

多解法 MULTAN 失败原因的探讨

姚 家 星

(中国科学院物理研究所)

1983年1月25日收到

提 要

本文从三相角结构不变量的概率分布出发, 分析了多解法的根本缺点——小起始套是如何影响相角的推导和修正的, 从而说明程序系统 MULTAN 和 SHELX 中多解法失败的原因及其克服这一缺点的途径. 并简要说明随机多解法 RANTAN 的原理, 优点及其应用, 在四个结构实例中加以比较.

引 言

自从多解法 MULTAN 提出以来^[1], 在小分子单晶结构测定中的应用是十分有效和成功的, 很多单晶结构的解决都是使用了包括多解法的程序系统 MULTAN^[2] 和 SHELX^[3], 在一般情况下, 常无需人为的干预便可解出结构. 但在某些情况下, 对于并不太大的结构, MULTAN 却无能为力, 得不到正确的结构信息. 另外, 随着结构的复杂性增加, 则多解法往往更容易失败.

我们知道多解法的基础是三相角结构不变量, 其概率分布由下式计算^[4]:

$$P(\Phi_3) = \frac{1}{2\pi I_0(K)} \cdot e^{(K \cos \Phi_3)}, \quad (1)$$

式中

$$\Phi_3 = \varphi_h + \varphi_{h'} + \varphi_{h-h'}, \quad (2)$$

$$K = 2N^{-\frac{1}{2}} |E_h E_{h'} E_{h-h'}|, \quad (3)$$

N 等于单胞中的原子数目, $I_0(K)$ 为变型零阶贝塞耳函数. 图 1 给出了不同 K 值的三相

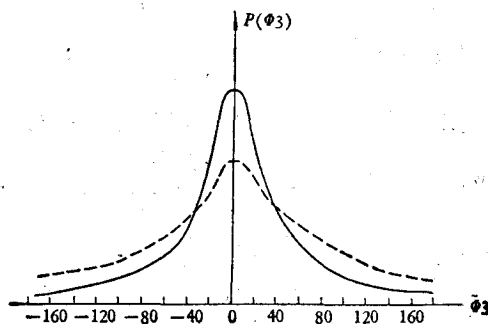


图 1 实线为 K_1 ; 虚线为 K_2 ; $K_1 > K_2$

角结构不变量的概率分布曲线,由此可以看出:

1) 概率分布的极大值在 $\Phi_3 = 0$ 处.

2) K 值越大,概率分布 $P(\Phi_3)$ 越尖锐,即当单胞中原子数目 N 越小和三个归一化结构振幅的乘积越大,则 $\Phi_3 = 0$ 的概率分布越高.在多解法中是用收敛图的方法来保证在推导相角的初期使用 K 值高的相角关系式,以得到较高的可靠性.

3) 无论 K 值的大小, $P(\Phi_3)$ 在 $\pm 180^\circ$ 附近总是有一定的概率分布,这部分的相角关系式的误差较大,由它推出的相角也具有相同的误差,往往造成多解法的失败.

在推导相角的初期使用个别相角关系式是多解法小起始套不可避免的缺陷, MAGEX^[5,6] 和 YZARC^[7] 等方法已将起始套衍射点的数目提高到 60 个左右,解结构的能力得到相应的提高.随机多解法 RANTAN^[8] 是解决小起始套缺点的另一个方法.我们知道在多解法中是允许起始相角具有 $\pm 45^\circ$ 的误差,那末,根据概率理论可得,一个在 $\pm 180^\circ$ 范围内均匀分布的随机变量在误差 $\pm 45^\circ$ 范围内的概率为 $0.25 (= 45^\circ/180^\circ)$.在 RANTAN 中给所有的衍射点(除确定原点和对映体的衍射外)赋以这样的随机变量作为起始相角值,并取它的概率值作为该相角的权因子(0.25),直接用正切公式对整套随机相角进行修正,根据计算的权函数对新的相角值进行取舍.这样不再使用个别相角关系式来确定相角,正切公式得到最好的满足,有效地克服了多解法 MULTAN 的缺点,从而解结构的能力显著提高,对于较复杂和较大的结构,例如包含 100 个独立原子的结构也获得了成功^[9].

下面从四个结构实例中,计算了相角关系式的分布,比较了 MULTAN 和 RANTAN 的结果,说明造成 MULTAN 失败的原因.在计算中我们认为 Φ_3 等于零的关系式是准确的.

相角关系式分布的计算

为了考查相角关系式在实际结构中的分布,并表示相角关系式的质量,引入以下品质因子:

1) 单一相角关系式的品质因子 M_i ,

$$M_i = \cos \Phi_3; \quad (4)$$

2) 一个衍射点所包含的关系式的品质因子 M_h

$$M_h = \sum_{h'} K_{h,h'} \cos \Phi_3 / \sum_{h'} K_{h,h'}; \quad (5)$$

3) 在结构测定中所使用的全部相角关系式的品质因子 M_s ,

$$M_s = \sum_h \sum_{h'} K_{h,h'} \cos \Phi_3 / \sum_h \sum_{h'} K_{h,h'}. \quad (6)$$

可以看出这三个品质因子的取值范围都是从 -1 到 +1. 如果等于 +1, 则表示相应的相角关系式的 Φ_3 都等于零. 具体计算步骤如下: 首先和结构测定相同, 建立相角关系式, 然后用已知的结构参数计算出各衍射点的正确相角值, 分别代入 (4), (5) 和 (6) 式, 得到品质因子 M_i , M_h 和 M_s .

计算结果及讨论

上述计算分别在四个结构实例中完成,表 1 给出了这些结构的参数、品质因子 M_s 及 MULTAN 和 RANTAN 的结果。表 2 给出两个中心对称结构的相角关系式的分布情况,用关系式的数目及百分比来表示。对于非中心对称结构的相角关系式的分布曲线,如图 2 所示。可以看出品质因子 M_s 反映了关系式分布的情况, M_s 值越高,相角关系式的误差越小。表 3 和表 4 给出了这些结构在 MULTAN 中使用的起始衍射的品质因子 M_h ,由此可以看出所选用的起始衍射的质量。

表 1

例	分子式	空间群	单胞分子数	使用的衍射数	关系式数目	品质因子 M_s	MULTAN	RANTAN
1	$C_{11}H_{16}NO_2PSCl_2$	P_1	2	272	2723	0.9968	成功	成功
2	$C_{10}H_{17}NO_3$	P_1	2	256	2741	0.8950	失败	成功
3	$C_{21}H_{28}O_3$	$P_{2_1}2_12_1$	4	204	3183	0.5223	成功	成功
4	$C_{22}H_{26}O_2$	P_{2_1}	4	292	3395	0.1710	失败	成功

表 2

例	品质因子 M_s	$\Phi 3 = 0^\circ$		$\Phi 3 = 180^\circ$	
		关系式数	百分比	关系式数	百分比
1	0.9968	2717	99.78	6	0.22
2	0.8950	2506	91.43	235	8.57

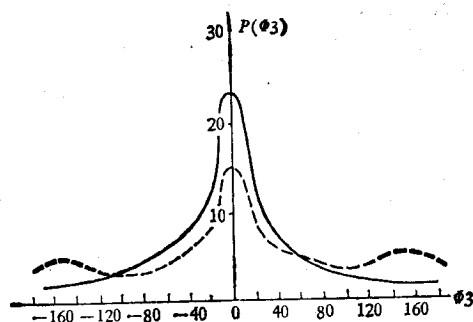


图 2 实线为例 3; 虚线为例 4。

由于 MULTAN 小起始套的缺点,使用个别或少量相角关系式来确定相角是不可避免的,因此较大的相角关系式的误差将给相角推导带来同样大的误差。对于某些像例 1 那样的结构,它们具有较高的品质因子 M_s ,相角关系式本身的误差就很小,所以小起始套的缺点并不显露出来, MULTAN 可以毫无困难地解出这些结构,而不论在起始套中选用什么样的衍射。随着品质因子 M_s 的降低,如例 2 和例 3 的情况,相角关系式的误差变

表 3 例 1 和例 2 的起始点及其品质因子 M_h

例 1				例 2			
序号	指标	关系式数	品质因子 M_h	序号	指标	关系式数	品质因子 M_h
1	0 5 1	66	1.0	1	1 0 2	169	1.0
5	4 5 0	67	1.0	3	8 -1 -1	45	1.0
7	4 5 -2	52	1.0	6	6 -4 3	34	1.0
11	3 -1 -5	34	1.0	7	2 1 -4	49	1.0
14	5 2 3	69	1.0	10	3 7 -4	32	1.0
19	4 -1 4	38	1.0	16	2 7 -3	34	0.926
49	2 -4 4	42	1.0	18	6 0 1	56	0.933
63	10 0 6	37	1.0	29	3 -3 3	57	0.849
				37	4 -6 3	42	1.0

表 4 例 3 和例 4 的起始点及其品质因子 M_h

例 3				例 4			
序号	指标	关系式数	品质因子 M_h	序号	指标	关系式数	品质因子 M_h
2	1 15 0	41	0.546	1	8 0 -9	47	0.282
12	6 1 0	55	0.750	2	1 4 0	71	0.713
16	1 4 4	79	0.594	15	7 0 -9	30	0.286
19	0 1 6	51	0.717	18	1 1 -1	82	0.044
21	5 8 6	57	0.488	19	6 0 -8	36	0.392
41	1 18 3	62	0.616	34	10 2 8	57	0.227
45	0 4 3	47	0.689	54	9 1 8	58	0.028
				56	16 2 0	47	0.308
				270	8 1 7	45	0.011

大,此时,在推导相角初期使用的相角关系式起着很关键的作用,起始衍射的选择直接影响到 MULTAN 的成功与否。所以当 MULTAN 失败后,还可以使用不同的起始衍射的方案,仍有可能获得成功。但由于在结构测定出以前,我们不可能知道相角关系式的分布情况,故这些尝试方案具有一定的盲目性,有时多次尝试仍不能成功,必然要花去大量的计算机时间。

如果结构的相角关系式的品质因子 M_h 很低,如例 4 的情况,这时误差大的相角关系式太多,无论尝试什么方案都不能避免使用这些相角关系式来确定相角,对这样的结构, MULTAN 便无能为力了。从实验结果可以看出, RANTAN 对这样的结构仍是有效的,因为它从来不使用个别和少量的关系式来确定相角。随着结构复杂性的增加,原子数目的增多,其相角关系式的分布变钝,品质因子 M_h 变小,所以 MULTAN 更容易失败,而 RANTAN 大起始套的优越性更为显著。

在 RANTAN 中对起始套的要求是和 MULTAN 不一样的,在 MULTAN 中,无论起始套的大小,均认为总能得到一套正确的起始相角,所以使用 1.0 作为起始相角的权因子,这样起始衍射的数目受到很大的限制。而在 RANTAN 中,不再要求起始相角是正确值,而是采用了与随机相角相适应的低权因子(0.25),故在起始衍射的数目上没有任何限制,大大提高了解结构的能力。另外,我们知道随机误差的修正要比系统误差的修正

更容易, 而且如果误差相同, 随机误差的相角套比系统误差的相角套拥有更多的结构信息^[10].

目前在中国科学院物理研究所的 IBM 370 计算机上, 除了安装有 MULTAN 的修改版本用来解超结构以外^[11], 还配有 RANTAN, MAGEX 和 YZARC 等程序供解一般结构和某些困难结构使用. 另外, 在美国 NICOLET 公司生产的四圆衍射仪所附的程序版本 SHELXTL 的 3.0 版本中也引入了 RANTAN 方法, 为解某些困难结构提供了方便.

本文所使用的数据(例 2, 例 3 和例 4)由 M. M. Woolfson 教授提供, 中国科学院物理研究所计算机室对计算工作给予了大力支持, 在此一并致以衷心的感谢.

参 考 文 献

- [1] G. Gernain & M. M. Woolfson, *Acta Cryst. B*, **34**(1968), 91.
- [2] P. Main, S. J. Fiske, S. E. Hull, L. Lessinger, G. Germain, J. P. Declercq & M. M. Woolfson, (1980), MULTAN-80, A System of Computer Programs for the Automatic Solution of Crystal Structure from X-Ray Diffraction Data. University of York, England.
- [3] G. M. Sheldrick, (1976), SHELX 76, A Program for Crystal Structure Determination. University of Cambridge, England.
- [4] W. Cochran, *Acta Cryst.*, **8**(1955), 473.
- [5] J. P. Declercq, G. Germain & M. M. Woolfson, *Acta Cryst. A*, **31**(1975), 367.
- [6] S. E. Hull, D. Viterbo, M. M. Woolfson & Zhang Shao-hui, *Acta Cryst. A*, **37**(1981), 566.
- [7] R. Baggio, M. M. Woolfson, J. P. Declercq & G. Germain, *Acta Cryst. A*, **34**(1978), 883.
- [8] Yao Jia-xing, *Acta Cryst. A*, **37**(1981), 642.
- [9] Yao Jia-xing, *Acta Cryst. A*, **39**(1983), 35.
- [10] A. M. Silva & D. Viterbo, *Acta Cryst. A*, **36**(1980), 1065.
- [11] Fan Hai-fu, Yao Jia-xing, P. Main & M. M. Woolfson, *Acta Cryst. A*, **39**(1983), 566.

A COMMENT ON THE REASONS FOR THE FAILURE OF MULTAN

YAO JIA-XING

(*Institute of Physics, Academia Sinica*)

ABSTRACT

MULTAN——multisolution tangent formula procedure is based on the probability distribution of three phase invariants. Based on this distribution, we analyzed how a small starting set, the main weakness of MULTAN, effects the process of phase development and refinement by the tangent formula. The reason of MULTAN failure and the approach to overcome the disadvantage of small starting set are discussed here. The principle and advantage of RANTAN——random MULTAN is elucidated briefly and compared with MULTAN in four examples.