

直接法中相位退化问题的处理

范海福 千金子

(中国科学院物理研究所)

1980年6月10日收到

提 要

相位退化是直接法应用中的一大问题。所谓相位退化，系指非中心对称晶体结构因子的相位在推引或修正过程中收敛为中心对称类型的相位。属于 $P2_1$ 、 $P4_1$ 等空间群的晶体，常会出现这一情况。相位退化使相应的 E 图添加了一个对称中心，这使 E 图中“鬼峰”的数目大增，因而往往导致直接法的失败。本文提出一种两步处理方法：先将晶体当作中心对称求解以获得一个同时包含真正结构及其对映体的、中心对称的赝结构；然后在此基础上利用所谓分量关系式从赝结构解出真正的结构。此法经用已知结构的晶体检验，证明效果良好。

一、引 言

在单晶体结构分析中，直接法的应用目前已经非常广泛。使用像 MULTAN^[1] 或 SHELX^[2] 等程序，在多数情况下可以顺利地解出相当复杂的结构。但是对于一些看来不太复杂的非中心对称结构有时却并不成功。有一些文章^[3,4] 对失败的原因作过分析。许多例子表明，失败往往是相位退化引起的。这里所说的相位退化，是指非中心对称晶体结构因子的相位在推引或修正的过程中收敛为中心对称类型的相位。其结果是使所得的 E 图同时包含真正的结构及其对映体。相位退化既可以发生在测定含重原子晶体的过程，也可以发生在测定不含重原子晶体的过程。我们已经在一些文章中探讨过前一种情况的解决办法^[5,6]。这里只讨论后一种情况。

一些人把相位退化归咎于选择对映体的办法不当或者是用以指定对映体的结构因子相位在修正时不够稳定。Hauptman^[7] 用一组相互联系紧密的衍射代替单个的衍射以保证对映体选择的恰当和稳定；Hull^[4] 则用所谓概率加权正切式修正 (SWTR) 方法以防止相位向错误方向漂移。Schenk^[3] 认为相位退化同直接法所用判据本身的缺陷有关，并证明凡属 $P2_1$ 、 $P4_1$ 、 Pc 、 $C2_1$ 、 $P2_1nb$ 、 $C2ca$ 等空间群的晶体往往容易发生相位退化。他提出了一种称为“ Σ_2P ”的事后处理方法。这种方法首先把非中心对称的结构当作中心对称结构求解，推引出结构因子的正负号(其实只是相应实分量的正负号)，这样就可以得到同时包含真正结构以及其对映体的 E 图。然后将这样的 E 图同 Patterson 图重迭，以解出真正的结构。这种方法的特色在于不把包含对映体的 E 图看作失败的终结，而使它成为通向成功的阶梯。这样，就可以设法顺应相位退化的趋势，化弊为利，干脆把推引相位简化为

推引正负号, 反而大大节省了时间。但是对于如何从一张真伪混杂的 E 图中辨认出真实的结构这一问题, Schenk 求助于 Patterson 函数, 而不是着眼于充分发挥直接法本身的潜力。本文提出另一种事后处理的方法。它采用了 Schenk 法的前半部。但是在后半部中用分量关系式^[8,9]代替 Patterson 函数, 不在实空间而在倒易空间进行处理。这个方法较之 Schenk 法, 途径更为直接; 它实质上包含了强化对映体选定的措施, 而比 Hauptman 所用的方法更可靠; 它同样以 MULTAN 法为基础, 但较之 Hull 的 SWTR 方法节省时间。

二、退化相位及相应的 E 图

令归一化结构因子 $E_H = \alpha_H + i\beta_H$, 则 \sum_2 关系式

$$E_H = K \sum_{H'} E_{H'} E_{H-H'}$$

可以写成如下的分量关系式:

$$\alpha_H = K \sum_{H'} \alpha_{H'} \alpha_{H-H'} - K \sum_{H'} \beta_{H'} \beta_{H-H'}, \quad (1)$$

$$\beta_H = 2K \sum_{H'} \alpha_{H'} \beta_{H-H'}, \quad (2)$$

其中 K 是一个常数。这时, 正切式

$$\tan \phi_H = \frac{\sum_{H'} |E_{H'}| |E_{H-H'}| \sin(\phi_{H'} + \phi_{H-H'})}{\sum_{H'} |E_{H'}| |E_{H-H'}| \cos(\phi_{H'} + \phi_{H-H'})}$$

可以写成如下形式:

$$\tan \phi_H = \frac{2 \sum_{H'} \alpha_{H'} \beta_{H-H'}}{\sum_{H'} \alpha_{H'} \alpha_{H-H'} - \sum_{H'} \beta_{H'} \beta_{H-H'}} = \frac{\beta_H}{\alpha_H}, \quad (3)$$

其中 ϕ_H 是 E_H 的相位。

当相位发生退化时, 各归一化结构因子的虚分量 β 均趋近于零。这时,

$$\phi_H = \cos^{-1} \left[S \left(\sum_{H'} \alpha_{H'} \alpha_{H-H'} \right) \right] = \cos^{-1} [S(\alpha_H)],$$

其中 $S(\)$ 表示 $(\)$ 的正负号。

由此可见, 在发生相位退化时, 结构因子的相位蜕变为相应的实分量的正负号。另外, 如果将非中心对称晶体按中心对称情况推引结构因子的正负号, 首先建立起强符号关系的衍射点将是那些 β 接近于零的衍射点。因此所得结果将基本上是结构因子实分量的正负号。

现在再来考虑用退化相位计算的 E 图。实际上, 它就是 $S(\alpha_H) |E_H|$ 的傅里叶反变换:

$$\Phi^{-1}[S(\alpha_H)|E_H|] = \frac{1}{V} \sum_H S(\alpha_H)|E_H| \exp(-2\pi i H \cdot r), \quad (4)$$

其中 V 是晶胞体积。

令

$$|E_H| = |\alpha_H| + \Delta_H,$$

则(4)式可以写成

$$\begin{aligned} \Phi^{-1}[S(\alpha_H)|E_H|] &= \frac{1}{V} \sum_H \alpha_H \exp(-2\pi i H \cdot r) \\ &+ \frac{1}{V} \sum_H S(\alpha_H) \Delta_H \exp(-2\pi i H \cdot r). \end{aligned} \quad (5)$$

$$\begin{aligned} (5) \text{ 式右边第一项} &= \frac{1}{2V} \sum_H (E_H + E_H^*) \exp(-2\pi i H \cdot r) \\ &= \frac{1}{2} [\rho_E(r) + \rho_E(-r)], \end{aligned}$$

其中 E_H^* 是 E_H 的共轭复数; $\rho_E(r)$ 是代表真正结构的 E 图; $\rho_E(-r)$ 是代表对映体的 E 图。因此(5)式右边第一项相当于经过锐化的、同时包含真正结构以及其对映体的电子密度图。

$$\begin{aligned} (5) \text{ 式右边第二项} &= \frac{1}{V} \sum_H [S(\alpha_H)|E_H| - S(\alpha_H)|E_H| |\cos \phi_H|] \exp(-2\pi i H \cdot r) \\ &= \frac{1}{V} \sum_H [S(\alpha_H)|E_H|(1 - |\cos \phi_H|)] \exp(-2\pi i H \cdot r). \end{aligned}$$

注意

$$\frac{1}{2\pi} \int_0^{2\pi} |\cos \phi_H| d\phi_H = \frac{2}{\pi} \int_0^{\pi/2} \cos \phi_H d\phi_H = \frac{2}{\pi},$$

将(5)式右边第二项按 ϕ_H 取平均可得

$$\left\langle \frac{1}{V} \sum_H S(\alpha_H) \Delta_H \exp(-2\pi i H \cdot r) \right\rangle_{\phi_H} \cong 0.363 \frac{1}{V} \sum_H S(\alpha_H) |E_H| \exp(-2\pi i H \cdot r).$$

由此可知,统计而言,(5)式右边第二项的贡献约占总贡献的 36.3%。加之 $S(\alpha_H)$ 同 Δ_H 之间是一种无规的联系,可以预期(5)式右边第二项的贡献一般将表现为较小的随机分布的鬼峰。

综上所述,由退化相位计算的 E 图可以近似地看作是真正结构同其对映体混合的、经过锐化的电子密度图。从这样的 E 图不难读出成对的双解原子坐标 x_j, y_j, z_j 及 $-x_j, -y_j, -z_j$ 。

三、分量关系式的应用

在由退化相位计算的 E 图中可以读出全部或大部分原子的双解坐标 x_j, y_j, z_j 和 $-x_j, -y_j, -z_j$ 。从每一对坐标中任选一个就可以按下式算出归一化结构因子的实分量

$$\alpha_H = \sum_{j=1}^N \frac{Z_j}{\left(\varepsilon \sum_{j=1}^N Z_j^2 \right)^{1/2}} \cos 2\pi H \cdot r_j, \quad (12)$$

式中 Z_j 是晶胞中第 j 个原子的原子序; r_j 是同 x_j, y_j, z_j 或 $-x_j, -y_j, -z_j$ 相应的坐标矢量; ε 是同空间群对称性及衍射指标类型有关的一个整数^[10]。由 α_H 及 $|E_H|$ 又可以算出归一化结构因子的虚分量的绝对值

$$|\beta_H| = \sqrt{|E_H|^2 - \alpha_H^2}. \quad (13)$$

由分量关系式(2)可得

$$S(\beta_H) = S\left(\sum_{H'} \alpha_{H'} \beta_{H-H'}\right). \quad (14)$$

根据(12),(13)和(14)式就可以推引出 β_H 的正负号。(14)式的使用方法同中心对称情况下 Σ_2 关系式的用法相仿。但(14)式有一个更高、更好的起点。这首先是因为(14)式中所有 $\alpha_{H'}$ 包括绝对值和正负号都是已知的。因此,与 Σ_2 关系式只能提供每三个符号之间的关系不同,(14)式可以直接给出结构因子虚分量两两之间的符号关系。另一方面,使用 Σ_2 关系式时,一般只能从一个不满 10 个衍射点的起始套出发。但使用(14)式时,从一开始就可以有一个包含上百个衍射点的特大起始套。这使符号推引的过程既简单又可靠。根据作者的经验,只要挑选出一个合适的结构因子虚分量 $\beta_{H-H'}$, 任意给定它的正负号(这相当于选定对映体)就不难推出其余 $\beta_{H-H'}$ 的正负号。这个过程不像附加记号法^[4]那样需要在起始套中引入未知正负的记号;也无需像 MULTAN 法^[1]那样在起始套中考虑各种可能的相位或正负号的组合。

具体推引 β_H 的正负号可按如下步骤进行:

1. 利用 MULTAN 程序求出收敛图;根据双解原子坐标算出收敛图底部约 100 个衍射点的 α_H 。
2. 从收敛图底部挑选约 10 个 $|E_H| \gg |\alpha_H|$ 的衍射点(一般要求 $|E_H|$ 比 $|\alpha_H|$ 大 2—3 倍^[5]);按(13)式求出相应的 $|\beta_H|$ 。
3. 将上面所得的 α_H 和 $|\beta_H|$ 按(2)式所示的分量关系相互组合,挑选出一个关系式最多、最强的 β_H , 指定其符号为 +1。
4. 从这个 β_H 和已知的约 100 个 α_H 出发,按(14)式用迭代法求出约 10 个 β_H 的正负号。
5. 用上述约 10 个 β_H 和相同指标的 α_H 算出相应的 10 个 ϕ_H (结构因子相位),再从收敛图底部挑选出约 10 个 $|E_H| \sim |\alpha_H|$ 的衍射点,令其 $\phi_H = \cos^{-1}(S(\alpha_H))$ 。将这 20 个左右的结构因子相位代入正切式作相位扩展及修正。

四、实例试验

L-丙氨基-甘氨酸 (L-Alanylglycine), 化学式为 $C_5O_3N_2H_{10}$; 空间群是 $P2_1$; 晶胞参数 $a = 5.283 \text{ \AA}$, $b = 11.810 \text{ \AA}$, $c = 5.510 \text{ \AA}$, $\beta = 101.58^\circ$, 该晶体的结构原由 Koch 和 Germain 测定过^[12]。Schenk^[3] 指出这个结构在测定过程中发生了严重的相位退化,并用这

个结构作为检验 Σ_2P 方法的实例。他把这个结构当作属于中心对称空间群 $P2_1/m$ ，用 MULTAN 法求解。在 16 种相位组合的结果中可以挑选出最佳的一套。由此计算的 E 图同时包含真正的结构及其对映体。将这个 E 图同 Patterson 图作 4 次重迭就解出了真正的结构。

本文的试验从 Schenk 的对映 E 图出发。在这个 E 图中所有的峰都是成对出现的(见图 1)。因为晶胞中有 10 个独立的非氢原子，我们从 E 图中挑选出最大的 10 对峰以作为起始双解模型。其中只有 8 对同真实的原子相应，其余两对是鬼峰(图中有 \times 者)。把这 10 对峰统统当作氮原子，计算出 76 个衍射点的 α_H 和 $|\beta_H|$ 。从中挑选出 12 个较大的 $|\beta_H|$ ，用分量关系式求出它们的正负号，从而获得这些衍射点的相位 ϕ_H (表 1)。另外，又挑选出 5 个 $|E_H| \sim |\alpha_H|$ 的衍射点，令其 $\phi_H = \cos^{-1}[S(\alpha_H)]$ (表 2)。加上用于规定原点的三个衍射点(表 3)共 20 个衍射作为起始套，用正切式作相位扩展和修正获得了 76 个 $|E_H| \geq 1.4$ 的衍射点的相位，并算出相应的 E 图(图 2)。从图 2 可见，两个对映体中的一个几乎完全消失。此外，在最大的 10 个峰当中，只包含了一个鬼峰(图中有 \times 者)。这样，经过上述处理之后，不仅消除了中心对称性，而且使鬼峰的比例从出发时的 20% 降至 10%。显而易见，要从图 2 导出正确的结构是毫不困难的。

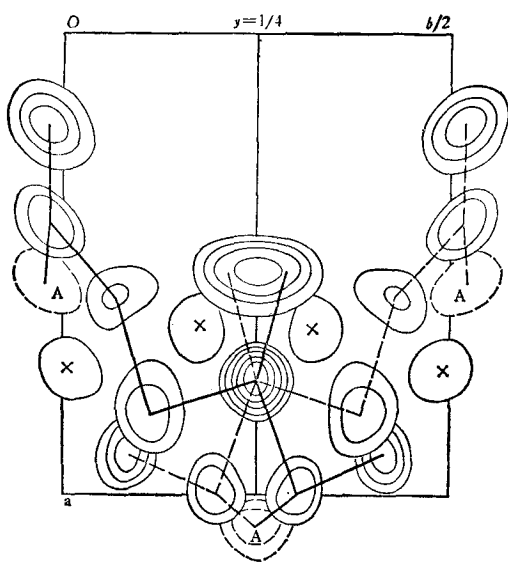


图 1

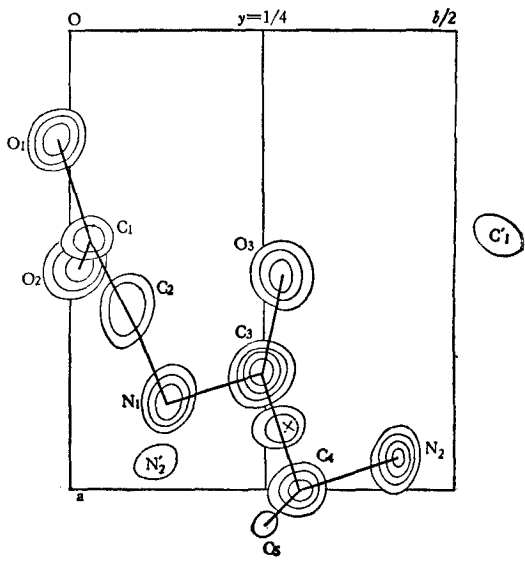


图 2

表 1

h	k	l	$\phi_H(^{\circ})$	h	k	l	$\phi_H(^{\circ})$
0	7	4	252	3	9	$\bar{1}$	273
1	7	4	108	3	10	$\bar{2}$	73
1	9	1	71	4	2	3	285
1	11	$\bar{1}$	70	4	7	$\bar{3}$	281
2	1	$\bar{5}$	105	5	1	1	81
3	7	$\bar{1}$	78	5	2	$\bar{2}$	262

表 2

h	k	l	$\phi_H(^{\circ})$
0	0	5	0
3	0	3	0
4	0	2	180
5	0	0	0
5	0	1	0

表 3

h	k	l	$\phi_H(^{\circ})$
1	0	4	180
2	0	$\bar{5}$	0
5	1	$\bar{2}$	180

五、双解结构模型的逐次电子密度逼近

本文的方法在使用分量关系式之前必须先获得各原子的双解坐标, 它们组成一个中心对称的双解结构模型。这个模型的好坏当然要影响到运用分量关系式的效果。因此在必要时可以先行用逐次电子密度逼近的办法改善原始的双解模型。与通常的电子密度逐次逼近不同, 双解结构模型的逼近只能使用一个近似的电子密度表达式:

$$\rho'(r) = \frac{1}{V} \sum_H S(\alpha_H) |F_H| \exp(-2\pi i H \cdot r). \quad (15)$$

如第二节所述, 这样的电子密度图将包含约占总贡献 36% 的鬼峰。但尽管如此, 实际试验表明用(15)式仍然可以改善原始的双解模型。

上节实例所用的原始模型中包含了 8 对双解原子和两对鬼峰。另有两对双解原子未归入原始模型之中(图中有 A 者) 从这个原始模型出发, 利用(15)式, 经过 4 轮逼近, 就排除了所有鬼峰并找回原来丢掉的两对双解原子(图 3)。这个试验说明用(15)式逐步改善原始的双解结构模型是可行的。但考虑到这样做就要显著地增加计算量, 而分量关系式一般地说又可以容许相当大的起始误差, 上节的试验就是明显的一例, 因此, 除非遇到特殊的困难, 原始双解模型的逐次逼近是不必要的。

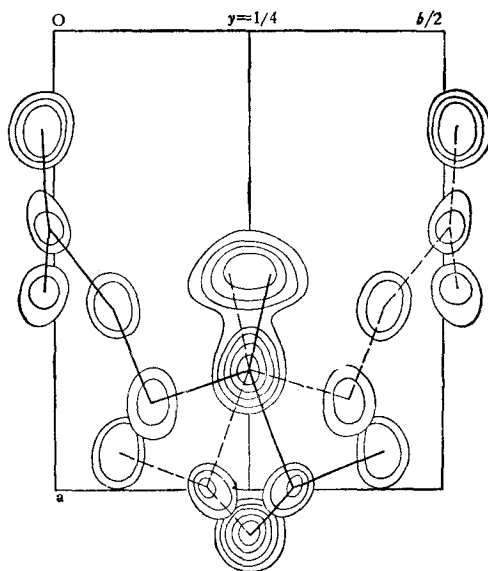


图 3

六、讨 论

本文的方法同 Σ_2P 方法的主要区别在于如何从混合两个对映体的 E 图中分离出其中一个对映体。 Σ_2P 方法并没有用直接法来解决这个问题。它虽然从直接法开始, 但等到获得一个混合的 E 图之后, 就把问题转交给 Patterson 法去解决。本文的方法则是一个以直接法贯串始终的方法。它用分量关系式在倒易空间中解决了 Schenk 用 Patterson 函

数在实空间中所解决的问题；证明了在直接法运用中所产生的相位退化问题可以从直接法本身找到更为直捷的解决途径。

相位退化问题也可以认为是对映体不稳定的问题。本文的方法所以能够有效地解决相位退化的问题是因为它包含了一项特殊的强化对映体选定的措施。这是通过分量关系式来实现的。它首先保证用于规定对映体的衍射点具有较大的 $|\beta_H|$ ，也就是保证所用衍射点对于对映体的选择是敏感的。另外，它不只是用一个衍射点来规定对映体。实际上，所选对映体的稳定性是由起始套中约 10 个较大的 $|\beta_H|$ 通过相互间的强分量关系来共同保证的。以上两点在 MULTAN 法中都是没有做到的。Hull^[4] 的 SWTR 法也具有稳定对映体的作用，但是它并不能保证起始套中的衍射对于对映体的选择是敏感的。而如果起始套本身对于对映体的选择并不敏感，SWTR 法就可能失效。Hauptman^[7] 用一组结构不变式的余弦来规定对映体。从理论上讲，它应该具有同本文方法一样的效果。但是，结构不变式余弦的数值在测定出正确的结构之前，只能用统计的方法近似地估算。这不但手续麻烦而且结果也不够可靠。相形之下， $|\beta_H|$ 的计算及 $S(\beta_H)$ 的推引则既简便又可靠。

本方法的另一特点是把推引衍射点相位的过程分解为推引实分量符号和推引虚分量符号两个阶段。这样可以大量节省时间。以 MULTAN 法为例，若在起始套中包含三个未知相位的衍射点，每一未知衍射取 4 种可能相位，其中一个兼用于规定对映体，则起始相位排列可有 32 种。同样条件下，若将结构因子的实分量和虚分量分别处理，则起始相位排列只有 8—12 种可能。这样，如果不要求减少相位排列的种数，就可以增加起始套中衍射点的数目。例如，在上例中将未知相位的衍射点数增至 4 个，将结构因子实、虚分量分别处理时的相位排列数不会超过 24 种。这对于解决复杂结构是很有好处的。有鉴于此，本文介绍的方法除用于处理相位退化问题以外，对于求解不发生相位退化的非中心对称晶体结构亦很值得一试。

郑启泰同志就本文内容同作者进行过有益的讨论，特此致谢。

参 考 文 献

- [1] G. Germain, P. Main and M. M. Woolfson, *Acta Cryst.*, **A27**(1971), 368; P. Main *et al.*, "MULTAN-80. A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data", Univ. of York, England.
- [2] G. M. Sheldrick, "Program for Crystal Structure Determination", Univ. of Cambridge, England.
- [3] H. Schenk, *Acta Cryst.*, **A28**(1972), 412.
- [4] E. Hull and M. J. Irwin, *Acta Cryst.*, **A34**(1978), 863.
- [5] 范海福、郑启泰, 物理学报, **27**(1978), 169.
- [6] 韩福森、范海福、古元新, 物理学报, **30**(1981), 111.
- [7] H. Hauptman and W. L. Duax, *Acta Cryst.*, **A28**(1972), 392.
- [8] 范海福, 物理学报, **21** (1965), 1114.
- [9] J. Karle, *Acta Cryst.*, **21**(1966), 273.
- [10] International Tables for X-ray Crystallography **4**, pp. 350—351.
- [11] I. L. Karle and J. Karle, *Acta Cryst.*, **16**(1963), 969.
- [12] M. H. J. Koch and G. Germain, *Acta Cryst.*, **B26**(1970), 410.

PROCEDURE FOR SOLVING THE PROBLEM OF PHASE DEGENERATION RESULTING FROM DIRECT METHODS

FAN HAI-FU QIAN JIN-ZI

(Institute of Physics, Academia Sinica)

ABSTRACT

A serious problem in the application of direct methods is the problem of phase degeneration. That is the phases of structure factors of a non-centrosymmetric crystal converge to a set of centrosymmetric phases. This would occur frequently in case of space groups $P2_1$, $P4_1$ etc. Phase degeneration gives rise to an additional centre of symmetry and hence a considerable number of spurious peaks in the resulting E -map leading often to the failure of direct methods. In this paper, a two step procedure of solving this ambiguity is proposed. In the first step, the crystal is treated as a centrosymmetric one and a mixed image including both enantiomorphs is obtained, while in the second step, the actual structure is then derived from the mixed image by making use of the so-called "component relations"^[1]. This method has been verified with a known structure and proved to be effective.