模型也是按尺度分为几个不同层次的。第一最粗的层次是“凝固设计”（ $10\mu\text{m}$ ）。采用例如 THERMOCALC 热力学程序，可以模拟 $10\mu\text{m}$ 尺度内的凝固过程。第二是“转变设计”（ $1.0\mu\text{m}$ ），目的是确定和控制加工温度，使得所希望的微观结构能够形成，而其他相竞争的微观结构能被抑制。第三是“微观力学设计”（ $0.1\mu\text{m}$ ）。例如在这个尺度上“grain refining”。要求采用一定模型模拟在材料变形和断裂情况下微观结构的演化过程。第四是“纳米设计”（ 1.0nm ）。近年由于高分辨率测量仪器的进步，促使新的纳米设计能力大为发展。纳米设计模型随着纳米技术的发展将会迅速发展。第五是“量子设计”（ 0.1nm ）。这是电子水平的设计。SRG 也需要用“计算量子力学”在亚原子水平上设计钢材料。

美国科学基金会（NSF）支持的 Workshop on Material Design Science and Engineering^[6]认为，应将 SGR 的设计工作推广到其他种类材料的设计上。

23.7 我国材料设计研究的进展概况

我国“863”新材料领域，自 1987 年开始设立了“材料微观结构设计和

性能预测”专题, 10 多年来坚持在这个方向上开展研究。1996 年又设立了“863 新材料模拟设计实验室”, 开展原子水平的模拟计算。多年来所开展的工作, 既有数据库基础上的专家系统和基于经验方法的性能预报, 也有第一性原理的计算。应该说, 在国家级的科研计划中设立这样的专题, 在全国范围内组织优势力量开展这类工作, 是具有我国特色的。从总体来说, 这项专题研究经过有关专家多年的努力, 取得了良好的成绩, 为我国材料设计研究的进一步发展打下了一定基础。

在设立材料微观结构设计专题研究之初, 材料界老一辈专家就曾指出: 材料科学的理论与计算研究是一项应用目标十分明确而带有基础性的工作, 需要有一支相对稳定的队伍, 坚持不懈地努力, 才能积累经验, 形成我国特色, 并在实践中有所创新。经过十多年的辛勤工作, 我国材料设计方面的研究进展, 已被编写成书, 书名为“21 世纪新材料丛书”之一: 材料设计^[17]。该书的内容, 大体上反映出当前我国材料设计的发展情况。

现代材料科学技术的发展潮流, 已将材料理论设计与计算推上了日程。我们相信, 在知识经济占主导地位的 21 世纪中, 我国材料科学技术领域要实现创新, 无疑应将材料设计放在应有的位置上。

参 考 文 献

- 1 Computational and Theoretical Techniques for Materials Science
http: //www2.nas.edu/20fa.html. National Academy of Sciences. 1995
- 2 Computational Materials Science: A Scientific Revolution about to Materialize, White Paper, March 1999, by Materials Component strategic Simulation Initiative, DOE, USA
- 3 National Nanotechnology Initiative. 见 www.nano.gov
- 4 三岛良绩. 新材料开发和材料设计学. Tokyo: Soft Science Inc, 1985
- 5 Materials Science and Engineering for 1990. Washington DC. National Academic Press, 1989
- 6 Olson G.B.. Designing a New Material World. Science. 2000, 228: 993
- 7 Baokas M., et.al.. Atomic-Scale Simulation in Materials Science. MRS Bulletin, Feb. 1998. 28
- 8 Davison Lee. Continuum Modeling, MRS Bulletin. Feb. 1988. 16
- 9 Quantum Theory of Real Materials. Chelikowsky J R and Louie S G ed. Kluwer Academic Publishers. 1996
- 10 Kohn W. Density Functional Theory: Fundamentals and Applications. In: Bassani F, Fumi F and Tosi M P ed. Highlights of Condensed Matter Theory. North Holland. 1985. 1
- 11 Sham L. Density Functional Theory and Computational Materials Physics. In: Chelikowsky J R and Louie S G ed. Quantum Theory of Real Materials. Kluwer Academic Publishers. 1996. 13
- 12 Hybertsen M S and Louie S G Phys. Rev. B. 1985, 34: 5390
- 13 Louie S. G. Quasiparticle Theory of Electron Excitations in Solid, in: Chelikowsky J. R. and Louie S G ed. Quantum Theory of Real Materials, Kluwer Academic Publishers. 1996. 83
- 14 Car R and Parrinello M. Phys. Rev. Lett. 1985, 55: 2471

- 15 Car R. Modeling Materials by ab initio Molecular Dynamics, in: Chelikowsky J R and Louie S.G ed. Quantum Theory of Real Materials. Kluwer Academic Publishers, 1996. 23
- 16 McDowell D. L. and Story T. L. . New Directions in Material Design Science and Engineering. National Science Foundation Workshop. Report, Georgia Institute of Technology Mater. Council, Atlanta. GA. 1999
- 17 熊家炯主编. 材料设计. 天津: 天津大学出版社, 2000

第 24 章 计算相图及其进展

撰稿人 金展鹏 刘立斌

相图计算就是运用热力学原理计算系统的相平衡关系并绘制出相图的科学研究。相图计算的关键就是选择合适的热力学模型模拟各相的热力学性质随温度、压力、成分等的变化。

所谓模拟，就是通过确定一些假设来确定模型，然后利用这些模型来计算系统的性质的过程。通常我们所说的模拟分为两种，一种情况下，模型能用完整的解析表达式来表示系统的性质随实际条件改变所产生的变化，这种情况，我们称为“modeling”，相图计算就是其中一种；另一种情况下，模型不能用完整的解析表达式，但是我们可以用一些假设来进行数值迭代，当这种迭代在某种程度上相似于真实物理体系的性质时，我们称为“simulation”，如 Monte-Carlo 模拟。随着电脑的功能越来越强大，模拟的功能也越来越强大。

模拟的目的有两个，从科学的角度来讲，模拟可以帮我们了解自然是怎样运作的，我们获取知识的一种手段就是通过假设来确定模型，然后通过比较模型的预测结果与实测结果来检测假设与模型的正确性。我们并不在乎这种预测是解析计算还是通过数值迭代来获得的。从技术的角度来讲，模拟能帮助我们预测实际体系的性质以便我们控制与优化某一工艺过程或预测与延长产品的寿命。

通常情况下，模型能用解析式表示的话，运用起来会更方便。最简单的情况下，模型来自于一些有用的数学表达式，有的表达式可能有物理意义，有的可能没有物理意义。但是实际经验表明，有坚实物理基础的模型比没有物理基础的经验模型通常更有用，运用这样的模型我们可以对实测范围以外的地方做出恰当的预测。

1908 年 J·J·范拉 (Van Laar) 就曾作过计算相图的尝试^[1]。1957 年 J·L·梅杰林 (Meijering) 应用规则溶液模型的边界二元系，成功地外推计算了 Ni-Cr-Cu 三元系在 1203K 下的等温截面^[2]。20 世纪 60 年代，L·考夫曼 (Kaufman) 及其合作者发表了一系列的计算二元相图与三元相图的文章，1970 年考夫曼和 H·伯恩斯坦 (Bernstein) 出版的《相图计算机计算》一书，给出了由高熔点金属组成的 72 个合金二元系和 286 个合金三元系的热

力学参数,并给出了计算图例,为相图计算机计算提供了基本思路^[3]。1973年成立了相图计算国际组织“相图计算”(CALPHAD),该组织1977年开始出版专门刊物《CALPHAD——相图与热化学的计算机耦合》,每年召开一次国际会议。目前国际上已有许多成熟的相图计算软件,如:Thermal-Calc^[4]、Lukas Program^[5]等,建立了许多相图热力学数据库,如:SGTE 纯物质数据库、溶液数据库、钢铁数据库、铝基合金数据库、钛基合金数据库、镍基合金数据库、半导体材料数据库、氧化锆基结构陶瓷材料数据库与人工晶体材料数据库等。许多国际上著名的公司如:美国的通用电器公司、惠普公司;英国劳斯莱斯公司;日本住友金属都开发了用于自己材料设计的相图热力学数据库。

运用 CALPHAD 方法进行相图计算主要分为 4 个步骤。首先,根据体系中各相的结构特点,来一一选择合适的热力学模型描述其 Gibbs (吉布斯)自由能,这些与温度、压力和浓度有关的自由能表达式中含有一定数量的待定参数;然后,利用实测的相图与热力学数据优化出 Gibbs 自由能表达式中的待定参数;再采用适当的算法和相应的计算机程序在计算机上按相平衡条件计算相图;最后将计算结果与实验数据比较和分析,如果两者相差较大,则调整待定参数或重新选择热力学模型,再进行一次优化计算,直至计算结果与大部分相图数据与热力学数据在实验误差范围内吻合。

实验测定相图过程中需要耗费大量的人力与物力,在高温、高压、有腐蚀性气体参与反应的条件下,还将面临成分控制、容器选择和高温测量等方面的困难,而且实验测定总是有限的、片面的,无法对体系的相图和热力学性质作一个完整、全面的了解。引入相图计算后,只需要对体系中相图的部分关键区域和某些关键相的热力学数据进行实验测量就可以优化出 Gibbs 自由能模型参数,外推计算出整个相图,建立起该体系完整的相图热力学数据库,从而大大减少了相图研究的工作量,有可能避开可能出现的实验困难。与实测相图相比,计算相图有以下显著特点。

① 可以用来判别实测相图数据和热力学数据本身及它们之间的一致性,从而对来自不同作者和运用不同实验方法所获得的实验结果进行合理的评价,为使用者提供准确可靠的相图信息。

② 可以外推和预测相图的亚稳部分,从而得到亚稳相图。

③ 可以外推和预测多元相图,计算多元相平衡,为实际材料设计与加工工艺的制订作参考。

④ 通过计算等 Gibbs 自由能曲线 (T_0 线),可以预测无扩散相变的成分范围。

⑤ 可以提供相变动力学研究所需的相变驱动力、活度等重要信息。

⑥ 可以方便地获得以不同热力学变量为坐标的各种相图形式, 以便用于不同条件下的材料制备与使用过程的研究与控制。

24.1 相图计算的基本原理——Gibbs 法则

为了进行相图计算必须选取一个热力学状态函数作为热力学模型的模型函数。由于实际相图与热力学测量数据恒温恒压数据比较多, 因此一般选 Gibbs 自由能作为模型函数。如果我们知道了 Gibbs 自由能, 根据以下关系式可以求出其他热力学量:

$$S = -(\partial G / \partial T)_p;$$

$$H = G + TS = G - T(\partial G / \partial T)_p;$$

$$V = (\partial G / \partial p)_T;$$

$$F = G - pV = G - p(\partial G / \partial p)_T;$$

$$U = G + TS - pV = G - T(\partial G / \partial T)_p - p(\partial G / \partial p)_T;$$

$$C_p = -T(\partial^2 G / \partial T^2)_p;$$

$$\alpha = -T(\partial^2 G / \partial T \partial p);$$

$$C_p = -\frac{1}{V}(\partial^2 G / \partial p^2)_T;$$

$$\mu_i = -(\partial G / \partial n_i)_{p, T, n_j};$$

$$a_i = \exp\left(\frac{\mu_i - \mu_i^\circ}{RT}\right)。$$

而根据 Gibbs 法则: 对于物质一定但与外界有能量交换的体系 (封闭体系), 恒温恒压过程总是朝 Gibbs 自由能降低的方向进行, 平衡状态下体系总的 Gibbs 自由能最低, 每一组元在各相中的化学位相等。如果我们知道在所有温度下自由能-成分曲线, 通过求自由能最小或解化学位相等方程我们就可以计算出相图 (图 24-1)。Gibbs 自由能成为了联系相图与各种热力学数据的桥梁。在相图计算软件中, 为了数值计算的需要, 程序设计者又对这两种法则进行了一些变形处理^[6,7]。为了利用某些温度与成分下的实测值来确定所有温度下的自由能-成分曲线, 我们必须采用 Gibbs 自由能模型。

Gibbs 自由能模型就是 Gibbs 自由能随温度、压力与成分变化的规律。为合理地描述体系中各种不同类型的相的 Gibbs 自由能与温度、压力与成分的关系, 经过许多学者的摸索与总结提出了一系列热力学模型。

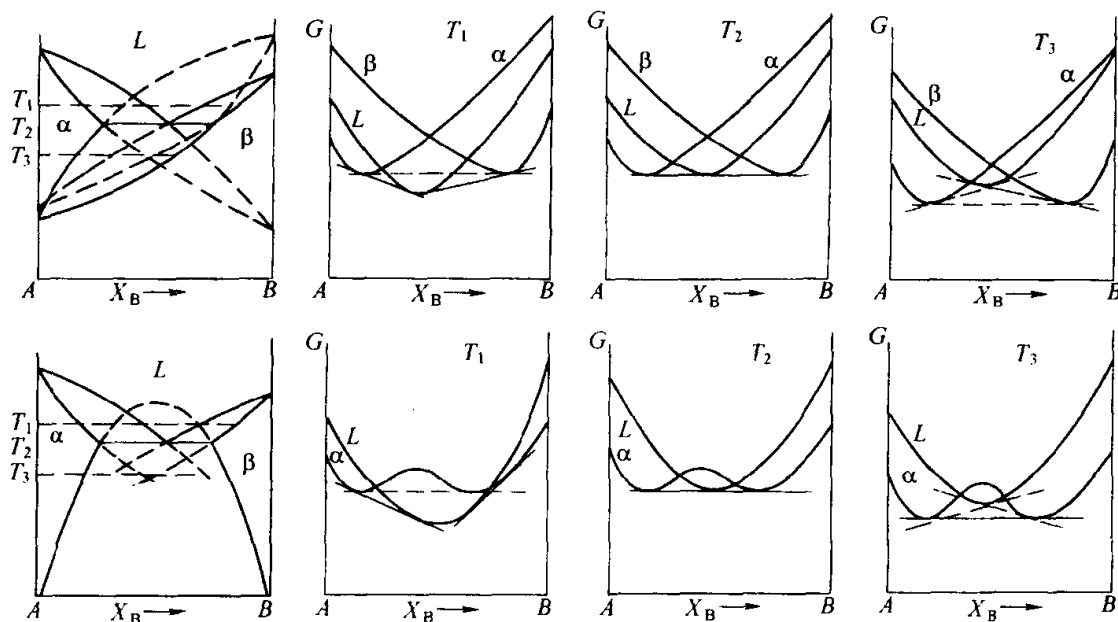


图 24-1 由自由能曲线求两端际固溶体结构
相同与结构不同两种情况下的共晶相图

24.2 纯物质的自由能——点阵稳定性常数

纯物质的自由能只与温度和压力有关而与成分无关。点阵稳定性常数就是纯物质两种可能组态的自由能差。自由能没有绝对值，影响两相平衡相图形状的是两组态的自由能差，而不是它的绝对值。之所以讲可能组态，是因为在相图计算时，我们不仅要计算稳定平衡，而且要计算亚稳平衡。

24.2.1 温度关系模型

根据描述热容的幂级数表达式：

$$C_p = a + bT + cT^{-2} + dT^2$$

描述 Gibbs 自由能与温度的关系最简便的方法是用幂级数组合：

$$G = A + BT + CT \ln T + DT^2 + ET^3 + FT^{-1}$$

其中温度为模型变量， A 、 B 、 C 、 D 、 E 、 F 为模型参数。事实上，上式往往只能适应于一个有限的温度区间，为了避免在 Gibbs 自由能的温度表达式中不得不增加参数的个数，通常我们采用划分温度区间的办法，即对各个温度区间，热容与 Gibbs 自由能表达式的形式相同但是值不同。温度区间的划分必须有理论依据，如相变点等，不能随意划分。通常我们只对 298K 以上的温度感兴趣，因此将此时稳定元素 0.1MPa 下的焓假定为零，用作各种能量数据的参考态。利用这一参考态，欧洲热力学研究组 (Scien-

tific Group of Thermal Europe) Dinsdale 等优化计算了所有纯元素的自由能-温度表达式, 建立了相应的点阵稳定性数据库^[8]。

24.2.2 压力关系模型

压力的影响像体积、热膨胀等性质的影响一样, 在热力学模型中常常被忽略, 仅在压力极高的情况下才予以考虑。对气相, 除接近临界点如沸点, 一般用 $RT \ln p/p^\circ$ 一项就足以描述压力的贡献了, 其中 p° 为纯物质在 298K 的蒸气压, 一般凝聚态材料体系通常不考虑压力对 Gibbs 自由能的贡献。

24.2.3 纯组元相图的计算

有些纯元素或化学计量比的化合物, 在不同的温度与压力下能以不同的形式存在, 如果我们可以找到其高温高压的实测相图与热力学数据, 这时就可以用 p - T 图来确定其晶格稳定性的温度与压力表达式, 反过来也可以用晶格稳定性常数的温度与压力表达式来计算 p - T 图。

以钛、锆为例, 常温常压下, 钛和锆都是密排六方结构 (ϵ), 高温下钛与锆都是体心立方结构 (β), 高压下钛与锆是一种密度比密排六方结构更密

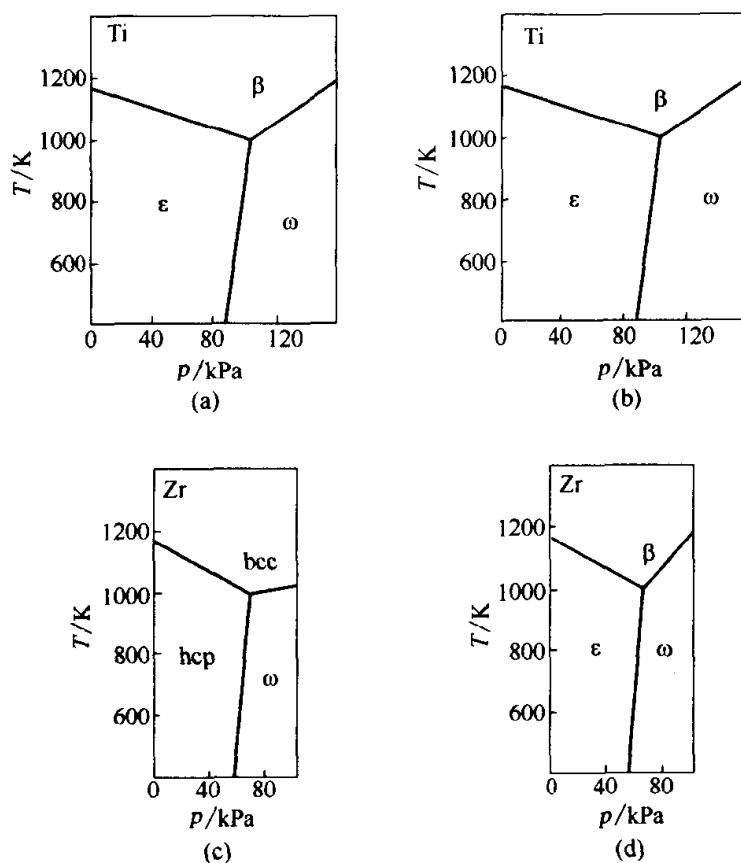


图 24-2 钛与锆计算相图与实测相图的比较

- (a) 钛的测量相图; (b) 钛的计算相图;
 (c) 锆的测量相图; (d) 锆的计算相图

的复杂立方结构 (ω)。根据钛、锆的 p - T 图以及:

$$\Delta G = \Delta H - T\Delta S = \Delta U + p\Delta V - T\Delta S$$

并假定 ΔU 、 ΔV 、 ΔS 为常数, 我们可以求得:

$$\Delta G_{\text{Ti}}^{\epsilon \rightarrow \beta} \cong 4393.2 - 3.81T - 115.68p(0.060)$$

$$\Delta G_{\text{Ti}}^{\omega \rightarrow \epsilon} \cong -1506.24 - 0.33T + 115.68p(0.190)$$

$$\Delta G_{\text{Ti}}^{\omega \rightarrow \beta} \cong 2886.96 - 4.14T + 115.68p(0.130)$$

$$\Delta G_{\text{Zr}}^{\epsilon \rightarrow \beta} \cong 4393.2 - 3.81T - 115.68p(0.096)$$

$$\Delta G_{\text{Zr}}^{\omega \rightarrow \epsilon} \cong -1506.24 - 0.33T + 115.68p(0.320)$$

$$\Delta G_{\text{Zr}}^{\omega \rightarrow \beta} \cong 2845.12 - 4.14T + 115.68p(0.224)$$

其中 ΔG 的单位为 J/mol, ΔS 为 J/(mol·K), 温度 T 为 K, ΔV 为 cm^3/mol , p 为 kPa。图 24-2 为利用上式计算相图与实测相图的比较。

24.3 热力学模型

24.3.1 二元化学计量比相的热力学模型

化学计量比相的自由能就等于纯物质的自由能加上化学计量比相的生成

自由能, 如图 24-3。那么化合物相 θ 的自由能表达式可写成:

$$G^\theta = X_A {}^\circ G_A^\alpha + X_B {}^\circ G_B^\beta + \Delta G^\theta$$

式中 A , B 原子的参考态分别为 α 、 β , 两者可以相同也可以不同, ${}^\circ G_A^\alpha$, ${}^\circ G_B^\beta$ 为纯 A 纯 B 分别在 α 态和 β 态的自由能, ΔG^θ 为 1mol 原子化合物的形成自由能。

24.3.2 替换溶液模型

替换溶液 (如液相与置换固溶体相) 的自由能一般通过对理想溶液进行修正来实现, 其自由能表达式可以

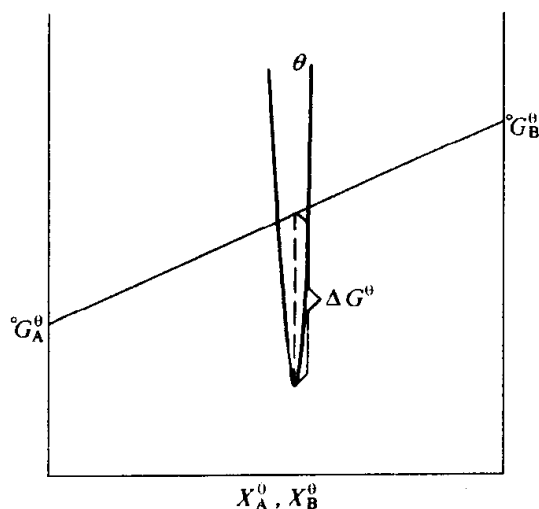


图 24-3 化合物的自由能

写成:

$$G_m^\alpha = X_A {}^\circ G_A^\alpha + X_B {}^\circ G_B^\beta + RT[X_A \ln(X_A) + X_B \ln(X_B)] + G^{\text{ex}}$$

式中, $X_A {}^\circ G_A^\alpha + X_B {}^\circ G_B^\beta$, 即 X_A 摩尔 α 态的 A 与 X_B 摩尔 α 态的 B 的机械混合的自由能 (注意与化合物不同, 对于溶液相纯物质必须用同一个参考态), $RT[X_A \ln(X_A) + X_B \ln(X_B)]$ 为理想混合熵引起的自由能增量, 两者之和为理想溶液的自由能。在理想溶液中 A - A 键的键能等于 B - B 键的键能也

等于A—B键的键能，A原子与B原子随机混合，混合热为零。 G^{ex} 为超额自由能，表示溶液偏离理想溶液的程度。对于替换溶液，超额自由能与成分的关系一般可以用多项式来表示，如：Redlich-Kister多项式^[9]：

$$G^{\text{ex}} = X_A X_B \sum I_i (X_A - X_B)^i$$

I_i 称为相互作用参数。当 $i=0$ 时上式变为

$$G^{\text{ex}} = X_A X_B I_0$$

称为规则溶液模型^[10]。在规则溶液中，A—A键、B—B键、A—B键的键能不相等，溶液中原子仍然是随机混合。设 ν 为破坏A—A、B—B键形成一个A—B键或B—A键的键能（即 $2\nu = 2u_{\text{A-B}} - u_{\text{A-A}} - u_{\text{B-B}}$ ，其中 $u_{\text{A-B}}$ 、 $u_{\text{A-A}}$ 、 $u_{\text{B-B}}$ 分别为A—B、A—A、B—B键的键能）， N 为总原子数，当物质量为 1 mol 时， N 即阿佛加德罗常数， Z 为配位数，在随机混合的条件下，形成A—B键与B—A键的总数为 $NZX_A X_B$ ，因此，由于键能不同而导致的自由能增量，即超额自由能为 $\nu NZX_A X_B$ ， $I_0 = \nu NZ$ 。

当 $\nu > 0$ ， $I_0 > 0$ 时， $G^{\text{ex}} > 0$ ，在自由能-成分曲线为马鞍形，有两个最低点，溶体相将产生溶解度间隙，溶体相内有偏聚（图 24-4）。可以证明溶解度间隙的顶点温度 $T_c = \nu Z / 2k_B$ ，其中 k_B 为玻耳兹曼常数。

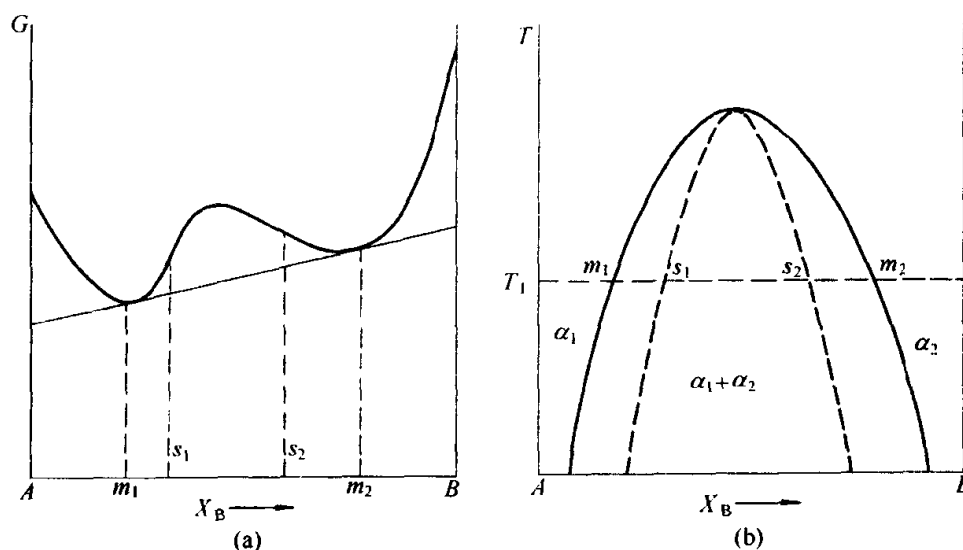


图 24-4 溶解度间隙的形成

(a) 自由能曲线；(b) 溶解度间隙（虚线为拐点曲线）

当 $\nu < 0$ ， $I_0 < 0$ 时， $G^{\text{ex}} < 0$ ，自由能-成分曲线只有一个最低点，溶体相内出现短程有序。

当 $\nu = 0$ ， $I_0 = 0$ 时， $G^{\text{ex}} = 0$ ，即理想溶液。

二元系中，符合规则溶液模型的固液两相相互作用参数的变化对相图的

影响见图24-5，由图可见，随相互作用参数的变化，相图可以产生巨大的变化。

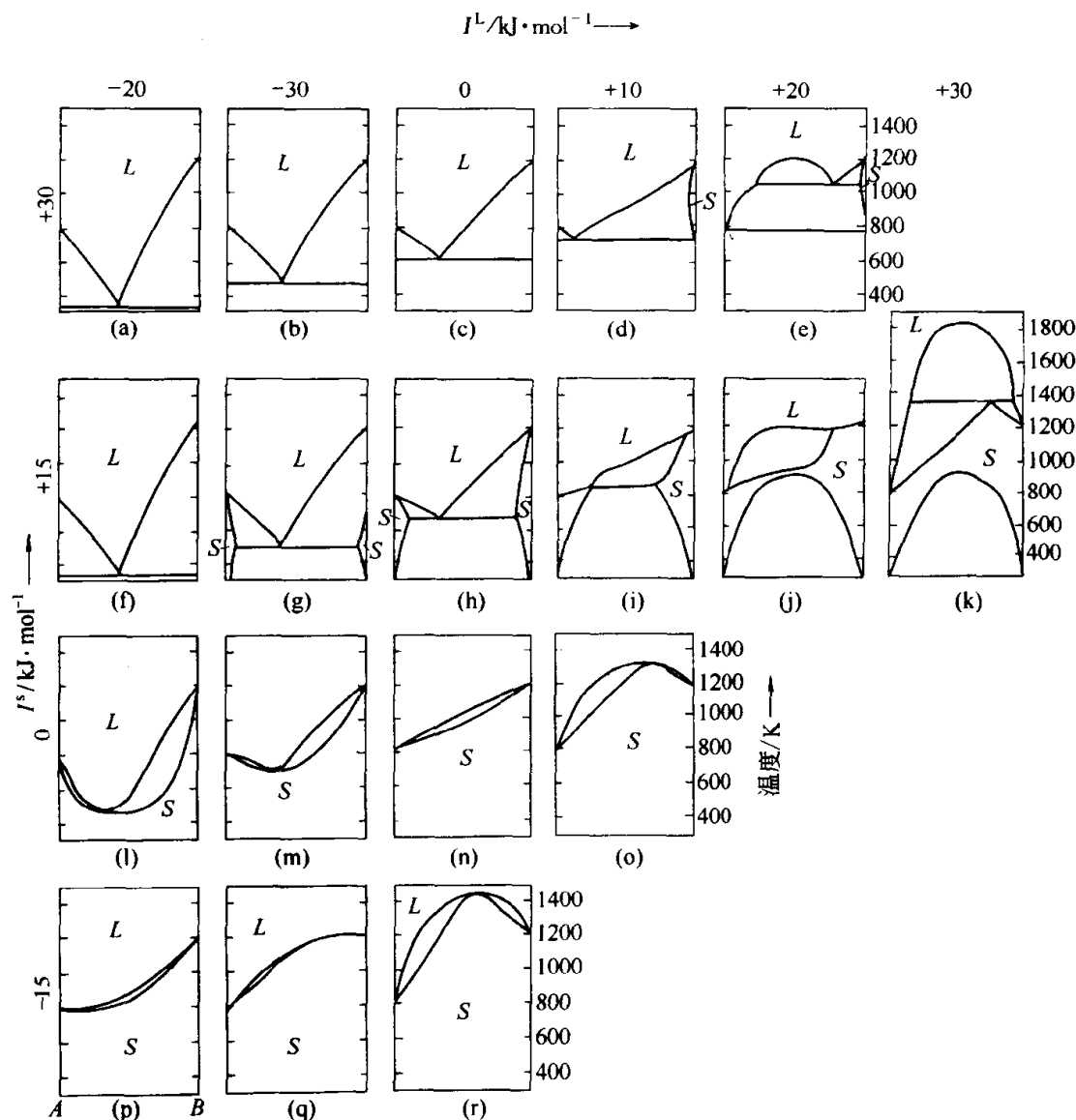


图 24-5 两规则溶液相互作用参数对相图的影响^[11]

只要 $\nu \neq 0$ ，溶体相内原子就不是随机分布的，因此规则溶液模型只是一个近似模型，为了描述非随机性，也就是原子环境对成键的影响，我们可以用次规则溶液模型 ($i=1$)^[12]，甚至次次规则溶液模型 ($i=2$)。次规则溶液模型的超额自由能表达式为：

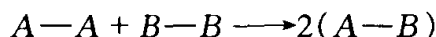
$$G^{\text{ex}} = X_A X_B [I_0 + I_1 (X_A - X_B)]$$

次次规则溶液模型的超额自由能表达式为：

$$G^{\text{ex}} = X_A X_B [I_0 + I_1 (X_A - X_B) + I_2 (X_A - X_B)^2]$$

24.3.3 准化学模型

如前所述, 在 $\nu < 0$ 的体系中, 将出现短程有序, 这种短程有序将对混合熵产生一定的影响, 为了更准确地描述这种变化, Bente 提出了准化学模型^[13]。在这一模型中, 假设 A 原子与 B 原子形成的键是相互独立并随机分布的, 键与键之间通过下面的化学反应式相连:



在平衡态有:

$$\frac{X_{AB}X_{BA}}{X_{AA}X_{BB}} = \exp(-2\nu/kT)$$

根据反应式可知 $X_{AB} = X_{BA}$, 对 1mol 原子, 混合键数为 $\frac{1}{2}NZ$ mol, 混合自由能变化为:

$$\begin{aligned} {}^M G = & \nu NZ X_{AB} + \frac{1}{2} NZ k T [X_{AA} \ln(X_{AA}) \\ & + X_{BB} \ln(X_{BB}) + X_{AB} \ln(X_{AB}) + X_{BA} \ln(X_{BA})] \end{aligned}$$

混合熵为:

$${}^M S = -\frac{1}{2} NZ k [X_{AA} \ln(X_{AA}) + X_{BB} \ln(X_{BB}) + X_{AB} \ln(X_{AB}) + X_{BA} \ln(X_{BA})]$$

但是, 事实上, 溶液中的无序度不会如此之大, 结合键不是相互独立并随机分布的, 全部 Z 个键都是有一定的指向, 它们指向某一位置必须与那里的原子类型相一致。我们可以增添一个负的修正项来校正前面的表达式。对 ν 比较小的体系可以用理想溶液来修正。在理想溶液中, 混合熵为理想混合熵, $\nu = 0$, 原子随机混合, $X_{AA} = X_A^2$, $X_{BB} = X_B^2$, $X_{AB} = X_A X_B = X_{BA}$, 有:

$$\begin{aligned} {}^M S = & -\frac{1}{2} NZ k [X_A^2 \ln(X_A^2) + X_B^2 \ln(X_B^2) + 2X_A X_B \ln(X_A X_B)] + \text{修正项} \\ = & -Nk [X_A \ln(X_A) + X_B \ln(X_B)] \\ \text{修正项} = & (Z-1) Nk [X_A \ln(X_A) + X_B \ln(X_B)] \end{aligned}$$

修正后, 混合自由能为:

$$\begin{aligned} {}^M G = & \nu NZ X_{AB} + \frac{1}{2} NZ k T [X_{AA} \ln(X_{AA}) + X_{BB} \ln(X_{BB}) + X_{AB} \ln(X_{AB}) \\ & + X_{BA} \ln(X_{BA})] - (Z-1) Nk T [X_A \ln(X_A) + X_B \ln(X_B)] \end{aligned}$$

由于修正项来自理想溶液, 对有序度比较大的体系, 上述表达式不成立。不过这时准化学模型也已不成立, 溶液相中有可能出现更大的集体。

24.3.4 缔合物模型

在有些情况下, 在某一成分处, 相界线、焓与熵与成分曲线将出现尖点

(图 24-6)^[14,15]。这时我们可以假设在相中存在某些固定成分的原子集团(缔合物), 如 Ag_2Te , 溶液是由缔合物与端际组元共同构成的, 组元之间及

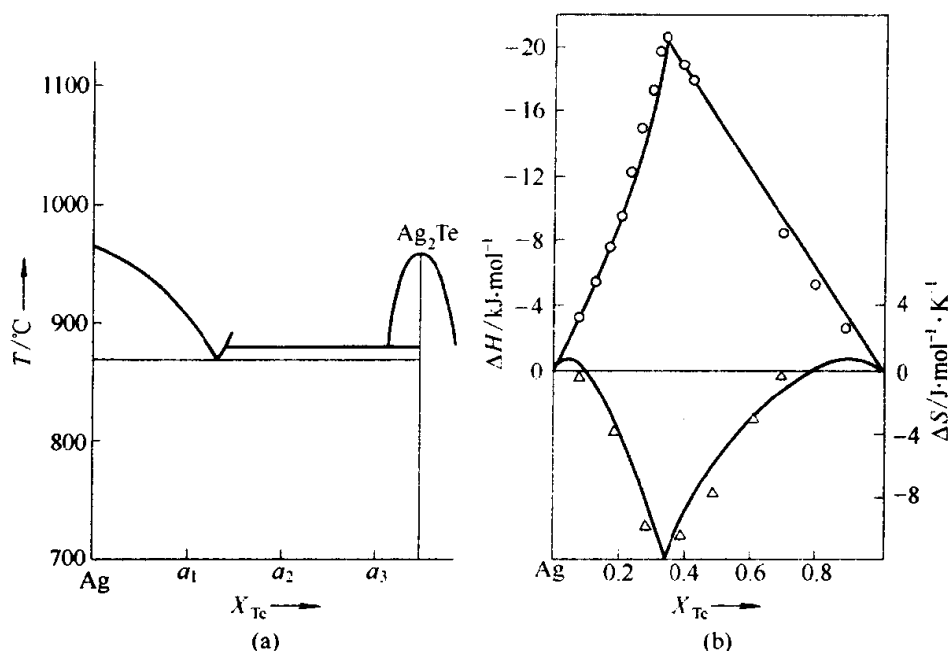


图 24-6 在 Ag-Te 二元相图及其焓与熵-成分曲线

(a) Ag-Te 二元相图; (b) 焓、熵-成分曲线

组元与缔合物之间的相互作用符合规则溶液模型。例如: 我们假设 $A-B$ 二元系中存在 A_iB_j 这样一个缔合物, $A-B$ 、 A_iB_j 在相中所占的摩尔分数分别为 y_A 、 y_B 、 $y_{A_iB_j}$, $y_A + y_B + y_{A_iB_j} = 1$, 那么有:

$$X_A = (y_A + i y_{A_iB_j}) / [1 + (i + j - 1) y_{A_iB_j}] X_B$$

$$= (y_B + j y_{A_iB_j}) / [1 + (i + j - 1) y_{A_iB_j}]$$

该相的自由能可以表示成:

$$G_m^a = y_A^a G_A^a + y_B^a G_B^a + y_{A_iB_j}^a G_{A_iB_j}^a + RT [y_A \ln(y_A) + y_B \ln(y_B) + y_{A_iB_j} \ln(y_{A_iB_j})] + I_{A:A_iB_j} y_A y_{A_iB_j} + I_{B:A_iB_j} y_B y_{A_iB_j} + I_{A:B} y_A y_B$$

其中:

$$^a G_{A_iB_j}^a = i^a G_A^a + j^a G_B^a + \Delta G_{A_iB_j}$$

为 1mol 分子缔合物的自由能。

$I_{A:A_iB_j}$, $I_{B:A_iB_j}$, $I_{A:B}$ 分别为溶液中 A 组元与 A_iB_j 、 B 组元与 A_iB_j 及 A 组元与 B 组元的相互作用参数。

这就是缔合物模型, 当然一个溶液相中可能不止一种缔合物, 缔合物与端际、组元及端际组元之间的相互作用可以假设成次规则溶液模型甚至次次规则溶液模型, 但是总的思路不变。

24.3.5 团簇变分法 (CVM 法)^[16]

在这种模型中, 我们不仅像准化学模型中一样考虑双原子键, 也考虑三原子、四原子、五原子甚至更多原子的原子团簇 (cluster)。与缔合物模型一样, 每一种原子团簇都有各自的形成自由能, 原子团簇的形成自由能由团簇中 AB 键的数量来估计。与缔合物模型不同, 考虑到团簇共面、共边、共角的事实, 团簇变分法中对熵进行了修正。团簇变分法是目前处理有序化最成功的方法。

例如^[17]: 对于 A 、 B 两种原子形成的面心立方固溶体, 我们可以假设其 A 、 $A_{0.75}B_{0.25}$ 、 $A_{0.5}B_{0.5}$ 、 $A_{0.25}B_{0.75}$ 、 B 5 种原子团簇, 每一原子团簇都由 4 个原子构成, 形成 1 个正四面体, 那么理想混合熵为:

$$-R[y_A \ln(y_A) + y_{A_{0.75}B_{0.25}} \ln(y_{A_{0.75}B_{0.25}}) + y_{A_{0.5}B_{0.5}} \ln(y_{A_{0.5}B_{0.5}}) + (y_{A_{0.25}B_{0.75}}) \ln(y_{A_{0.25}B_{0.75}}) + y_B \ln(y_B)]$$

考虑到包含一个 A 原子或一个 B 原子的 $A_{0.25}B_{0.75}$ 与 $A_{0.75}B_{0.25}$ 有 4 种不同的原子排列, A 原子或 B 原子在每一顶角的几率只有 $y_{A_{0.25}B_{0.75}}$ 或 $y_{A_{0.75}B_{0.25}}$ 值的 $1/4$, 包含两个 A 原子与两个 B 原子的 $A_{0.5}B_{0.5}$ 有 6 种不同的原子排列, 每种排列的几率只有 $y_{A_{0.5}B_{0.5}}$ 的 $1/6$ 。又

$$m(y/m) \ln(y/m) = y \ln(y) - y \ln(m)$$

我们可以引入一个熵的修正项:

$$R[(y_{A_{0.75}B_{0.25}} + y_{A_{0.25}B_{0.75}}) \ln(4) + y_{A_{0.5}B_{0.5}} \ln(6)]$$

再考虑到团簇共面、共边、共角的事实, 并进行近似处理, 团簇变分法所得到的最后熵的关系式为:

$$\begin{aligned} S = & -2R[y_A \ln(y_A) + y_{A_{0.75}B_{0.25}} \ln(y_{A_{0.75}B_{0.25}}) + y_{A_{0.5}B_{0.5}} \ln(y_{A_{0.5}B_{0.5}}) \\ & + y_{A_{0.25}B_{0.75}} \ln(y_{A_{0.25}B_{0.75}}) + y_B \ln(y_B) - (y_{A_{0.75}B_{0.25}} + y_{A_{0.25}B_{0.75}}) \ln(4) \\ & - y_{A_{0.5}B_{0.5}} \ln(6)] + 6R[p_{AA} \ln(p_{AA}) + p_{AB} \ln(p_{AB}) + (p_{BA}) \ln(p_{BA}) \\ & + p_{BB} \ln(p_{BB})] - 5R[X_A \ln(X_A) + X_B \ln(X_B)] \end{aligned}$$

可以证明, 在假设原子随机混合的情况下, 缔合物模型的熵不等于规则溶液模型的熵, 而 CVM 模型的熵却等于规则溶液模型的熵。

24.3.6 亚点阵模型^[18]

对于有一定成分范围的化合物相, 我们既可以认为是在化学计量比处存在缔合物, 这些相的溶解度是由端际组元与缔合物相互混合而形成的; 也可以认为在化合物相中 A 、 B 原子本来各自占据某一固定的位置 (亚点阵), 形成化学计量比相, 但是由于 A 、 B 原子互相串位而且 A 原子串到 B 原子位置的量不一定等于 B 原子串到 A 原子位置的量, 因而形成了溶解度。实

际晶体结构中也确实存在以某一种原子或某几种原子所占据的亚点阵。另外在间隙固溶体中，大原子占据点阵位置而间隙位置由小原子所占据，从而形成了点阵位置与间隙位置两种不同的位置，我们也认为是两个亚点阵。

如果我们假设：

① 每一亚点阵内的原子只与其他亚点阵内的原子相邻（这一点可以通过亚点阵的选取来保证）；

② 最近邻相互作用是常数；

③ 各亚点阵之间的相互作用忽略不计，过剩自由能是描述同一亚点阵内原子的相互作用，并与另一亚点阵内原子种类有关；

④ 亚点阵内原子遵循规则溶液模型。

则我们就可以根据实际相中亚点阵的形式，写出亚点阵模型的自由能表达式。

例如，如果我们假设亚点阵的形式为 $(A, B)_P(B, A)_Q$ ，在第一个亚点阵中以 A 原子为主，在第二个亚点阵中以 B 原子为主。A 在第一个亚点阵中的点阵分数为 y_A^1 、B 在第一个亚点阵中的点阵分数为 y_B^1 、A 在第二个亚点阵中的点阵分数为 y_A^2 、B 在第二个亚点阵中的点阵分数为 y_B^2 ， $y_A^1 + y_B^1 = 1$ ， $y_A^2 + y_B^2 = 1$ ，那么：

$$X_A = \frac{y_A^1 P + y_A^2 Q}{P + Q}$$

$$X_B = \frac{y_B^1 P + y_B^2 Q}{P + Q}$$

根据第①项与第②项假设，最近邻相互作用为：

$$y_A^1 y_A^2 \cdot G_{A_P A_Q} + y_A^1 y_B^2 \cdot G_{A_P B_Q} + y_B^1 y_A^2 \cdot G_{B_P A_Q} + y_B^1 y_B^2 \cdot G_{B_P B_Q}$$

根据第②项与第③项假设，混合熵与次近邻相互作用为：

$$PRT [y_A^1 \ln(y_A^1) + y_B^1 \ln(y_B^1)] + QRT [y_A^2 \ln(y_A^2) + y_B^2 \ln(y_B^2)]$$

$$+ P(I_A^1 y_A^1 y_B^1 y_A^2 + I_B^1 y_A^1 y_B^1 y_B^2) + Q(I_A^2 y_A^2 y_B^2 y_A^1 + I_B^2 y_A^2 y_B^2 y_B^1)$$

$(A, B)_P(B, A)_Q$ 的 Gibbs 自由能就等于这两项自由能之和。

其中 $G_{A_P A_Q}$ ， $G_{A_P B_Q}$ ， $G_{B_P A_Q}$ ， $G_{B_P B_Q}$ 为每一个亚点阵只含一个原子的情况下，4个假想化合物的自由能。

I_A^1 ， I_B^1 为第二个亚点阵分别为 A 或 B 的情况下，第一个亚点阵中 A，B 的相互作用参数。

I_A^2 ， I_B^2 为第一个亚点阵分别为 A 或 B 的情况下，第二个亚点阵中 A，B 的相互作用参数。

亚点阵模型可以广泛用于不同类型的相。

① 固定成分的化合物。对于固定成分的化合物，我们可以认为每一个亚点阵只包含一种原子，没有原子串位。 $y_A^1 = 1$, $y_A^2 = 0$, $y_B^1 = 0$, $y_B^2 = 1$

$$G_{A_P B_Q} = y_A^1 y_B^2 {}^\circ G_{A_P B_Q} = {}^\circ G_{A_P B_Q} = P {}^\circ G_A + Q {}^\circ G_B + \Delta G_{A_P B_Q}$$

其中 $\Delta G_{A_P B_Q}$ 为 $A_P B_Q$ 相的形成自由能。

② 规则溶液。对于规则溶液，原子随机混合，两亚点阵完全相等， $(A, B)_P (B, A)_Q$ 相当于 $(A, B)_{(P+Q)}$, $y_A^1 = y_A^2 = X_A$, $y_B^1 = y_B^2 = X_B$ ，而且只考虑最近邻相互作用

$$\begin{aligned} G_{(A,B)_{(P+Q)}} &= X_A X_A {}^\circ G_{A_P A_Q} + X_A X_B {}^\circ G_{A_P B_Q} + X_B X_A {}^\circ G_{B_P A_Q} \\ &\quad + X_B X_B {}^\circ G_{B_P B_Q} + (P+Q)RT[X_A \ln(X_A) + X_B \ln(X_B)] \\ &= (P+Q)[X_A X_A {}^\circ G_A + X_A {}^\circ G_A + X_B {}^\circ G_B + X_B X_B {}^\circ G_B \\ &\quad + X_A X_B (\Delta G_{A_P B_Q} + \Delta G_{B_P A_Q})] \\ &\quad + (P+Q)RT[X_A \ln(X_A) + X_B \ln(X_B)] \\ &= (P+Q)(X_A {}^\circ G_A + X_B {}^\circ G_B) + (\Delta G_{A_P B_Q} + \Delta G_{B_P A_Q})X_A X_B \\ &\quad + (P+Q)RT[X_A \ln(X_A) + X_B \ln(X_B)] \end{aligned}$$

其中 $\Delta G_{A_P B_Q} + \Delta G_{B_P A_Q}$ 即形成 $A-B$ 键导致的自由能增量，即相互作用参数。这样亚点阵模型就回到了规则溶液模型。当 $\Delta G_{A_P B_Q} + \Delta G_{B_P A_Q}$ 为零时，即理想溶液模型。

③ 间隙固溶体^[19]。常见的间隙固溶体例子是钢铁。在钢铁中，铁与合金元素占据点阵位置，碳、氮与其他小原子占据间隙位置。我们可以假设铁与合金元素占据的点阵位置为一个亚点阵，而碳、氮与其他小原子占据间隙位置为另一个亚点阵，间隙位置不可能全部填满，我们可以认为亚点阵的其他部分由空位所占据，这样奥氏体的模型的形式为 $(Fe, Cr, Mn, Ni \cdots \cdots) (C, N, Va \cdots \cdots)$ ，其中 Va 即空位。根据前面的假设，我们就可以写出其 Gibbs 自由能表达式。在自由能模型中 ${}^\circ G_{FeVa}$ 即纯铁的奥氏体态的自由能。

④ 金属间化合物^[20]。常见的金属间化合物相 σ 相是由体心立方 A 元素 (Cr, Mo, V) 和面心立方的 B 元素 (Fe, Ni, Co) 构成，根据 σ 相的晶体结构，每个元胞有 30 个原子分布在 5 个亚点阵上，第一个亚点阵有两个原子位置，第二个亚点阵有 4 个原子位置，其余 3 个有 8 个原子位置，第一个亚点阵与第三个亚点阵几乎只由 B 原子占据，第二个亚点阵几乎由 A 原子所占据，剩余的两个亚点阵由 A, B 原子共同占据。因此亚点阵的形式为： $(B)_2 (A)_4 (B)_8 (A, B)_8 (A, B)_8$ ，在实际计算中，为了减少参数的个数一般简化为 $(B)_{10} (A)_4 (A, B)_{16}$ ，根据前面的假设，我们就可以写出其 Gibbs 自

由能表达式。注意一旦选定了某一相，如 σ 相的亚点阵形式，那么在所有系统中必须使用同一亚点阵形式，以保证数据库的整体性。

⑤ 非计量离子化合物^[21]。对离子化合物一般阳离子周围总是阴离子，阴离子周围总是阳离子，形成密堆结构，因此我们通常假使阳离子为一个点阵，阴离子为一个点阵，与间隙固溶体一样，在亚点阵中可以引入空位来描述阴阳离子缺位引起的成分变化。例如： CeO_{2-x} 的亚点阵形式为 $(\text{Ce}^{4+}, \text{Ce}^{3+})(\text{O}^{2-}, \text{Va})_2$ 。为了满足电中性条件，必须保证 $y_{\text{Ce}^{3+}} = 4y_{\text{Va}}$ 亚点阵模型中，最近邻相互作用能必须保持一定的关系。用于化合物的亚点阵模型又称为化合物能量模型 (compound energy model)。

⑥ 离子溶液^[22]。液相中没有固定的点阵位置，故一般情况下，不能用亚点阵模型，但是对于熔盐-离子溶液，由于阳离子总是被阴离子包围，阴离子总是被阳离子包围，我们可以假设阳离子为一个亚点阵，阴离子为一个亚点阵，使用亚点阵模型描述其自由能。为了使模型能描述从金属熔体到非金属熔体的整个成分范围，在阴离子亚点阵中引入带电的空位与中性原子。由于液相不存在固定的点阵结构，不存在固定的 P 、 Q 值，一般假设， P 、 Q 等于阴离子的平均电荷数与阳离子的平均电荷数，空位的价数等于零。如阴阳离子都是一价的 $A-B$ 二元熔体相的亚点阵形式可写为 $(A^+)_{\text{P}}(B^-, B^0, \text{Va}^-)_{\text{Q}}$ 。 $P = y_{\text{B}} + y_{\text{Va}}$ ； $Q = 1$ 。用于离子溶液的亚点阵模型又称为双亚点阵离子溶液模型 (two sublattice ionic solution model)。

⑦ 特殊物理现象的热力学模型。对于特殊的物理现象，我们通常单独模拟其热力学性质的变化，也就是说，将包含这种现象的相的自由能分成两部分，一部分是特殊的物理现象引起的自由能变化，另一部分为这个相没有这种物理现象的假想态的自由能。铁磁材料的磁有序化就是其中一个典型例子。

图 24-7 为 $\alpha\text{-Fe}$ 磁性对热容的贡献与磁熵，热容曲线的顶点或磁熵曲线

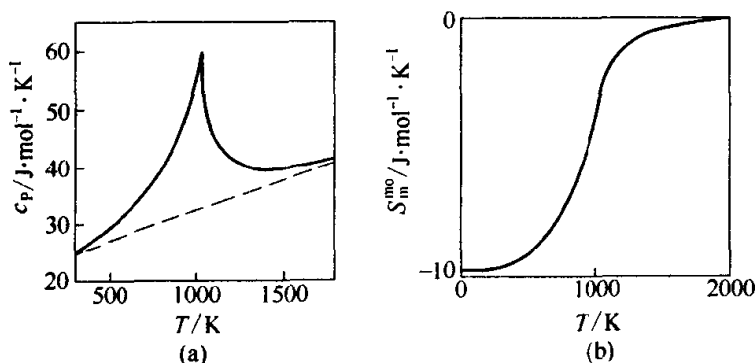


图 24-7 $\alpha\text{-Fe}$ 磁性对热容的贡献与磁熵 S_m^{mo}

(a) $\alpha\text{-Fe}$ 磁性对热容贡献；(b) 磁熵

的拐点即 $\alpha\text{-Fe}$ 由高温的顺磁状态转变为低温铁磁状态的居里点 T_c ，在顺磁状态下只存在短程磁有序，在铁磁状态下存在长程磁有序。

1977 年，Inden 率先推导出了磁性自由能与温度的关系的对数表达式^[23]，在此基础上，Hillert 等将对数表达式展开成幂级数并只取前三项推导出了磁性自由能随温度变化的表达式^[24]：

在居里温度以上

$$G_m^{\text{mo}} = -RT \ln(\beta + 1) \left(\frac{\tau^{-5}}{10} + \frac{\tau^{-15}}{315} + \frac{\tau^{-25}}{1500} \right) / \left[\frac{518}{1125} + \frac{11692}{15975} \left(\frac{1}{p} - 1 \right) \right]$$

在居里温度以下

$$G_m^{\text{mo}} = RT \ln(\beta + 1) \left\{ 1 - \left[\frac{79\tau^{-1}}{140p} - \frac{474}{497} \left(\frac{1}{p} - 1 \right) \left(\frac{\tau^3}{6} + \frac{\tau^9}{135} + \frac{\tau^{15}}{600} \right) \right] \right\} / \left[\frac{518}{1125} + \frac{11692}{15975} \left(\frac{1}{p} - 1 \right) \right]$$

其中 τ 等于 T/T_c ， p 为一个与晶体结构有关的常数，其物理意义为临界温度以上的磁焓占总磁焓的分数，对面心立方金属为 0.28，对体心立方金属为 0.4， β 为波耳磁子数。

T_c 与 β 都是成分的函数，对于 A、B 二元合金：

$$V = X_A^\circ V_A + X_B^\circ V_B + \sum_{i=0}^n W_{A,B}^i (X_A - X_B)$$

这里， V 即为 T_c 或 β ， W 为待优化参数。

⑧ 多元外推模型。实测的相图一般为二元相图或三元相图的一些截面，而实际材料通常由四五种以上的主要元素构成，因此，用低元系的相图与热力学数据外推计算多元相平衡对实际材料设计与工艺条件的优化很重要。与二元系相似多元系的机械混合自由能为： $X_A^\circ G_A + X_B^\circ G_B + X_C^\circ G_C + X_D^\circ G_D + \dots$ 。理想混合引起的自由能增量为：

$$RT [X_A \ln(X_A) + X_B \ln(X_B) + X_C \ln(X_C) + \dots]$$

关键是要确定超额自由能。多元系的超额自由能可以由低元系的超额自由能外推而得到，这样的外推公式有以下几种。

a. 科勒外推公式^[25]。只考虑二元超额自由能对多元超额自由能的贡献，科勒外推公式的通式为：

$$G^{\text{ex}} = \sum_{i=1}^N \sum_{j=i+1}^N (X_i + X_j)^2 G_{ij}^{\text{ex}} \left(\frac{X_i}{X_i + X_j}, \frac{X_j}{X_i + X_j} \right)$$

应用于 A-B-C 三元系，科勒外推公式可用图 24-8 (a) 来表示：

$$\begin{aligned}
G^{\text{ex}} = & (X_A + X_B)^2 G_{AB}^{\text{ex}} \left(\frac{X_A}{X_A + X_B}, \frac{X_B}{X_A + X_B} \right) \\
& + (X_B + X_C)^2 G_{BC}^{\text{ex}} \left(\frac{X_B}{X_B + X_C}, \frac{X_C}{X_B + X_C} \right) \\
& + (X_C + X_A)^2 G_{CA}^{\text{ex}} \left(\frac{X_C}{X_C + X_A}, \frac{X_A}{X_C + X_A} \right)
\end{aligned}$$

即三元系的过剩自由能等于二元边界上图示三个点的过剩自由能的加权平均值, 权重为 $(X_i + X_j)^2$ 。对于规则溶液:

$$\begin{aligned}
G^{\text{ex}} = & (X_A + X_B)^2 I_{AB} \frac{X_A}{X_A + X_B} \frac{X_B}{X_A + X_B} + (X_B + X_C)^2 I_{BC} \frac{X_B}{X_B + X_C} \frac{X_C}{X_B + X_C} \\
& + (X_C + X_A)^2 I_{CA} \frac{X_C}{X_C + X_A} \frac{X_A}{X_C + X_A} = I_{AB} X_A X_B + I_{BC} X_B X_C + I_{CA} X_C X_A
\end{aligned}$$

如果考虑三元超额自由能的贡献, 可以加入三元项, 科勒外推公式三元项的通式为:

$$\begin{aligned}
G^{\text{ex}} = & \sum_{i=1}^N \sum_{j=i+1}^N (X_i + X_j)^2 G_{ij}^{\text{ex}} \left(\frac{X_i}{X_i + X_j}, \frac{X_j}{X_i + X_j} \right) + \sum_{i=1}^N \sum_{j=i+1}^N \sum_{k=i+2}^N \\
& (X_i + X_j + X_k)^2 G_{ijk}^{\text{ex}} \left(\frac{X_i}{X_i + X_j + X_k}, \frac{X_j}{X_i + X_j + X_k}, \frac{X_k}{X_i + X_j + X_k} \right)
\end{aligned}$$

应用于 A-B-C 三元规则溶液:

$$G^{\text{ex}} = I_{AB} X_A X_B + I_{BC} X_B X_C + I_{CA} X_C X_A + I_{ABC} X_A X_B X_C$$

b. 科里内特公式^[26]。只考虑二元超额自由能对多元超额自由能的贡献, 科里内特公式的通式为:

$$G^{\text{ex}} = \frac{1}{2} \Sigma \left[\frac{X_2}{1 - X_1} G_{12}^{\text{ex}}(X_1, 1 - X_1) + \frac{X_1}{1 - X_2} G_{12}^{\text{ex}}(1 - X_2, X_2) \right]$$

其中 Σ 表示对所有二元系求和。

对 A-B-C 三元系, 科里内特公式可以用图 24-8 (b) 来表示,

$$\begin{aligned}
G^{\text{ex}} = & \frac{1}{2} \frac{X_B}{1 - X_A} G_{AB}^{\text{ex}}(X_A, 1 - X_A) + \frac{1}{2} \frac{X_A}{1 - X_B} G_{AB}^{\text{ex}}(1 - X_B, X_B) \\
& + \frac{1}{2} \frac{X_C}{1 - X_B} G_{BC}^{\text{ex}}(X_B, 1 - X_B) + \frac{1}{2} \frac{X_B}{1 - X_C} G_{BC}^{\text{ex}}(1 - X_C, X_C) \\
& + \frac{1}{2} \frac{X_A}{1 - X_C} G_{CA}^{\text{ex}}(X_C, 1 - X_C) + \frac{1}{2} \frac{X_C}{1 - X_A} G_{CA}^{\text{ex}}(1 - X_A, X_A)
\end{aligned}$$

即三元系的过剩自由能等于二元边界上图示 6 个点的过剩自由能的加权平均值。如果考虑三元超额自由能的贡献, 可以加入三元项, 科里内特外推公式中三元项的通式为:

$$G^{\text{ex}} = \frac{1}{2} \sum \left[\frac{X_2}{1-X_1} G_{12}^{\text{ex}}(X_1, 1-X_1) + \frac{X_1}{1-X_2} G_{12}^{\text{ex}}(1-X_2, X_2) \right] \\ + \frac{1}{3} \sum \left[\frac{X_3}{1-X_1-X_2} G_{123}^{\text{ex}}(X_1, X_2, 1-X_1-X_2) \right. \\ \left. + \frac{X_2}{1-X_1-X_3} G_{123}^{\text{ex}}(X_1, 1-X_1-X_3, X_3) \right. \\ \left. + \frac{X_1}{1-X_2-X_3} G_{123}^{\text{ex}}(1-X_2-X_3, X_2, X_3) \right]$$

其中第一个 \sum 表示对所有二元系求和, 第二个 \sum 表示对所有三元系求和。可以证明, 对于 A-B-C 三元规则溶液同样有:

$$G^{\text{ex}} = I_{AB} X_A X_B + I_{BC} X_B X_C + I_{CA} X_C X_A + I_{ABC} X_A X_B X_C$$

c. 姆加努方程^[27]。只考虑二元超额自由能对多元超额自由能的贡献, 姆加努方程的通式为:

$$G^{\text{ex}} = \sum \left[\frac{X_1 X_2}{V_{12} V_{21}} G_{12}^{\text{ex}}(V_{12}, V_{21}) \right] \\ V_{12} = X_1 + \frac{1-X_1-X_2}{2}, \quad V_{21} = X_2 + \frac{1-X_1-X_2}{2}$$

对 A-B-C 三元系, 该方程可用图 24-8(c) 来表示, 三元超额自由能可由三个垂直点的超额自由能外推而得到:

$$G^{\text{ex}} = \frac{X_A X_B}{V_{AB} V_{BA}} G_{AB}^{\text{ex}}(V_{AB}, V_{BA}) + \frac{X_B X_C}{V_{BC} V_{CB}} G_{BC}^{\text{ex}}(V_{BC}, V_{CB}) \\ + \frac{X_C X_A}{V_{CA} V_{AC}} G_{CA}^{\text{ex}}(V_{CA}, V_{AC})$$

其中 (V_{AB}, V_{BA}) 、 (V_{BC}, V_{CB}) 、 (V_{CA}, V_{AC}) 为图示 3 个垂直点的成分。

如果考虑三元相互过剩自由能的贡献, 则

$$G^{\text{ex}} = \sum \left[\frac{X_1 X_2}{V_{12} V_{21}} G_{12}^{\text{ex}}(V_{12}, V_{21}) \right] \\ + \sum \left[\frac{X_1 X_2 X_3}{V_{123} V_{231} V_{321}} G_{123}^{\text{ex}}(V_{123}, V_{231}, V_{321}) \right] \\ V_{123} = X_1 + \frac{1-X_1-X_2-X_3}{3}, \quad V_{231} = X_2 + \frac{1-X_1-X_2-X_3}{3},$$

$$V_{321} = X_3 + \frac{1 - X_1 - X_2 - X_3}{3}$$

对于 A-B-C 三元规则溶液同样有：

$$G^{\text{ex}} = I_{AB}X_A X_B + I_{BC}X_B X_C + I_{CA}X_C X_A + I_{ABC}X_A X_B X_C$$

以上三个外推公式均为对称外推，只适应于所有组元性质相近的情况。如果 B、C 两组元性质比较接近，而 A 的性质相差较远，那么 A-B、A-C 与 B-C 二元系的过剩自由能相距较远，在没有三元相图与热力学实测数据的情况下，用二元系的过剩自由能的热力学参数来外推计算三元相图就会带来比较大的误差。这时最好不用对称外推，而用下面的不对称外推方法。

d. 图普-希拉特方程^[28,29]。图普-希拉特方程根据最后一项的取法有 3 种形式。

① 最后一项取自科勒方程

$$G^{\text{ex}} = \frac{X_B}{1 - X_A} G_{AB}^{\text{ex}}(X_A, 1 - X_A) + \frac{X_C}{1 - X_A} G_{AB}^{\text{ex}}(X_A, 1 - X_A) \\ + (X_B + X_C)^2 G_{BC}^{\text{ex}}\left(\frac{X_B}{X_B + X_C}, \frac{X_C}{X_B + X_C}\right)$$

② 最后两项取自科里内特方程

$$G^{\text{ex}} = \frac{X_B}{1 - X_A} G_{AB}^{\text{ex}}(X_A, 1 - X_A) + \frac{X_C}{1 - X_A} G_{AB}^{\text{ex}}(X_A, 1 - X_A) \\ + \frac{X_C}{1 - X_B} G_{BC}^{\text{ex}}(X_B, 1 - X_B) + \frac{X_B}{1 - X_C} G_{BC}^{\text{ex}}(X_C, 1 - X_C)$$

③ 最后一项取自姆加努方程

$$G^{\text{ex}} = \frac{X_B}{1 - X_A} G_{AB}^{\text{ex}}(X_A, 1 - X_A) + \frac{X_C}{1 - X_A} G_{AB}^{\text{ex}}(X_A, 1 - X_A) \\ + \frac{X_B X_C}{V_{BC} V_{CB}} G_{BC}^{\text{ex}}(V_{BC}, V_{CB}) \\ V_{BC} = X_B + \frac{1 - X_B - X_C}{2}, V_{CB} = X_C + \frac{1 - X_B - X_C}{2}$$

3 种不对称外推示意图如图 24-9 所示。

不同的外推方法事实上是吉布斯-杜亥姆方程沿不同路径积分的结果。对规则溶液结果是一致的。对于有多元实测相图与热力学数据的体系，不同模型的差别可以通过多元参数的优化来弥补。因此为了简单起见，一般软件中只采用一种外推方法。

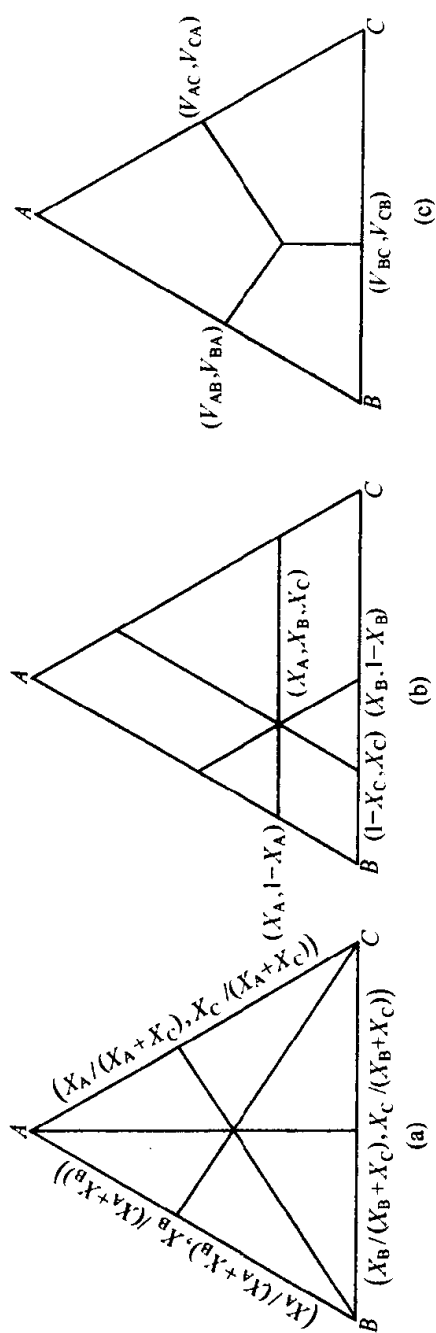


图 24-8 对称外推法示意图

(a) 科勒外推法; (b) 科里内特外推法; (c) 姆加努外推法

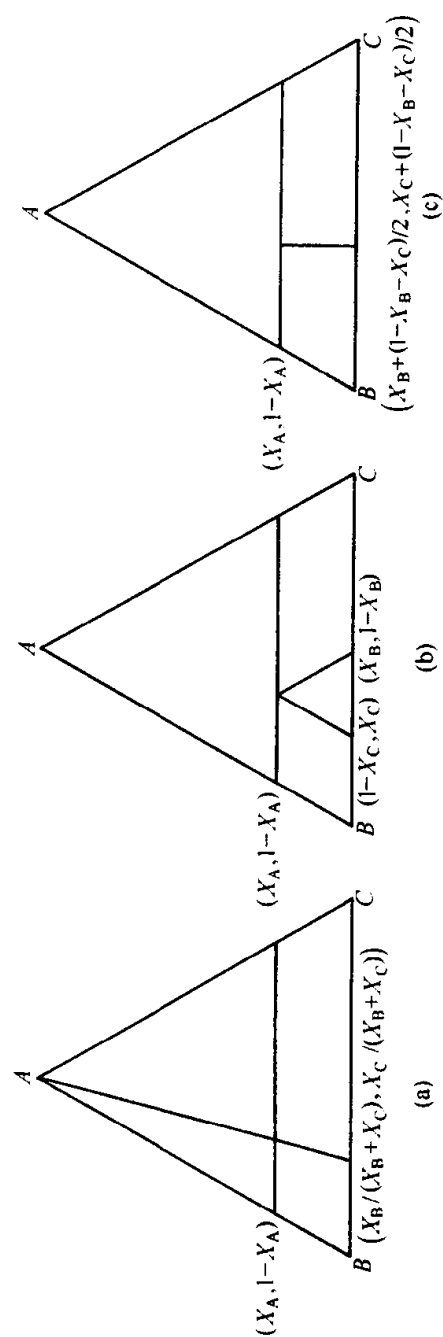


图 24-9 不对称外推示意图

(a) 最后一项取自科勒方程; (b) 最后一项取自科里内特方程; (c) 最后一项取自姆加努方程

24.4 相图优化和计算过程

在相图优化计算之前首先要进行实验数据的收集与评估,评估的目的主要根据作者所采用的实验方法判断实验数据的准确性,判断相互矛盾的实验数据哪一个更合理,作者对实验数据的解释有无违背相图与热力学原理。自由能模型的选用主要考虑相的结构,模型选择是否合理,可以通过外推计算多元系来验证。优化的过程是一个不断调整模型参数的过程,最后使得用所得参数计算的相图与热力学数据与实测数据之差小于某一临界值,参数的初值可以根据有关热力学原理来估算^[30]。一般先优化实验数据比较多的相的模型参数,然后再优化实验数据比较少的相的模型参数,最后整体优化。调整参数的方法即最优化方法。常见的有最小二乘法、牛顿法等,详情见有关《最优化原理》一书^[31]。优化后所有的相图与热力学数据通过热力学模型联结为一个自治的整体,最后以模型参数方式储存起来形成相图热力学数据库。

相图优化与计算的过程可以用图 24-10 来表示。

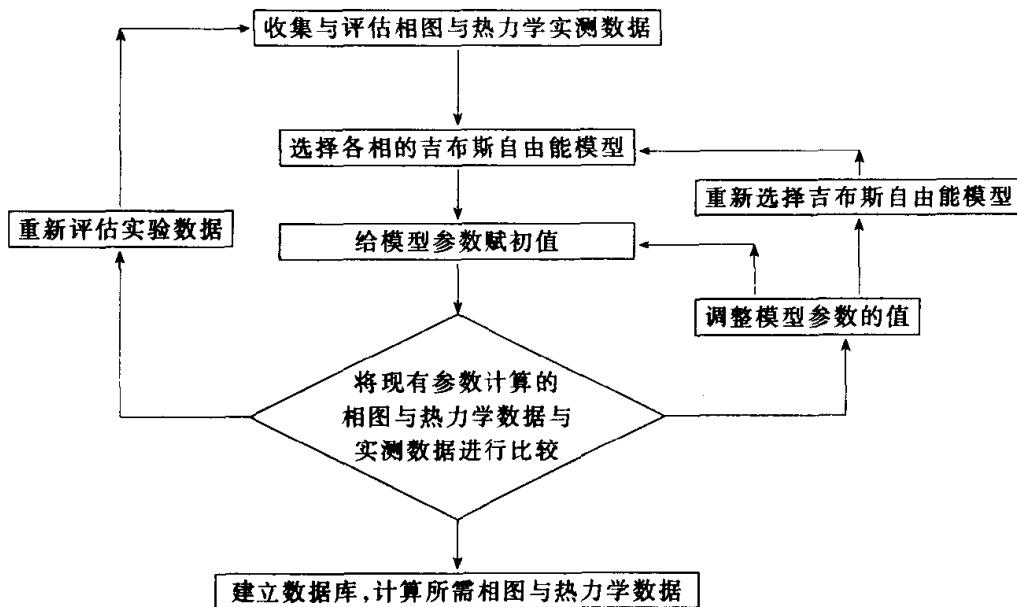


图 24-10 相图计算流程图

24.5 材料优化设计实例——双相不锈钢的成分设计^[32]

双相不锈钢中主要包含有 Fe-Cr-Ni-Mo-Mn-Si-C-N 等 8 种元素,利用已开发的相图热力学数据库, B.J.Lee 计算了成分为: 22Cr-5.5Ni-3Mo-1.7Mn-0.4Si-0.14N-0.02C (质量) 的 SAF2205 钢的奥氏体含量与温度的关系, 如图 24-11 所示。

奥氏体与铁素体中 Cr、Ni、Mo、Mn 含量与实验数据的比较, 如图 24-12

所示。计算结果与 F·H·海斯 (Hayes) 的实验结果^[32]符合得很好, 由此可以证明利用该相图数据库的可靠性。

用相图热力学数据库进行材料设计, 首先必须搞清材料的性能与结构的关系。通过查阅文献可知, 在设计双相不锈钢必须考虑的性能有强度、韧性、可焊接性与抗腐蚀性能。加工硬化与控制晶粒尺寸对于强度的影响比成分变化的影响大得多, 因此在成分优化设计时, 不考虑成分变化对强度的影响。对韧性影响最大的因素是 σ 相的析出; 对可焊接性能影响最大的是高温下 (从热处理温度到熔点) 奥氏体相的稳定性; 对双相不锈钢抗腐蚀性能影响最大的是由于 $M_{23}C_6$ 析出引起的奥氏体中 Cr 含量的降低引起的晶间腐蚀与点蚀。

用相图热力学数据库进行材料设计的第二步就是搞清材料的结构与成分

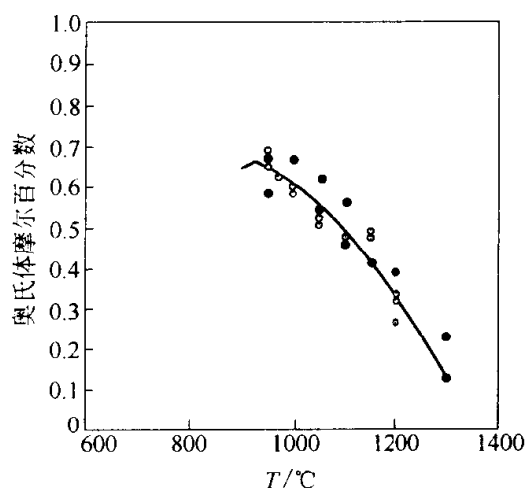


图 24-11 SAF2205 双相不锈钢中奥氏体的相含量与温度的关系
○—SEM-EDAX; ●—点计数

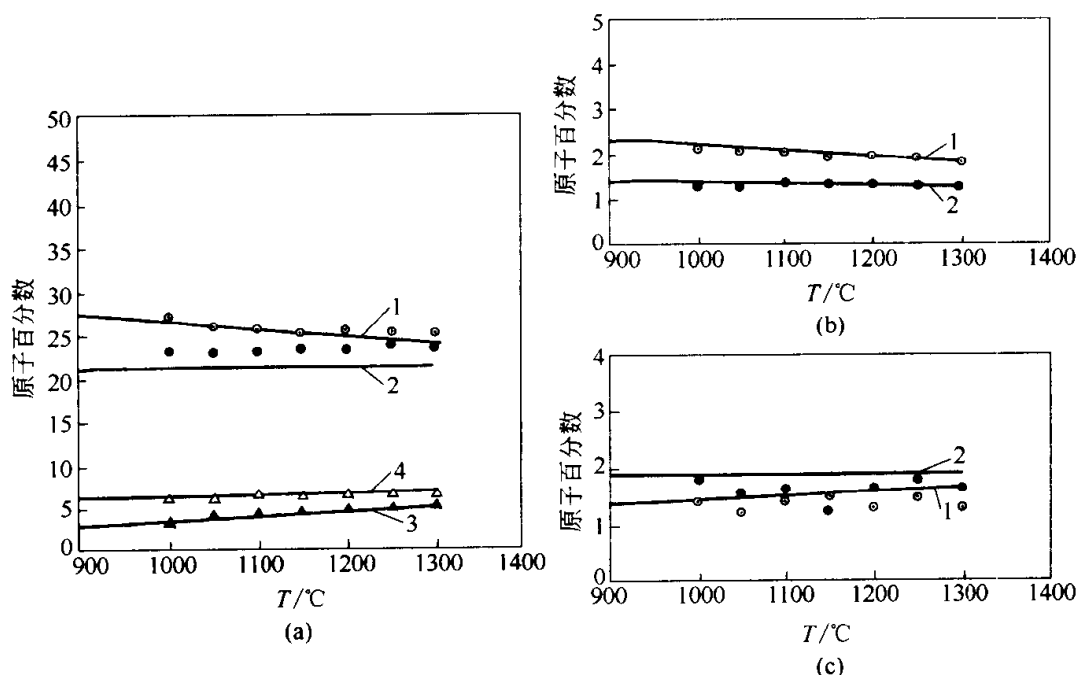


图 24-12 SAF2205 双相不锈钢奥氏体与铁素体中 Cr、Ni、Mo、Mn 的含量

(a) Cr、Ni 含量: 1—铁素体中的 Cr; 2—奥氏体中的 Cr;

3—铁素体中的 Ni; 4—奥氏体中的 Ni;

(b) Mo 含量: 1—铁素体中的 Mo; 2—奥氏体中的 Mo;

(c) Mn 含量: 1—铁素体中的 Mn; 2—奥氏体中的 Mn

的关系。高温下奥氏体的稳定性可以通过计算高温下不同成分的材料中奥氏体的相分数来实现，因为温度很高，我们可以假设材料处于平衡态。

由于很少量的 σ 相的析出就足以引起韧性的大幅降低，因此我们要考虑的是 σ 相的析出可能性，而不是平衡态 σ 相析出的多少。 σ 相的形核既牵涉到热力学条件也牵涉到动力学条件，在热力学条件中又牵涉到形核驱动力、界面能与应变能，为了简化，我们可以假设合金成分小的变化不会引起 σ 相形核过程中界面能、应变能以及动力学条件太大的变化， σ 相的析出可能性主要取决于 σ 相的形核驱动力。要计算 σ 相的形核驱动力，首先要搞清 σ 相是由什么相反应形成的，通过计算 SAF2205 双相不锈钢 200~1200℃ 各相的平衡摩尔分数可以发现，当合金冷却到 820℃ 时，铁素体的含量突然下降，奥氏体与 σ 相的含量突然上升，由此可见 σ 相是由铁素体共析分解而来（图 24-13），文献中的实验结果也证明了这一点^[33]，因此在计算 σ 相的驱动力时，所采用的成分为热处理温度下铁素体相的平衡成分。

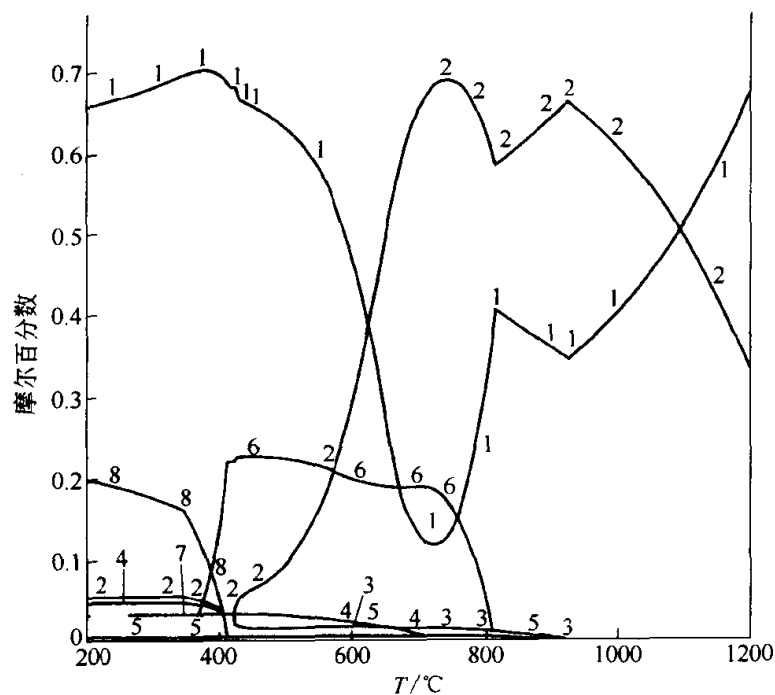


图 24-13 SAF2205 双相不锈钢中各相含量与温度的关系

1—BCC_A2; 2—FCC_A1; 3—HCP_A3#3; 4—LAVES;

5—M₂₃C₆; 6—SIGMA; 7—PI; 8—BCC_A2#2

不锈钢抗晶间腐蚀的能力取决于奥氏体中的 Cr 含量，一般认为当奥氏体中的 Cr 含量大于 12% 时，不锈钢就不会发生晶间腐蚀，但是事实并非如此，成分为 17Cr-12.5Ni-2.8Mo-1.5Mn-0.5Si-0.08N-0.02C（质量）的奥氏体不锈钢在 800℃ 中存在 99.7% 的奥氏体相和 0.3% 的 M₂₃C₆ 相，其奥氏体

Cr 含量为 17.4%，大于 12%，但是 800℃ 是这种钢的晶间腐蚀敏感化温度。可见影响晶间腐蚀的不是与 $M_{23}C_6$ 相平衡的奥氏体中的 Cr 含量，而是与 $M_{23}C_6$ 相的形成过程有关。

这个问题可以用图 24-14 所示的示意图来解释。假设 O 点为合金的总成分点，那么 M 与 A 分别为 $M_{23}C_6$ 与奥氏体平衡成分点。实际上在 $M_{23}C_6$ 的形成初期，由于合金元素扩散困难， $M_{23}C_6$ 中的 Cr 主要来自于 $M_{23}C_6$ 相附近的奥氏体，因此这些奥氏体中的 Cr 含量将远低于 A 点所示的值。假设刚形成的 $M_{23}C_6$ 的 Cr 含量接近于合金的总成分，即图中的 M' 点，那么与局部平衡的奥氏体成分为 A' ，实际的 $M_{23}C_6$ 成分就位于 MM' 之间，而奥氏体的成分则位于 AA' 之间。准确的成分需结合热力学平衡与扩散动力学理论进行相变模拟来计算，遗憾的是目前包含 8 个元素的相变模拟还无法进行。作为一种近似，我们假设除 Cr 外奥氏体中其他合金元素的成分在 $M_{23}C_6$ 的形成初期不会变化，计算奥氏体与 $M_{23}C_6$ 的局部平衡成分，即图中的 A'' 和 M'' 点。对上面所提到的奥氏体不锈钢其计算结果为：8.7%，低于 12%，这就解释了这种钢晶间腐蚀的原因。在设计双相不锈钢时我们也以同样的计算结果作为衡量抗晶间腐蚀的指标。对于点蚀我们用点蚀抗力经验关系式的值（ PRE 值）作为点蚀抗力指标。 PRE 值等于 $\%Cr + 3.3\%Mo + 30\%N + 0.5\%Ni - 0.5\%Mn - 30\%C$ （均为质量百分数）。

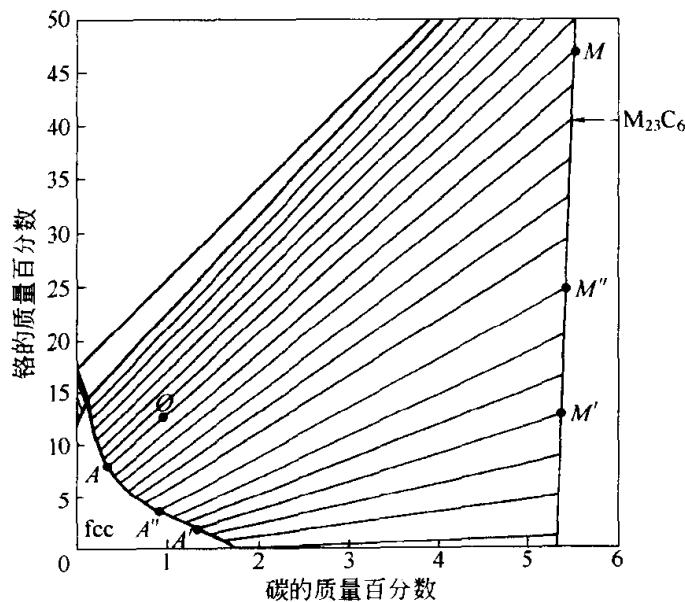


图 24-14 奥氏体与 $M_{23}C_6$ 两相平衡示意图

根据上面的讨论，我们最终将双相不锈钢的性能与成分（即相图计算可以计算的量）一一对应起来。我们进行双相不锈钢进行成分设计的最终目的

就是在保持双相钢的好的焊接性能与抗腐蚀性能的前提下，提高它的韧性。也就是寻找具有低的 σ 相形核相变驱动力，同时具有高的 高温（1050～1350℃）奥氏体相分数、高的与 $M_{23}C_6$ 局部平衡的奥氏体中的 Cr 含量与高的 PRE 值的合金成分。

首先 C 含量取具有代表性的 0.024%（质量），通过计算发现 σ 相形核相变驱动力随 Mn 含量的增加连续地增加，而随 Si 含量的变化，在含 Si 0.4%～1.0%（质量）处有一极小值，因此 Mn 含量取的一半，即 0.8%（质量），而 Si 含量提高到 0.7%（质量）。在此基础上在 SAF2205 附近选了 20 个成分，其中 N 含量为 0.14% 与 0.2%，Ni 含量为 4.4、5.5、6.5 和 7.5%（质量），Mo 含量为 0、1、2、3 和 4%（质量），Cr 含量随其他成分变化以保持 1050℃ 铁素体的含量与 SAF2205 钢相同。 σ 相形核相变驱动力、高温（1050～1350℃）奥氏体相分数、与 $M_{23}C_6$ 局部平衡的奥氏体中的 Cr 含量与 PRE 值。其计算结果见表 24-1，其中“×”表示该项指标比 SAF2205 差，“○”表示该项指标比 SAF2205 好，“—”表示由于 σ 相形核相变驱动力比 SAF2205 高未作进一步的计算，“√”表示该钢的全部 4 项指标都比 SAF2205 好。

表 24-1 双相不锈钢性能与成分的计算结果

合金成分/%(质量)				韧性	焊接性	抗腐蚀性	
N	Ni	Mo	Cr	σ 相形核相变驱动力	高温奥氏体相分数/%	奥氏体中 Cr 含量	PRE (37.28)
0.14	7.5	0	27.15	×	—	—	
0.14	7.5	1	26.21	×	—	—	
0.14	7.5	2	25.27	×	—	—	
0.14	7.5	3	24.33	×	—	—	
0.14	7.5	4	23.40	×	—	—	
0.14	6.5	0	25.93	×	—	—	
0.14	6.5	1	24.95	×	—	—	
0.14	6.5	2	23.97	×	—	—	
0.14	6.5	3	22.99	×	—	—	
0.14	6.5	4	22.01	○	○(3.2)	×	41.55
0.14	5.5	0	24.69	×	—	—	
0.14	5.5	1	23.67	○	×(0.3)	○	32.8
0.14	5.5	2	22.64	○	×(0.8)	○	35.08
0.14	5.5	3	21.62	○	×(1.4)	×	37.36
0.14	5.5	4	20.60	○	×(1.9)	×	39.63
0.14	4.5	0	23.43	○	×(0.0)	○	28.76
0.14	4.5	1	22.37	○	×(0.0)	○	31.00
0.14	4.5	2	21.29	○	×(0.0)	○	33.23
0.14	4.5	3	20.21	○	×(0.2)	○	35.45
0.14	4.5	4	19.13	○	×(1.0)	×	37.67

续表

合金成分/%(质量)				韧性	焊接性	抗腐蚀性	
N	Ni	Mo	Cr	σ 相形核相变驱动力	高温奥氏体相分数/%	奥氏体中Cr含量	PRE (37.28)
0.20	7.5	0	26.47	×	—	—	
0.20	7.5	1	27.54	×	—	—	
0.20	7.5	2	26.61	×	—	—	
0.20	7.5	3	25.69	×	—	—	
0.20	7.5	4	24.77	×	—	—	
0.20	6.5	0	27.24	×	—	—	
0.20	6.5	1	26.27	×	—	—	
0.20	6.5	2	25.31	×	—	—	
0.20	6.5	3	24.34	×	—	—	
0.20	6.5	4	23.38	×	—	—	
0.20	5.5	0	25.99	×	—	—	
0.20	5.5	1	24.99	×	—	—	
0.20	5.5	2	23.98	○	○(6.5)	○	38.22✓
0.20	5.5	3	22.97	○	○(7.2)	○	40.51✓
0.20	5.5	4	21.97	○	○(7.8)	○	42.8✓
0.20	4.5	0	24.73	○	○(3.6)	○	31.87
0.20	4.5	1	23.68	○	○(4.6)	○	34.12
0.20	4.5	2	22.63	○	○(5.4)	○	36.37
0.20	4.5	3	21.57	○	○(6.3)	○	38.61✓
0.20	4.5	4	20.51	○	○(7.0)	○	40.85✓

根据以上计算还可以发现,高的 N 含量低的 Ni 含量有利于韧性与可焊接性的提高,N 的含量的上限取决于 Cr_2N 对机械性能的影响。高的 Cr 含量有利于提高抗晶间腐蚀能力而高的 Mo 含量有利于提高抗点蚀能力。

虽然光凭以上的计算结果,不能判定材料一定就有好的性能,在实际材料设计中,计算必须与实验结合起来进行,但是计算可以大大减少实验的工作量,使实验目标更明确。材料科学通过几十年的发展积累了大量的实验数据,运用计算机处理这些数据,可以大大提高进一步的材料研究的速度与效率。

参 考 文 献

- 1 Van J.J.Laar,Z.Physik.Chem.1908,63:216;1908,64:257
- 2 Merjering J.L.,Philips.Res.Rept.1948,3:281
- 3 Kaufman L.and Bernstein H.Computer Calculation of Phase Diagram.New York:Academic Press,1970
- 4 Sundman B.,Jansson B.and Andersson J.O..Calphad.1985,9:153~190
- 5 Lukas H.L.,Henig E.Th.Zimmermann and B..Calphad.1977,1:155
- 6 Lukas H.L.,Weiss J. and Henig E.Th..Calphad.1983,7:229
- 7 Chen S.L.,Chou K.C.,Chang Y.A..Calphad.1993,17:237;1993, 17:287

- 8 Dinsdale A.T. .Calphad. 1991,15:317~425
- 9 Redlich,O. and Kister. A. Ind. Eng. Chem. 1948,40:345~348
- 10 Hildebrand J.H. .J. Amer. Chem. Soc. 1929,51:66
- 11 Pelton A.D. and Thompson W.T. .Prog. Solid State Chem. 1975,10(3):119
- 12 Hardy H.K. .Acta Metall. 1953,1:202;1954,2:348
- 13 Bente. Proc. Royal Soc. 1935,A(150):552
- 14 Sommer F. ,Z. Metallkd. 1982,73:72
- 15 Sommer F. ,Z. Metallkd. 1982,73:77
- 16 Kikuchi R. .Phys. Rev. 1951,81:998
- 17 Sudman Bo, Mohri T. .Z. Metallkd. 1990,81:251
- 18 Hillert M. ,Stanflasson L. I. .Acta. Chem. Scand. 1970,24:3618
- 19 Sundman B. ,Jansson B. .J. Phys. Chem. Sol. 1981,42:297
- 20 Ansara I. ,Chart T.G. ,Guillermet A.F. ,Hayes F.H. ,Kattner U.R. ,Pettifor D.G. ,Saunders N. ,
Zeng K. .Calphad. 1997,21:171
- 21 Hillert M. ,Jansson B. ,Sundman B. .Z. Metallkd. 1988,79:81
- 22 Hillert M. ,Jansson B. ,Sundman B. ,and Agren J. .Metall. Trans. 1985,A(16):661
- 23 Inden G. .Proc. Calphadv. Max-Plank Inst. Eisenforschung, Düsseldorf. 1976,1~13
- 24 Hillert M. ,Jarl M. .Calphad. 1978,2:11~21
- 25 Kohler F. .Monatsh fur Chemic. 1960,91,783
- 26 Colinet C. .Fac. Des. Sci. .Univ. Grenoble, France. 1967
- 27 Muggiann Y.M. ,Gambino M. ,and Bros I.P. .J. Chemic Physique. 1975,72:83
- 28 Toop G.M. .Trans. Met. Soc. Aime. 1965,233:850
- 29 Hillert M. .Calphad. 1980,4:1
- 30 Kubaschewski O. and Alcock C.B. .Metallurgical Thermohemistry. 5th ed. New York: Pergamn Press,
1979
- 31 Lee B.J. .Applications of Thermodynamics in the Systhesis and Processing of Materials. Edited by Nash
P. and Sundman B. , The Minerals. Metals & Materials Society. 1995,215
- 32 Hayes F.H. .Hetherington M.G. and Longbottom R.D. .Mater. Sci. & Eng. . 1990,6:263
- 33 Norstrom L.A. ,Pettersson S. and Nordin S. .Werkstofftech Z. . 1981,12:229

第 25 章 材料前沿的若干问题

撰稿人 冯 端 冯少彤

25.1 科学、技术与材料

现代科学与技术唇齿相依，关系密切。一般而言，它们之间存在有两种基本模式：①是由于生产实践的需要而创建了技术（如蒸汽机等热机技术），然后提高到理论上来（建立了热力学），再反馈到技术中去，促使技术进一步发展；②是先在实验室中揭示了某些基本规律，建立了比较完整的理论体系，然后再在生产技术中创建全新的技术部门。19 世纪电磁学的发展，提供了这一模式良好的范例，它为近代的电力工程和无线电技术奠定了基础。尽管这两种模式继续在当代社会中发挥作用，应该看到，第二种模式正在发挥愈来愈重要的作用。下面着重讨论材料物理与材料研究的问题。材料技术的核心问题在于新材料的研制和传统材料性能的提高。在过去，传统材料的发展依赖于长期的经验积累和不断炒菜式的试验，在其中理论的指导并不起太大的作用。即使在对于钢的相变尚一无所知的时代，熟练的技工已掌握了淬火使钢硬化的诀窍，就很足以说明问题。但自从 20 世纪 40~50 年代以来，固体物理学，随后深化和扩大为凝聚态物理学，发展成为成熟的科学，可以从基础理论上解释物质的物理性质，形成了材料物理学^[1]，遂使材料研究的面目大为改观。

25.1.1 技术对材料研制的促进

材料促进了技术的发展，反过来技术的发展又为材料的制备和表征提供了全新的手段：超高真空条件下的分子束，脉冲激光和磁控溅射沉积均为薄膜材料制备提供了过去难于设想的优越手段；高分辨电子显微术、扫描隧道显微术和同步辐射提供的 X 射线源、飞秒级的激光器等，有的在空间分辨率上达到原子尺度，有的时间分辨率达到飞秒的量级，已成为研究材料的结构和过程及其重要的手段（参阅第 22 章）。

探索和研制新材料的传统方法是凭借经验来进行“炒菜”，即配制许多成分有差异的试样，然后通过性能的表征来进行筛选。显然这种方法费时费功，效率不高，但仍然不失为一种有成效的研制新材料的手段。面临的问题在于如何利用当代技术来提高“炒菜”的效率。1995 年以来，组合合成法（combinatorial synthesis）被应用于材料的研制，提供了一种高效率的“超级

炒菜”的手段^[2]。制备方法一般采用气相沉积，其母体为多层膜，使组元数可以很多，再应用掩模技术，使沉积物限于小点，从而在基片上制备出成分各异的材料点构成的列阵，被称为集成材料芯片。再应用多种现代技术对材料芯片进行扫描，从而筛选一些性能优异的材料进行进一步的研究（图 25-1）。组合合成法最初用于药物的研制，在移植到材料研制后，它成绩斐然，曾用于筛选高温超导体、发光材料、高介电常数材料等研制工作^[2]。而居里点超过室温的铁磁半导体也是采用这种方法首次发现的。

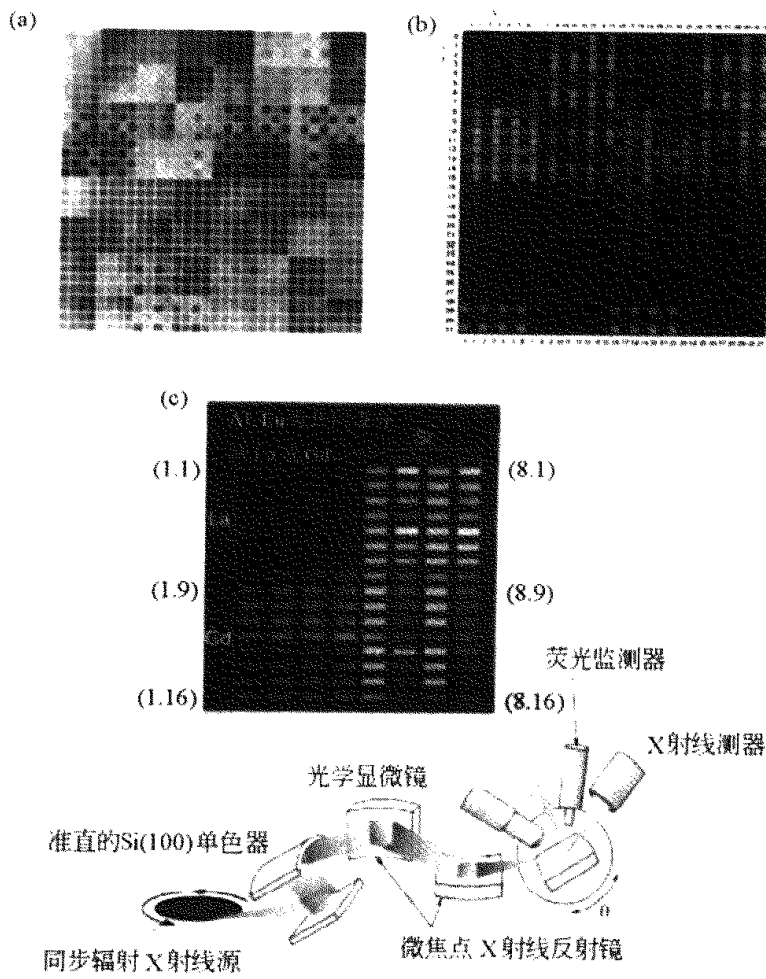


图 25-1 材料芯片

- (a) 在白天看到 4 元材料芯片，不同斑点的彩色是由元材料斑点的厚度及折射率不同所造成的；
- (b) 在多带发射（中心波长为 254nm）的紫外灯照射下的材料芯片荧光照片；
- (c) 同步辐射在 X 射线探针的示意图，在试样上产生 $3 \times 20\mu\text{m}^2$ 的 X 射线束斑。插图在紫外光普照下的发光材料芯片^[2]

25.1.2 当代技术的挑战与机遇

过去 100 多年以来，科学与技术相互影响，铸就了现代文明的支柱。具

体来说, 可以用图 25-2 (根据美国科学院与工程院联合组织对凝聚态与材料物理的综述^[1]) 来予以阐明。

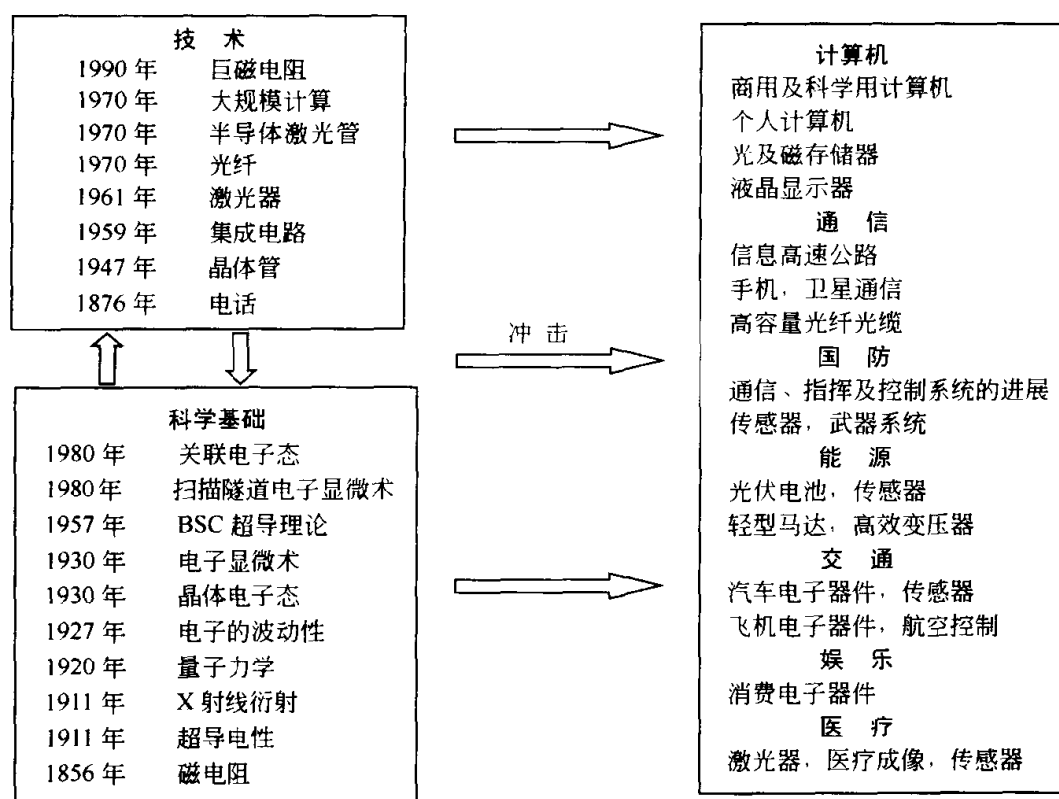


图 25-2 变革的时代

应该指出, 科学家的理论蓝图只有凭借合适的材料才能实现。如果没有可控掺杂的高纯度、高完整性的硅薄膜材料, 所有的晶体管和集成电路, 乃至基于这器件的电子计算机都将成为一句空话。如果不能制备出低损耗的光纤材料, 没有高效率的 GaAs-GaAlAs 量子阱激光器, 也决不可能实现光纤通信。因而新型材料的研制、开发和应用就成为当今信息技术的主流。

下面讨论一下 21 世纪面临的问题。大家知道计算机芯片的集成密度是随年代而呈指数式的增长, 服从一经验规律, 被称为摩尔 (Moore) 定律 (见图 25-3)。应该注意, 从 20 世纪 60 年代起, 芯片的发展的线路是单一的: 即缩小集成电路的尺寸。但从图上可以看出, 到了 2015 年左右, 集成电路的尺寸将达到其物理极限, 芯片厂的投资量也会超出了社会的承担能力 (参看图 25-4), 必须改弦易辙, 另觅途径, 这也对材料科学研究提出了重大的挑战和无限的机遇。除此之外, 光纤通信的容量, 磁存储的容量也是随年代作指数上升 (见图 25-5 与图 25-6), 也对材料科学的研究提供机会和挑战。如何进一步发展微电子材料、光电子材料和磁性材料不仅有学科上的重

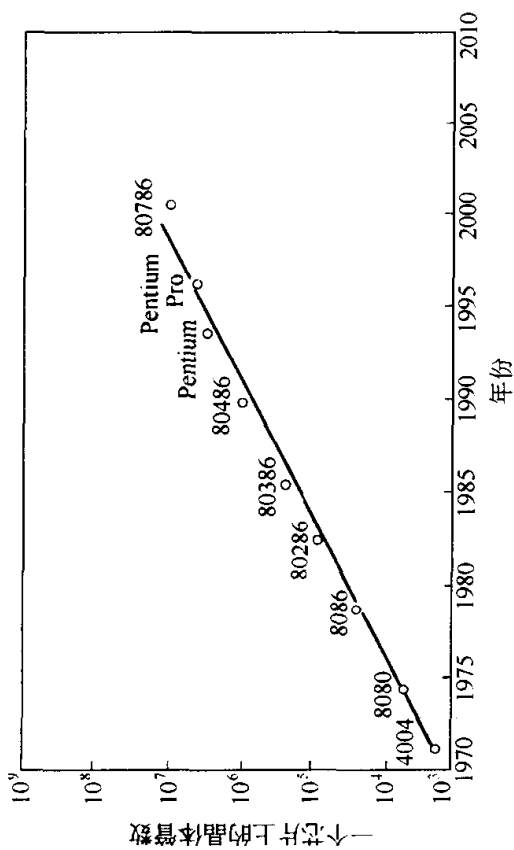


图 25-3 芯片发展年代曲线^[1]

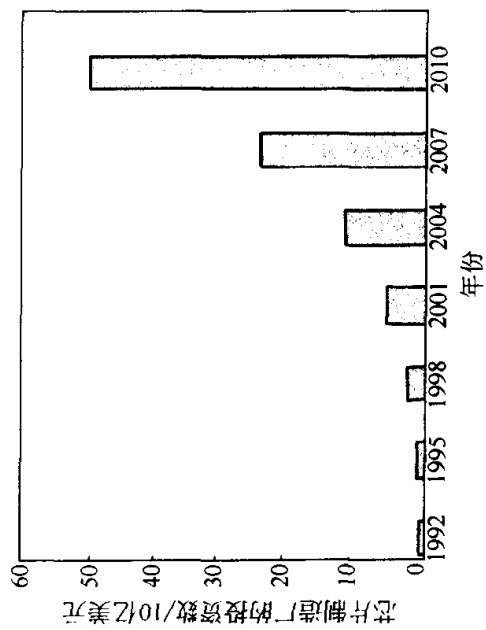


图 25-4 芯片厂投资的年代趋势^[1]

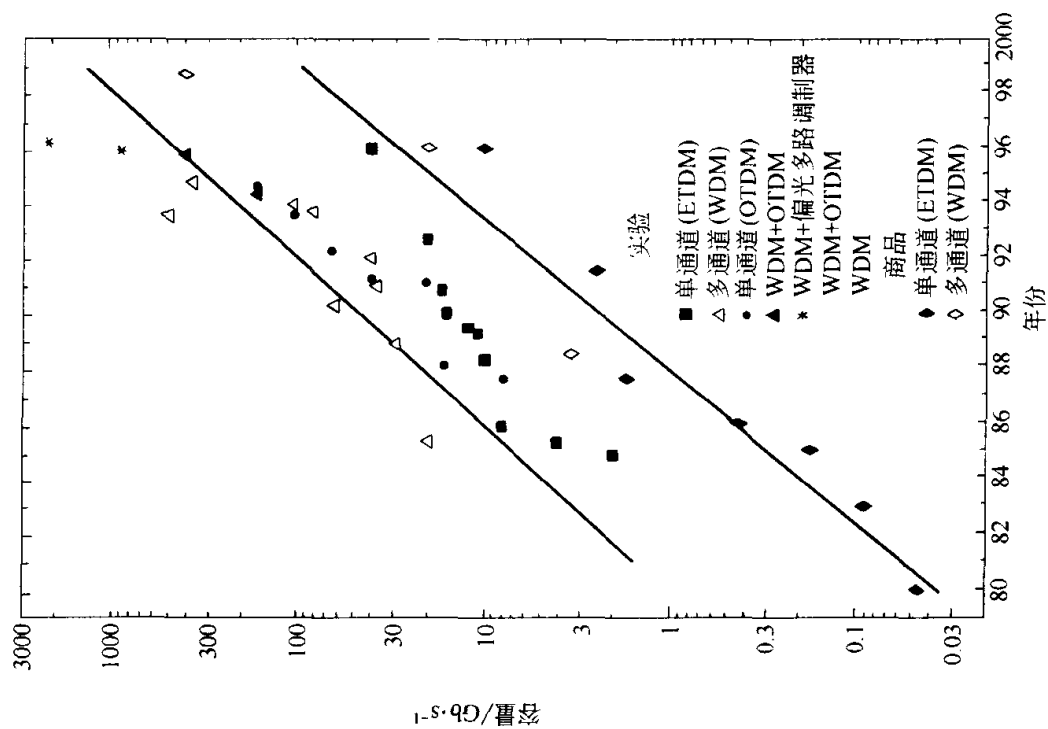


图 25-5 光纤通信容量发展的年代表^[1]

大意义，也成为 21 世纪科技发展的关键问题之一。

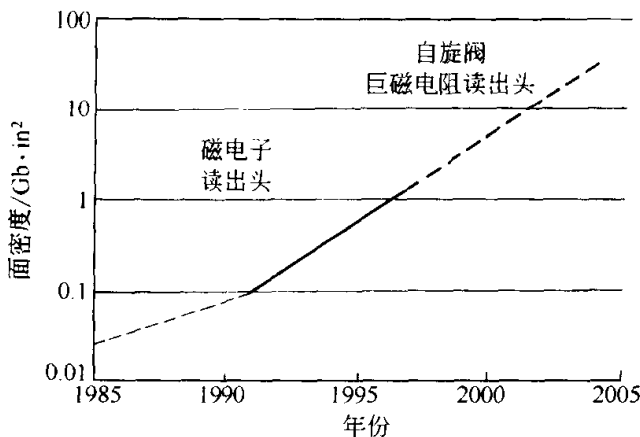


图 25-6 磁存储容量发展的年代表^[1]

25.2 新材料发展的主要动向

凝聚态物理学的一些基本理论，如固体能带理论、点阵动力学理论、对称破缺的相变理论、有序媒质的基态理论、原激发理论、缺陷理论和临界现象理论等，都是对于粒子数非常大的^[3,4]，这些理论很富有成效。

应该说我们对常规材料的物性已经基本了解了。比如，对 Cu, Ag, Al……的导电性质应该说是清清楚楚；Si, Ge, GaAs, InP……。这一类的半导体的性质也是清楚的；Nb, Sn 等金属乃至超导金属与合金的超导性质自从 BCS 理论提出来后，可以说是一清二楚了。关于材料的铁磁性也基本上是清楚的，虽然过渡金属巡游电子的铁磁性还有一些疑问，多半属于理论细节问题，对于材料的发展大体已经足够。对于 BaTiO₃, LiNbO₃, KDP 等的铁电性和非线性光学性质也是清楚的。所以可以说，许多物性的问题从大体上来说，已经解决了，而且目前计算材料科学得到了发展，已经可以进行所谓“从头计算”（ab initio calculation），根据物质结构的基本数据来计算出材料的物理性质（请参阅第 23 章材料设计）。

下面进一步的发展主要反映在材料的复杂化上，大体上有 3 个途径。

25.2.1 开拓超微结构的领域

通过不同材料的复合来获得优异的性能，首先在结构材料领域中崭露头角。在宏观尺度上进行，复合钢铁混凝土提供了一个绝好的例证。而在显微尺度上进行复合，沉淀强化的硬铝合金就是一个例子。铝具有良好的塑性，但强度略嫌不足。但过高强度金属间化合物的弥散分别使其强度提高，但仍然保留其良好的塑性。第二相粒子的尺度和其平均间距是强化效应的关键参量。这可以从位错理论上得到说明。随后又发展了纤维强化材料，使硬的纤

维（例如碳纤维）埋在软的基质（例如树脂）中，从而硬软兼备，成为优异的航天材料。关键问题在于阻止裂口的萌生和发展。

而在电子功能材料的发展方面，则更需要材料物理理论乃至凝聚态物理的介入。下面就着重讨论这一方面的问题。

简单结构的材料通过微结构化获得不同尺度（特别是纳米尺度，约 1~100nm）的复合结构。这个问题的基本观念，著名的理论物理学家 R. Feynman 在 1957 年就讲过：“如果我们得以在细微尺度上控制事物的话，毫无疑问，这将使得材料所具有的性质范围大为扩充”^[5]。这就是近年来非常热门的所谓纳米结构的基本构想：就是拿不同的常规材料组合起来构成了具有特征长度为纳米量级（约 1~100nm）的复合结构。纳米结构的关键在于引入有物理意义的特征长度。

量子力学早就说明了粒子可以翻越势垒而呈现隧道效应。可应用这一原理制出隧道结：两导电层间夹有绝缘层，其厚度应小于电子的隧穿长度，而此长度约为 1nm 的量级。早在 20 世纪 50 年代，江琦就据此发明了半导体隧道二极管，20 世纪 60 年代初，Giever 发明了单电子超导隧道结。

另外有一类宏观的隧道现象。20 世纪 60 年代初超导体的约瑟夫森 (Josephson) 效应的被理论预言、实验证实和技术应用就是最早的例证，这里涉及的特征长度是超导体的相干长度，在这种情形下作隧穿的粒子是库珀对。这些工作对于后来半导体电子学和超导电子学均产生了深远的影响，他们三人合得了 1972 年的诺贝尔奖。

值得注意，1975 年 Julliere 发现了与自旋有关的电子隧道效应，但未受学术界的重视。一直到 20 世纪 90 年代科学家重新研究了这一问题，制出了具有巨隧道磁电阻的器件，可能发展为新型的不可擦除的磁随机存储 (MRAM) 技术。

如果结构的尺寸进入电子费米波长的范围，就会呈现另一类的量子力学现象，即量子限制效应：量子阱体现了一维的限制效应，即在受限制的方向出现了德布罗意波驻波所引起的离散能级（参看图 25-7）。一维的量子限制效应可以通过两种不同的化合物半导体（如 GaAs 与 GaAlAs，具有不同的能隙宽度），复合成为三明治式的结构来实现。早期的工作可溯源于江琦与朱兆祥的理论设想，他们于 1970 年首先提出了半导体超晶格的概念（参看图 25-8）。随后利用卓以和所发展的分子束外延新技术，张立纲与江琦等制备了能隙高低不同的半导体多层膜，在实验室中实现了量子阱、多量子阱和超晶格，观测到极其丰富的物理效应^[7]。这项研究从基础理论来看，为半导体物理学打开了新局面，经过多年发展，已经成为当今半导体物理学的主流^[4]。

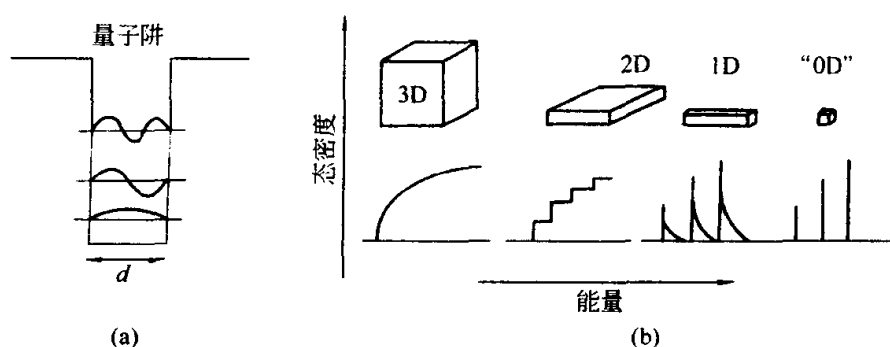


图 25-7 量子限制效应

(a) 量子限制效应的能量（波数）示意图；(b) 不同维度量子限制效应的态密度与能量的关系

量子阱与超晶格不仅有重大的理论意义，也在实用的器件上大显身手：Z. J. Alferov 和 H. Krömer 就由于利用量子阱发展了高迁移率晶体管和高效率激光二极管而获得了 2000 年的诺贝尔物理奖；而基于超晶格概念而发展起来的量子级联激光器，采用 AlInAs/GaInAs 的超晶格，可以获得 $3 \sim 13 \mu\text{m}$ 波长范围红外激光器，也受到技术界的重视。而中村修二发展了晶格失配度较大的 GaN 薄膜的分子束外延技术，制成了基于 GaN 材料的多量子阱发光二极管和激光器，将发光的波长延拓到蓝光和紫光的领域，并在实用化方面取得了巨大成功，不仅提高了光盘的效果还可能引发一场照明技术的革命。

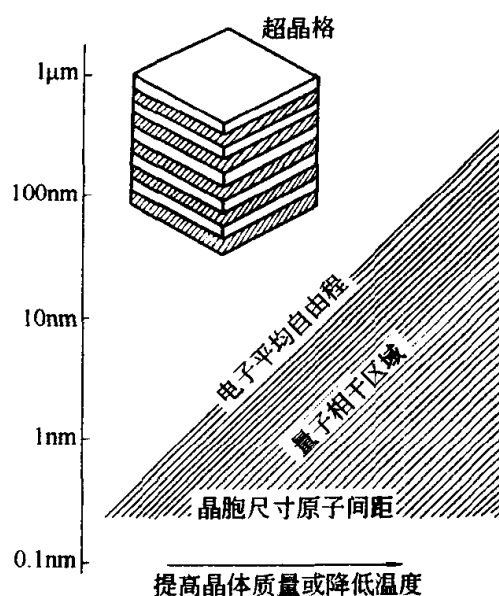


图 25-8 量子相干区域与超晶格示意图

这一类型的研究往往融物理设计、材料制备与表征和器件发展为一炉，宛然成为当今新材料发展最先进技术的标志。

令人惊奇的多孔硅可见光发射的发现，又将发展硅基材料光电子学器件的课题提到议事日程上来了。块状的硅是间接带隙材料，只能发射微弱的红外光，因而虽在微电子学材料中独占鳌头，而在光电子学中一直无立足之地。多孔硅的光发射改变了这一情况，虽则多孔硅由于结构脆弱，可能难以发展为实用的材料，但在其指引下，纳米硅基材料的研究方兴未艾，值得注意。已有实验证实了在硅中传播的光可以增益，这表明硅激光器的出现已指日可待。由于硅片是微电子技术的支柱，将发光管和激光管都集成于硅片之

上, 其技术上重要性是显而易见的。

由于金属中传导电子浓度高, 相应地使其费米能级的波长要短得多, 约为 1nm 量级, 仅为半导体的 1%。因而, 要在多层金属中实现量子阱就比半导体困难很多。然而薄膜制备技术的进展克服了这一困难, 而巨磁阻效应的发现就是一个证明。由于这一现象已在第 10 章论述过, 这里就不再重复了。值得注意的是以往电子学只应用了电子具有电荷这一特征, 各种器件也只是对于电荷输运加以控制和影响。巨磁阻效应及相关的巨磁隧道电阻效应可以视为自旋电子学 (或磁电子学) 的开端, 其理论与实际意义十分重大^[7]。美国 IBM 公司已利用巨磁电阻效应的自旋阀作为磁存储器的读出磁头, 已产生每年百亿美元的产值, 如果 MRAM 能够实用化, 则其产值还要大。这是基础研究通过新型材料来实现重大经济效应的又一实例。而当前的一个重要问题在于如何将自旋极化的电流注入常规的半导体之中, 这就需要将磁性金属和半导体界面之间结构加以改善, 而更加理想的是发展具有高居里点的铁磁性半导体。科学家利用分子束外延方法对 GaAs 薄膜超常规掺锰, 取得可观的结果, 锰含量可高达原本固溶度的数十倍, 居里点达到 110K; 另外, 在具有锐钛矿 (anatase) 结构的掺 Co (6%~8%) 的 TiO_2 , 获得高于室温的居里点。这些结果表明自旋电子学的发展和新材料的研究与开发息息相关。

从二维量子阱朝向更低的维度开拓, 到达一维的量子线和零维的量子点, 理论和实验研究都十分活跃。一维量子线的研究以碳纳米管最为活跃, 随着蜷角的不同, 它可以是半导体, 也可以是金属, 已经进行了原型器件的研制, 有可能开拓碳电子学这一新领域^[8]。

具有零维结构的量子点, 在 3 个维度上均呈现量子限制效应, 使其能谱为一系列离散的能级, 与原子中电子的能级十分相似, 因而被比拟为人工原子^[9]。从基础理论上来说, 如何使人工原子具有天然原子所没有的性能也是很有兴趣的研究课题, 而且量子点所载电荷也是量子化的, 正如早年 Millikan 油滴实验就已验证了的。因而要加电子到量子点上, 会受到库仑斥力的作用, 产生库仑阻塞效应, 使电子只能一个一个地加上去。利用这一效应可以制作单电子晶体管或存储器等原型器件, 相当于极度微型化的电子器件。另外, 如果单个的量子点相当于人工原子, 一对耦合起来的量子点就相当于人工分子, 而一大群量子点耦合在一起就相当于人工固体。科学家已经开始对这些问题进行了初步的实验演示, 是否会对新材料的研制和开发具有重要意义尚待进一步的工作方见分晓。

早在 1962 年久保就提出了超微颗粒的量子限制理论, 推动实验物理学家向纳米尺寸的微粒进行探索。在 20 世纪 80 年代, 处于大块凝聚态物质和单个原子之间的原子团簇 (clusters) 受到了重视。实验科学家采用了多种方法, 如

蒸发和超喷注、激光蒸发等，制备了一系列团簇，原子数从几个到几千见图 25-9 (a)。质谱分析表明团簇的丰度具有一系列幻数 (magic number)，对应于特别稳定的团簇结构。在金属中幻数对于电子壳层填满的情况，而在惰性元素，团簇的幻数则可归结为几何壳层的效应。特别有意义的是在共价键的碳团簇中，发现了丰度特高的幻数 60。Kroto 和 Smalley 认为价键完全饱和的 60-碳原子分成于一个截面 20 面体的顶角上，像一个足球状的分子，即 C_{60} ，被称为巴基球或富勒烯。随后科学家又在氦气氛中用石墨电极的电弧放电法，制备了大量的 C_{60} ，有的形成了固体，引起科学界的轰动。 C_{60} 分子可以衍生出许多化合物，有人推测它对有机化学的意义可能不亚于苯环的发现。掺碱金属的 C_{60} 固体具有超导性， T_c 达到 30K 的水平，是氧化物超导体以外， T_c 最高的材料。 C_{60} 是团簇组装材料的一个特例，其特点在于有序度非常高。20 世纪 80 年代初 Gleiter 利用金属在超高真空中蒸发后冷凝，再原位压结，获得另一种不规则的三维团簇组装材料，见图 25-9 (d)，被称纳米结构材料，其晶粒尺寸、形状和取向均不一致，而晶界上的原子数达到原子总数的 50%，观测到异常的性能，包括力学性能。其他的方法，如非晶晶化、机械合金化，乃至改进传统的陶瓷工艺也可以制备这类材料。利用这种纳米结构材料来获得韧性陶瓷也是颇有前景的，人们也许要问，除 C_{60} 以外，是否还有其他的有序的团簇组装材料？这是一个很具有挑战性的问题。通过一些理论计算，得出了一些设想的稳定团簇，如何制备大量的单一尺寸的这类团簇，再组装起来，是一个很值得花力气来解决的实验技术问题，近年来发展的自组装方法开辟了一条可行的道路（参阅第 21 章）。

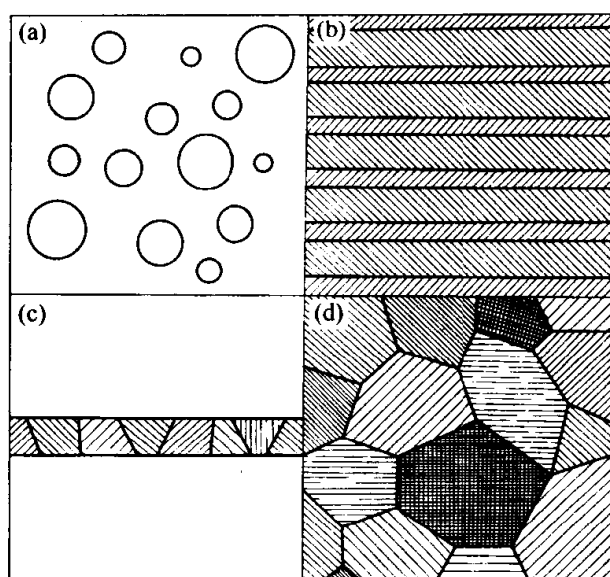


图 25-9 纳米尺寸的微结构材料的示意图

(a) 团簇或纳米微粒；(b) 多层膜或超晶格；(c) 颗粒膜；(d) 块状纳米材料

如果说半导体技术中体块材料向薄膜材料的过渡在 20 世纪 50 年代就已经完成了, 而铁电压电材料, 非线性光学材料, 这一过渡则尚在进行之中, 集成光学和集成铁电学的发展指明了这一趋向: 相应地这类材料通过微结构设计来获得优异性能也就提到议事日程上来了。早在 1980 年我国科学家就利用铌酸锂规则的畴结构来取得激光倍频增强效应, 实现了非线性光学晶体的准位相匹配。近年来这一领域已经成了国际上的热门领域, 这类微米尺度的微结构设计也在高频超声换能、电光效应等领域发挥作用 (参看第 13 章)。

纳米尺度并不是材料微结构化的惟一尺度, 微米和亚微米也是在光波范围内运作的特征尺度。有关光子晶体问题亦可参阅第 13 章。通过光子晶体来截断和引导光波的传播正在当前激光技术中发挥愈来愈重要的作用^[10]。

应该指出, 制备纳米结构的技术, 除了分子束外延等高精密的物理技术外, 利用传统化学方法进行不同尺度 (从微米到纳米) 自组装也取得了优异的结果, 请参阅第 21 章。

25.2.2 探测电子关联体系中的宝藏

凝聚态理论的长足进展已经搞清楚了许多材料的物性问题, 但是还存在一些疑难问题悬而未决, 其中最突出的莫过于关联电子体系的问题。所谓电子关联就意味着电子和电子之间存在库仑相互作用, 这一点也不稀奇。传统的能带理论在处理固体中的电子系统时, 首先是忽略了电子之间相互作用, 将电子系统视为相互独立的理想气体, 考虑单电子与晶格的周期结构之间的相互作用, 从而得到了固体的能带结构, 然后再引入电子间的相互作用加以修正。这样的理论处理显然适用于弱关联的电子体系, 即高浓度的电子体系, 其电子间并能大大超过了电子之间库仑相互作用的势能, 也就是通常的宽能带物质的情况。能带理论的巨大成功每每掩盖了这一理论的不足之处, 即它并不适用于强关联的电子体系, 即电子浓度甚小的物质, 标志电子之间库仑相互作用的关联能的重要性被突出出来了。这类材料往往具很窄的能带, 只比完全局域化的能级略宽一点。早在 1937 年科学家就发现 NiO, MnO, CoO 等氧化物并不是能带理论所预言的金属, 而是能隙甚大的绝缘体。这一结果引起固体理论学家的关注^[4]。Mott 引进关联能来解释这一物理问题, 认为是 d 电子间库仑相互作用抑制了极化涨落, 产生了关联能隙。后来这一类绝缘体即被称为莫特绝缘体。莫特进一步讨论 VO_2 , V_2O_3 等材料因温度或压力改变所引起的绝缘体到金属的相变, 认定它们也是电子关联导致的相变, 后来被称为莫特转变。

莫特绝缘体为数不少, 几乎占 $3d$ 过渡金属二元氧化物中的一半, 还有难以数计的多元复杂氧化物和 $4f$ 稀土化合物和 $5f$ 锕系化合物。

关联物质往往处于金属与绝缘体的界限附近, 即电子处于完全离域化的

扩展态和完全局域化的能级之间。要判断电子是离域化还是局域化,就要看 f 电子或 d 电子波函数的分布范围是否和近邻产生重叠。图 25-10 画出了量子力学计算给出的 $4f$, $5f$, $3d$, $4d$, $5d$ 元素的 l 壳层体积与晶体的 Wigner-Seitz 原胞体积之比值。可以看出电子壳层体积以 $4f$ 为最小, $5f$ 次之, 其后乃是 $3d$, $4d$ 与 $5d$ 。这样, 可以确认, 电子态局域化程度的顺序大致为 $4f > 5f > 3d > 4d > 5d$; 换言之, 电子态形成的能带宽度按此顺序增加, 而关联性则按此顺序递减。还可以看出, 从左往右穿过周期表时, 部分填充电子壳层的半径逐步降低, 而关联性则逐步增强。利用以上规则, 科学家重新画出了一张周期表 (见图 25-11)^[3]。表上画出了两根斜线, 将线外划分为两个区域: 一是左下角, 离域能的能带效应占优势; 另一是右上角, 反过来, 局域化和关联性占优势。在两条斜线的中间或近邻, 标志了强关联 (窄能带) 的区域。在这区域内, 材料常会出现反常的物理性质。

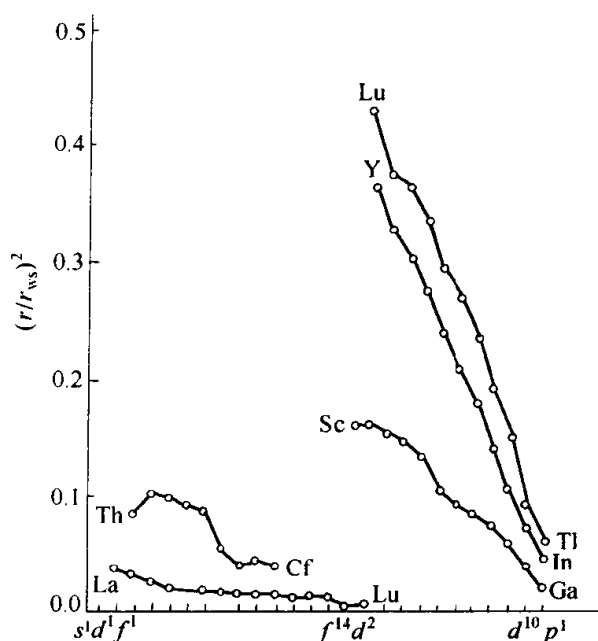


图 25-10 $4f$, $5f$, $3d$, $4d$ 与 $5d$ 元素的 l 壳层体积与 Wigner-Seitz 原胞体积之比值随它们在周期表中位置的变化关系

的半径逐步降低, 而关联性则逐步增强。利用以上规则, 科学家重新画出了一张周期表 (见图 25-11)^[3]。表上画出了两根斜线, 将线外划分为两个区域: 一是左下角, 离域能的能带效应占优势; 另一是右上角, 反过来, 局域化和关联性占优势。在两条斜线的中间或近邻, 标志了强关联 (窄能带) 的区域。在这区域内, 材料常会出现反常的物理性质。

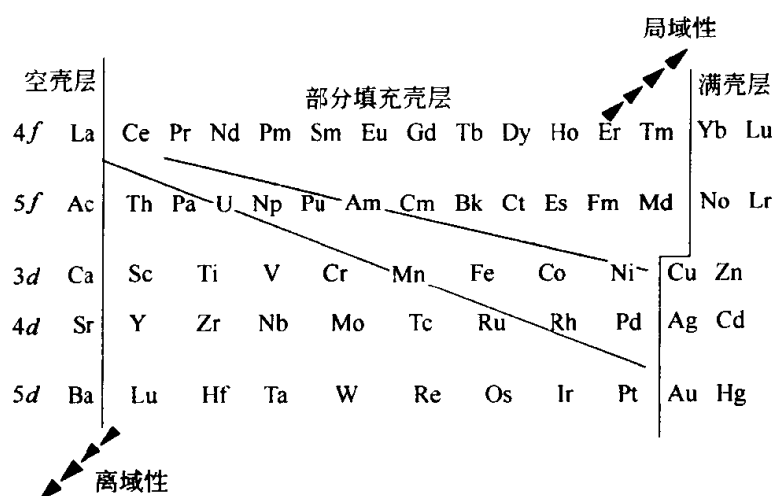


图 25-11 显示稀土、钢系及过渡金属的 f 、 d 轨道局域化趋势而重新绘制的周期表

我们来视察一下周期表和它的斜线区：具有强铁磁性的金属 Fe, Co, Ni 都在斜线区域之内，它们的能带具有强关联性，它们的电子巡游相互作用是其铁磁性的根源；但由于窄的 d 带和宽的 s 带交叠，因而它们的导电性仍接近于正常金属。和斜线邻近的 Mn 和 Cr，在低温下这些金属具有反铁磁性。而以 Fe_2O_3 为重要成分的铁氧体，如尖晶石型的 NiFe_2O_4 ，石榴石型的 $\text{Y}_3\text{Fe}_5\text{O}_{15}$ (YIG)，磁铅石型的 $\text{BaFe}_{12}\text{O}_{19}$ 和钙钛矿型的 YFeO_3 ，都是具良好绝缘性能的强磁性材料。它们通常是亚铁磁性，即磁矩不相互抵消的反铁磁序，其反铁磁序的根源在于是磁性离子通过间隔的氧离子的超交换相互作用而实现的，而超交换相互作用也和强关联有关，如 Anderson 的超交换理论所指出。另一方面强关联还可以导致 Zener 提出的双交换作用，使某些复杂氧化物具有铁磁性和金属导电性，如 $\text{La}_{1-x}\text{Ca}_x\text{MnO}_3$ ($0.2 < x < 0.4$)，即为一例。值得注意 Mn, La 均在周期表上为斜线区域的近邻。

具有特殊能带结构半金属磁性材料也引起了科技界的重视。早在 1983 年 R.A.de Groot 对于复杂结构的磁性合金进行了能带计算，其结果表明可能存在一种具有特殊能带结构的磁性材料，其一种自旋（例如向上）的电子具有金属特征，而另一种自旋（向下）的电子却具有绝缘性，称之为半金属磁体。这是计算材料设计成功的一个良好范例，20 年后，理论成为现实。

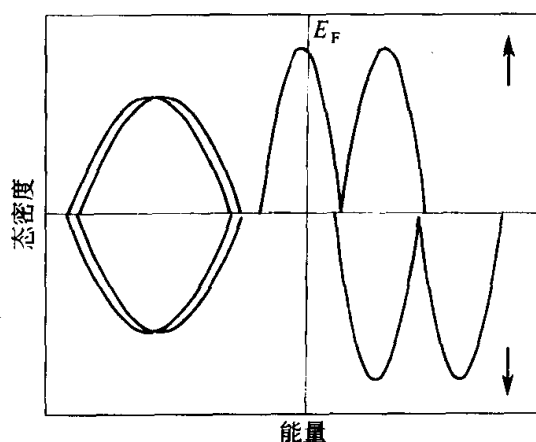


图 25-12 CrO_2 的态密度对能量的示意图，显示了自旋半金属性的能带结构

图中箭头表示电子自旋的取向

如今已经确证了一系列的化合物和合金具有半金属铁磁体或亚铁磁体的特征^[11]，例如 CrO_2 、 $\text{Sr}_2\text{FeMnO}_6$ 、 Cu-CrO_4 、 Co_2MnSi (Heusler 合金) 等。 CrO_2 的能带结构如图 25-12 所示，在费米面只有自旋向上的电子的态密度不为零。因此传导电子的自旋极化率可能达到 100%，因而这类材料对自旋电子学的重要性是不言而喻的。

我们再来看重电子金属^[4]，斜线区中或近邻的 Ce, U, Np 等元素，成为这些合金的必要成分。它们的特征在于低温比热容异常地大，从而可以推断其电子的有效量 m^* 与自由电子的 m 之比值异常的高 ($m^*/m \approx 100 \sim 1000$)，名符其实的是重电子。例如此比值， CeAl_3 为 600， CeCu_6 为 740， CeCu_2S 为 460， UBe_{13} 为 300， NpBe_{13} 为 280， UPt_3 为 178，而 U_2Zn_7 与 UCd_{13} 均大于 100。如此重的电子应接近于局域态，但这些合金均具有金属导电性，使人困惑不解。重电子合金中

CeCu₂Si, UBe₁₃, UPt₃ 在低温具有超导电性, 但 T_c 不高, 均小于 1K, 而其超导电性质, 由于偏离常规, 因而被科学家所注意研究。在常规超导体中, 少量磁性杂质就可以破坏超导电性; 但在重电子超导体中, 反电磁关联可以和超导电性共存。看来它的电子配对机制与 BCS 理论的有些不同, 可能是由于电子间的直接相互作用。

如果说研究重电子合金的兴趣主要是由于理论上有意义, 那么, 氧化物高温超导体的发现, 就因其潜在的巨大实际意义而轰动全球^[13]。La-Sr-Cu-O, Y-Ba-Cu-O, Bi-Sr-Ca-Cu-O, Te-Ba-Cu-O, Hg-Ba-Ca-Cu-O 等相继问世, T_c 值已达到 135K。在这些材料中准二维的 CuO₂ 平面起了关键性作用。从 La-Sr-Cu-O 系的相图来看, 在 x 为零的材料是反铁磁序的莫特绝缘体, 随着掺杂浓度 x 的增大, 发生了绝缘体到金属的转变, 而低温就具有超导电性, 随着掺杂浓度的增大, T_c 达到一峰值之后, 又逐渐下降。高温超导体的正常态的电子性质都十分异常, 这也是强关联材料一种特征^[12,13]。

值得注意, 近年来在 La-Ca-Mn-O 系的材料之中又发现了庞 (colossal) 磁电阻效应^[14], 加上磁场后的电阻变化率 $\Delta R/R$ 值可达到 103~106 之巨。这种材料的铁磁性的根源是双交换相互作用, 而且磁性转变与绝缘体金属转变相邻近。磁场在这类材料引起电阻变化的量级甚至于超过了超导转变中电阻的变化。显而易见, 这种异常的物理性质将会得到技术上的应用。

应该指出, 我们对于强关联体系的科学认识尚不完备。这方面的科学的研究, 不管是实验上还是理论上的都尚有待深入。从材料研究的角度来看, 这些多元复杂结构的氧化物是尚未勘探清楚的新材料的富矿区。过去, 像铁氧体这类适用高频技术的磁性材料的开发实早于超交换相互作用理论的提出, 更不要说 Fe, Ni, Co 具有强铁磁性是早为众所周知, 在 Bednorz 与 Müller 发现氧化物超导体过程之中, 机遇起了重要的作用。随着对关联电子体系物理学的较深入的了解, 将会大大地促进这一领域中新材料的勘探和开发工作。

25.2.3 开拓有机材料的新领域

材料进一步复杂化也体现在从无机材料到有机材料的发展。有机材料是基于有机分子, 主要成分是碳, 可能再加上氢、氧、氮……等元素。虽然组成的元素不多, 但其化合物的品种繁多, 而结构上的复杂性也远远超过了无机材料。有机材料的早期发展也侧重于结构材料: 塑料、橡胶和各种人造纤维材料。塑料滞弹性行为则可用长键分子沿管道的蛇行理论来解释。橡胶的高弹性, 是交联高分子体系的特征, 则用形变导致有序化造成熵的减少来予以说明。要提高分子材料的强度, 使长链分子取向一致乃是一条有效途径: 尼龙纤维依靠拉制来获得高强度; 高分子液晶的取向脱溶则又是另一种强化

的途径。聚酰胺纤维（商品名 kevlar）就是一个突出的例子，其强度业已超过了最强的钢丝（钢琴弦），而其强度密度比更是高达钢琴弦的 8 倍，是制造防弹衣的良好材料。这里侧重讨论有关有机电子和光电子材料方面的问题。

在 20 世纪 60~70 年代，有关简单有机分子晶体（如萘、蒽及并多苯）的电子性质的研究就已开始，明确了分子本身的电子结构仍然大体上遗存在晶体中发挥作用，如最高已占分子轨道（HOMO）与最低未占分子轨道（LUMO）的电子跃迁决定光吸收与发射的波长，当然在固体中能级略有展宽，这些结果已总结在 Pope 与 Swenberg 的专著之中^[15]。图 25-13 显示一些简单分子的结构，光吸收区的波长和根据 Hückel 分子轨道近似计算出来的萘能级图。

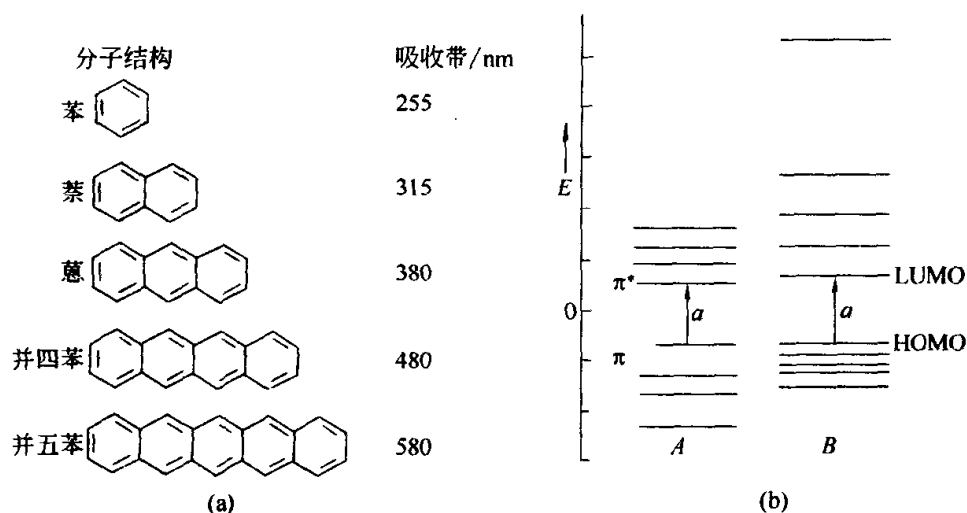


图 25-13 苯及其衍生物的结构与能级

(a) 一些简单分子的结构，光吸收区的波长；(b) 萘的电子能级（A 及 B 分别为根据 Hückel 分子轨道近似未加和已加交迭修正的计算结果）

在 1964 年 W. Little 提出惊人的理论设想，即认为具有旁支的聚合物中电子通过激子配对产生室温以上的高温超导体。这一设想引起了一些物理学家研究有机材料的导电性质的兴趣，其中包括了 A. J. Heeger 的研究组。在 20 世纪 70 年代 Heeger 的研究组发现了一组有机电荷转移盐类 TTF-TCNQ 在 70K 具有导电异常现象，最初认为可能是高温超导体，进一步研究确认为这是一维晶体中金属-绝缘体的 Peierl 相变。随后转而研究共轭聚合物的导电性质。他与 MacDiarmid 与白川英树合作在反式聚乙炔（trans-polyacetylene）中通过电子受主如 AsF₅，或碘或电子施主如钠的掺杂，可以在室温获得金属性的导体，其电导率高达 10000S/cm。这样，就将有机金属导体引入了材料技术，而他们三人于 2000 年获得了诺贝尔化学奖。物理学家苏武沛、

Schrieffer 与 Heeger 也对其导电机制进行了研究, 提出了有影响的孤子与极化子等有机导体理论。

在有机光电显示器件中首先得到广泛应用的是液晶材料。1971 年 M.Schadt 与 W.Helfrich 首先提出利用扭曲 90° 的相列型液晶盒, 用加电场的方法使偏振光通过与否来实现液晶的数字显示, 已广泛地应用于手表、计算器的显示。而近年来铁电液晶材料的进一步发展又导致了平板式大屏幕的彩色显示。

到 20 世纪 80 年代中有机薄膜制备有了长足的进展, 为发展有机半导体提供了条件。1987 年邓青云与 Van Slyke 演示了高效电致发光器件, 它是由芳香族六胺的空穴输运层和 Alq_3 (8-羟基喹啉) 的发射层, 使有机半导体受到科技界的重视。在器件方面主要集中在电致发光器件和场效应晶体管; 在材料方面, 即有用像 Alq_3 这类的小分子, 也有用共轭聚合物, 如 PPV (polyphenylenevinylene)、聚噻吩 (polythiophene) 等, 其中齐聚物 (oligomer) 起了很大的作用^[16]。例如 6-噻吩 (sexithiophene), 是聚噻吩的齐聚物, 每个分子都是由 6 段单体所构成的, 性能比较容易控制, 因而在共轭聚合物光电子性质研究中起了示范的作用。

在 20 世纪 90 年代后期, 不管是有机发光管及其列阵, 还是有机场效应管及相应的集成电路, 都有很大进展, 正在向实用化迈进。有机半导体有容易制备和价格便宜的优点, 其重要缺点在于迁移率不高, 例如 6-噻吩的迁移率最高也只达到 $7 \times 10^{-2} / (\text{V} \cdot \text{s})$, 接近非晶硅的水平, 比晶态半导体要低好几个数量级。

到 2000 年 H.Schön 研究组的一组工作, 又使有机导体的研究上了一个台阶。他们制备了纯度较高的有机单晶体 (并四苯、并五苯、 C_{60} 等), 而将这些绝缘体制成双级场效应管, 由改变极电压来对晶态表面层进行掺电子或空穴, 掺入电子或空穴仅限于晶体表面薄层内, 从而可以避免化学掺杂引入的无序和结构畸变等不利影响。这样一来, 有机晶体的迁移率有明显的提高参阅图 25-14。并四苯和并五苯在室温达到了 $3 \text{cm}^2 / (\text{V} \cdot \text{s})$ 的量级, 而且空穴高于电子, 而处于低温下, 达到 $10^5 \text{cm}^2 / (\text{V} \cdot \text{s})$ 的量级, 并在 1.7K 观测到分数量子霍尔效应, 这在无机半导体晶体中要实现也是十分困难的, 随后又在并四苯晶体同时掺入电子与空穴, 通过它们的复合而发射激光, 成为有机半导体电致激光发射的首例 (工作温度 5K, 阈值电流 $500 \text{A}/\text{cm}^2$)。

上面讲过, 有机导体的研究受到了高温超导理论设想的推动。这方面研究的实际情况究竟如何呢? 1980 年 Jerome 首先发现了一种有机电荷转移盐类 $(\text{TMTSF})_2$ 在 0.9K 以下具有超导性, 不是一个不起眼的开端。到 1987 年, 另一种有机导体 $(\text{BEDT-TTF})_2\text{Cu}(\text{SCN})_2$ 具有 11.4K 的 T_c , 开始有一

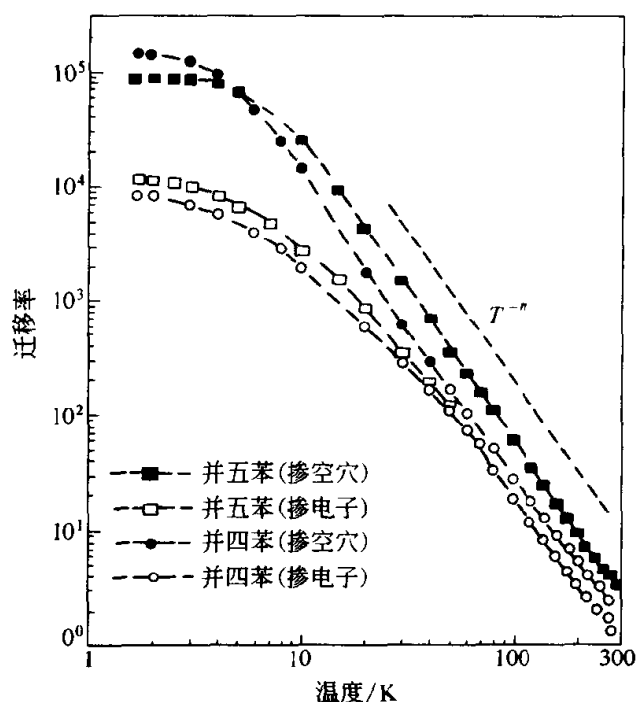


图 25-14 并四苯与并五苯单晶的迁移率与温度的关系

些苗头了。在掺杂碱金属的 C_{60} 固体中发现了更高的超导 T_c : K_3C_{60} ——18K, Rb_3C_{60} ——28K。显示了随着晶体参数的膨胀, T_c 有变高的趋势, 但化学掺杂使晶格不稳定, 因而难以获得高的 T_c 。Schön 的研究组在 2000 年底, 报道了他们对于 C_{60} 单晶体进行电荷掺杂的结果, 获得了掺电子只能达到 11K, 而掺空穴可以达到 52K 的优异结果。他们设想如果掺入一些有机分子使晶格膨胀, 当可进一步提高 T_c 。2001 年夏, 这一设想得以实现, 在 C_{60} 固体中掺了 ClH_3 , T_c ——80K; 掺了 BrH_3 , T_c ——117K^[17]。似乎 T_c 还存在进一步提高的可能性。就是 117K T_c 的材料已经可以在液氮温度下工作, 而且这一系统比氧化物超导体要干净得多, 而且便于微电子芯片兼容, 这是其重要优点。但这一类超导体的许多物性尚待研究。图 25-15 显示了 3 种不同类型超导体 (金属与合金、氧化物及有机物质) T_c 随年代提高的情况, 显示了有机材料与无机材料之间的竞争。

全部由轻元素组成的纯有机物中没有磁性离子, 似与铁磁性无缘。1991 年 M. Kahashi 等首先在准一维的纯有机物 P-NPNN ($C_{13}H_{16}N_3O_4$) 中取得了突破, 发现了铁磁性, 虽然居里点只有 0.65K。在 C_{60} 固体问世之后, 情况即有所改观。在 C_{60} 固体中掺入有机化合物 TADE [四(二甲胺)乙烯], 铁磁居里点提高到 17K。在纯有机化合物中要算是居里温度最高的, 但距离实用性还很远。2001 年秋, T. L. Makarova 等发现纯 C_{60} 固体在高压 (6GPa) 和高温 (1025~1050K) 处理后, 碳笼间键合由弱键转化为强的共价键, 变

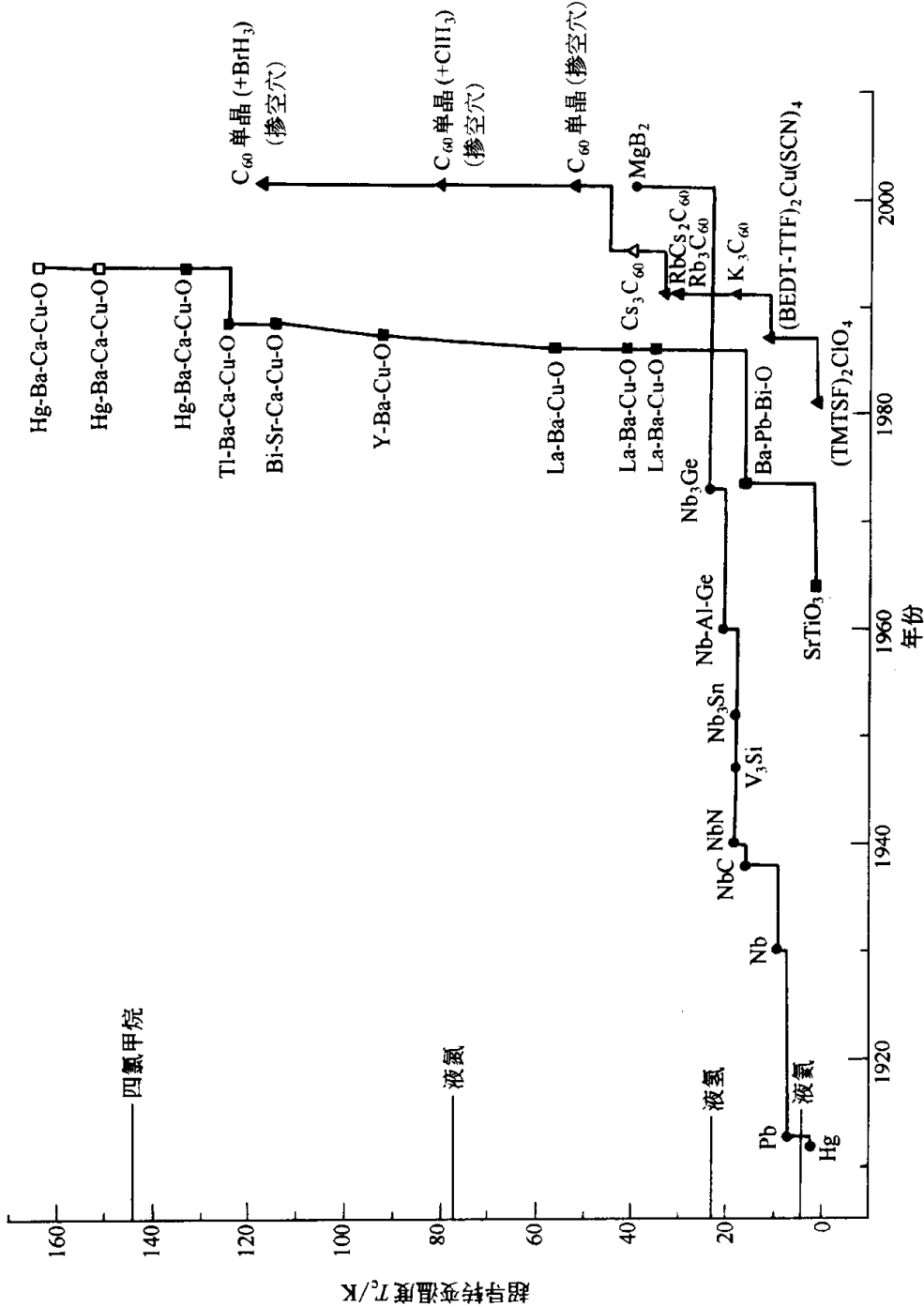


图 25-15 3 种材料 (金属与合金、氧化物及有机物质) T_c 随年代的提高

图中 3 种材料分别用 3 种不同符号来表示, 实心符号为常压的 T_c , 空心符号为高压下的 T_c , 图上也显示了 3 种致冷剂的沸点

为层状的菱面体相，其居里点高达 500K，成为第一种居里点在室温以上的铁磁体^[18]。当然，含过渡金属离子〔如钒（四丙烯腈）〕的有机材料，具有亚铁磁性，其居里点约为 400K，在 1994 年就已制出。但纯有机材料居里点高于室温，这是首次报道，从而产生了轰动性。有关 C_{60} 及其衍生物的新结果几乎层出不穷，令人应接不暇。许多结果还需要进一步证实和研究。有关机制问题也尚待探明。这说明了有机材料的丰富多彩，充满了机遇和挑战。

以上讨论的还限于一般的有机材料。至于生物聚合物，由于存储了特定的信息顺序，显然是更加错综复杂，而且灵巧机动。即以蜘蛛丝为例，细到肉眼不易看清楚，但其强度能挡住以 1m/s 速度飞翔的蚊子。就强度而言，超过了等直径的钢丝；不仅如此，它的长度延伸到 4 倍方才断裂。这种优异的力学性能显然和它复杂而且多层次的结构有关。图 25-16 显示了不同尺度下观察到的结构示意图^[19]。这充分阐明了材料复杂化的神通广大，而蜘蛛吐丝技术的高妙也使材料科学家们叹为观止。因而仿生材料成为当今材料科学的一个热门课题；而智能材料的发展也受到生物结构的影响。

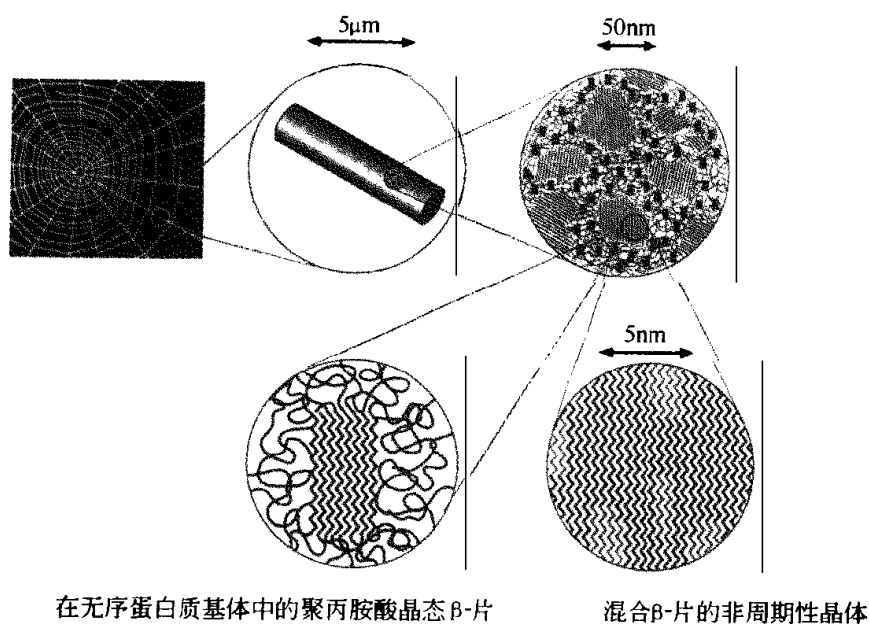


图 25-16 蜘蛛丝多层次结构的示意图

25.3 结语

将以上三方面的论述汇合起来看，有一条主线贯穿其间，即从简单到复杂，沟通了材料科学基础理论的发展和新材料的研制。第一部分所讨论的问题涉及结构（包括晶体结构和电子结构）比较简单的物质之中发展新材料的问题，其主要对策在于引入超微结构这一层次来增加其复杂性，从而使材料

的面目焕然一新；第二部分所论述的问题涉及电子结构具有复杂性的物质，一般而言，化学成分也相当复杂。随着物理机制的逐渐探明，新的材料亦将不断涌现；第三部分所论述的问题，复杂性主要体现在分子层次上，而且随着研究的深入，可以更加充分利用众多层次上的复杂性来优化材料的性能。目前已经探明的部分还只相当于冰山浮露在水面上的部分；大部分问题还潜伏在水面之下，尚有待于进一步的探测。

参 考 文 献

- 1 Committee on Condensed Matter and Materials Physics, National Research Council, Condensed Matter and Materials Physics, Washington D. C.: National Academic Press, 1999
- 2 Xiang X.D. Ann. Rev. Mater. Sci. 1999, 29: 149
- 3 黄昆原著. 韩汝琦改编. 固体物理学. 北京: 高等教育出版社, 1988
- 4 冯端, 金国钧. 凝聚态物理学新论. 上海: 上海科技出版社, 1992
- 5 Feynman R. Engineering and Science. 1960, (2): 22
- 6 夏坚白, 朱邦芬. 半导体超晶格. 上海: 上海科技出版社, 1995
- 7 Grünberg P. Phys. Today. 2001, 54 (5): 22
- 8 Dresselhaus M S, Dresselhaus G, Eklund P C. Science and Technology of Fullerenes & Carbon Nanotubes. New York: Academic Press, 1998
- 9 Yoffe A. D. Adv. Phys. 2001, 50: 1
- 10 Joannopoulos J. D. Villeneuve, Fan S H. Nature. 1997, 386: 143
- 11 Pikett W. E. Moodera J. S. Phys. Today. 2001, 54 (5): 39
- 12 韩汝珊. 高温超导物理. 北京: 北京大学出版社, 1998
- 13 周午纵, 梁维耀主编. 高温超导基础研究. 上海: 上海科技出版社, 1999
- 14 Rao C N R. Phil. Trans. Roy. Soc. London, A. 1998, 356: 23
- 15 Pope M, Swenberg C E. Electronic Processes in Organic Crystals. Oxford: Clarendon Press, 1982
- 16 Greenham N C. Friend R H. Solid State Physics. 1998, 49: 1
- 17 Schön J. H. Kloc C. Batlogg B. Science. 2001, 293: 2432
- 18 Makarova T. L. et al. Nature. 2001, 413: 716
- 19 Ball P. Made to Measure: New Materials for the 21st Century. Princeton: Princeton Uni. Press, 1997

[G e n e r a l I n f o r m a t
i o n]

书名 = 材料科学导论

作者 = 冯端 师昌绪

页数 = 7 4 1

S S 号 = 1 0 8 7 5 2 3 5

出版日期 = 2 0 0 2 年 0 5 月 第 1 版