

由重迭的穆斯堡尔谱求解超精细 参数分布的最大熵方法

徐祖雄 马如璋

(北京科学技术大学金属物理教研室)

1988 年 3 月 9 日收到

对于晶态或非晶态材料的严重重迭的穆斯堡尔谱, 本文导出了一种弱约束条件的最大熵拟谱方法. 求解的超精细参数分布具有正值性和整体光滑性. 通过放松约束可以改善局域光滑性. 最大熵方法使拟合结果最大限度地与已知信息符合, 而与未知信息无关, 从而保证了拟合结果的客观性和可靠性.

一、引 言

非晶态合金或某些晶态合金固溶体的穆斯堡尔谱具有重迭的展宽峰特点, 其超精细参数呈连续或准连续分布. 对于这种重迭的穆斯堡尔谱, 目前多采用基于最小二乘法的 Window 法^[1]和 Hess-Rübartsch 法^[2]或基于傅氏变换的 Vincze 法^[3]进行拟合. 上述方法均需适当的光滑或滤波处理, 以消除超精细参数分布曲线的不合理振荡. 然而, 光滑处理效果过大, 将使分布曲线上有价值的精细结构消失. Le Caer 和 Dubois^[4]在 Hesse-Rübartsch 方法基础之上, 曾提出非负几率做为附加的约束条件, 使拟合结果更为可信. 总之, 目前仍然存在这样一个问题有待进一步解决, 即如何从实验谱信息求得最客观和可信的超精细参数分布.

根据数据变换的 Jaynes 原则^[5], 由实测穆斯堡尔谱数据求解超精细参数分布, 应与全部已知信息相符合, 同时又最大限度地与未知信息无关. 本文提出一种最大熵拟谱方法, 试图用信息理论的最大熵原理在超精细参数分布的求解问题上实现 Jaynes 原则, 从而使拟谱结果在数据精度范围内更加可信. 有关最大熵原理及其在其他方面的应用, 可以参阅文献 [5—7].

二、方法原理

以下采用通用的拟谱基本假设: 只有一种超精细参数 x 的分布为主 (本文中均以超精细场 H_{hf} 为例), 忽略其他超精细参数分布的影响或仅考虑它们与超精细场 H_{hf} 的相关性^[2,4]. 若忽略厚度效应, 则透射穆斯堡尔谱可用下列理论公式表达:

$$Y(\nu) = E \int_0^{\infty} Ap(x)L(x, \nu)dx, \quad (1a)$$

$$\int_0^{\infty} p(x) dx = 1, \quad (2a)$$

式中的 v 为源与吸收体之间的相对速度, $L(x, v)$ 为归一化的洛伦兹型子谱函数, E 为谱线本底, $p(x)$ 为待求的超精细参数 x 的分布几率. 由 (1) 和 (2) 式可知 A 为谱线的总吸收面积, 因为

$$\begin{aligned} \int_{-\infty}^{\infty} [E - Y(v)] dv &= \int_0^{\infty} A p(x) \left[\int_{-\infty}^{\infty} L(x, v) dv \right] dx \\ &= A \int_0^{\infty} p(x) dx = A. \end{aligned} \quad (3a)$$

我们在 Hesse-Rübartsch 法的基础上建立最大熵拟谱方法. 设实测穆斯堡尔谱有 N 个实验点数据, 将超精细参数 x 的连续分布 $p(x)$ 用 M ($M \gg N$) 个分离的分布几率 $\{p_i\}$ ($i = 1, 2, \dots, M$) 代替, 则可用离散化的下列表达式来代替式 (1)–(3) 式

$$Y_k = E - \sum_{i=1}^M A_i L_{ki} \quad (k = 1, 2, \dots, N), \quad (1b)$$

$$\sum_{i=1}^M A_i = A, \quad (2b)$$

$$\sum_{k=1}^N (E - Y_k) = A, \quad (3b)$$

式中的 $A_i = A p_i$ ($i = 1, 2, \dots, M$), 它是第 i 个子谱的吸收面积. 不言而喻, $A_i \geq 0$ ($i = 1, 2, \dots, M$).

对于所研究的体系, 分布几率 $\{p_i\}$ 的熵为

$$S = - \sum_{i=1}^M A p_i \ln p_i, \quad (4a)$$

但为了计算方便, 我们采用子谱面积分布 $\{A_i\}$ 的熵

$$S = - \sum_{i=1}^M A_i \ln A_i. \quad (4b)$$

可以证明, 采用 (4a) 或 (4b) 式求得的结果是一致的. 按照最大熵原理, 可以求出在满足 (1b) 和 (2b) 式约束条件下使熵 S 取最大值的分布 $\{A_i\}$, 即用拉氏乘法求使

$$S_\lambda = S - \lambda_0 \left(\sum_{i=1}^M A_i - A \right) - \sum_{k=1}^N \lambda_k \left(Y_k - E + \sum_{i=1}^M A_i L_{ki} \right) \quad (5a)$$

取最大值时的分布 $\{A_i\}$.

上述最大熵法需要求解一个高达 $(N+1)$ 阶的非线性方程组^[7], 计算工作量过大, 而且, 从穆斯堡尔谱的分辨能力来说, 也没有必要划分如此众多 ($M > N$) 的子谱数.

实际上, 在有实验误差和 $p_i \geq 0$ 的条件下, (1b) 式不能严格成立, 应附加一个误差项而变为下列形式

$$Y_k = E - \sum_{i=1}^M A_i L_{ki} + \varepsilon_k \quad (k = 1, 2, \dots, N). \quad (1c)$$

因此, 应采用下述弱约束条件来代替 (1b) 式的强约束条件^[6]:

$$\chi_r^2 = \frac{1}{N_1} \left(\sum_{k=1}^N \varepsilon_k^2 / Y_k \right) = \alpha, \quad (6a)$$

式中的 N_1 为自由度, α 为一常数. 显然, α 若为零, (6) 式就等同于 (1b) 式的强约束条件; α 越大, 表示 (6) 式的约束程度越弱, 拟合的理论谱与实验谱之间的符合程度越差. 如果拟合的理论模型和选取的参数均完全正确, 则合理的 α 值应十分接近于 1^[6]. 对于重迭的穆斯堡尔谱, 采用下述弱约束形式似乎更为方便^[2]:

$$\chi_r^2 = \left(\sum_{k=1}^N \varepsilon_k^2 \right) / \left(\sum_{k=1}^N Y_k \right) = \alpha. \quad (6b)$$

以下, 我们将采用 (6b) 式形式的弱约束条件为例说明最大熵拟谱方法.

以 (6b) 式为约束条件, 则求解的 $\{A_i\}$ 应满足使

$$S_i = S - \lambda_0 \left(\sum_{i=1}^M A_i - A \right) - \lambda_1 [\chi_r^2 - \alpha] \quad (5b)$$

取最大值的条件. 由 $\partial S_i / \partial A_i = 0$ ($i = 1, 2, \dots, M$) 可知

$$-(1 + \ln A_i) - \lambda_0 - 2\lambda_1 \left(\sum_{k=1}^N \varepsilon_k L_{ki} \right) / \left(\sum_{k=1}^N Y_k \right) = 0 \quad (i = 1, 2, \dots, M). \quad (7)$$

(2b), (6b) 和 (7) 式就构成了求解 $\{A_i\}$, λ_0 和 λ_1 的基本非线性方程组, 可以用迭代法求解. 初值选取 $\lambda_1 = 0$, $\lambda_0 = -(1 + \ln A - \ln M)$, $A_i = A/M$ ($i = 1, 2, \dots, M$). 此时, 由于仅有两个约束条件和两个拉氏乘子 λ_0 和 λ_1 , 故 M 可以选取远小于 N 而又足够大的任意数值, 从而使得方程组的阶数和求解方程组的计算量大大减少. 这也是采用弱约束条件的一个优点.

由 (7) 式可知

$$A_i = \exp \left[-1 - \lambda_0 - 2\lambda_1 \left(\sum_{k=1}^N \varepsilon_k L_{ki} \right) / \left(\sum_{k=1}^N Y_k \right) \right] \quad (i = 1, 2, \dots, M). \quad (8)$$

将 (8) 式代入 (2b) 式可得

$$e^{1+\lambda_0} = \left[\sum_{i=1}^M e^{-2\lambda_1 \left(\sum_{k=1}^N \varepsilon_k L_{ki} \right) / \left(\sum_{k=1}^N Y_k \right)} \right] / A. \quad (9)$$

故

$$A_i = \frac{A \exp \left[-2\lambda_1 \left(\sum_{k=1}^N \varepsilon_k L_{ki} \right) / \left(\sum_{k=1}^N Y_k \right) \right]}{\sum_{i=1}^M \exp \left[-2\lambda_1 \left(\sum_{k=1}^N \varepsilon_k L_{ki} \right) / \left(\sum_{k=1}^N Y_k \right) \right]} \quad (i = 1, 2, \dots, M), \quad (10)$$

和

$$p_i = \frac{\exp \left[-2\lambda_1 \left(\sum_{k=1}^N \varepsilon_k L_{ki} \right) / \left(\sum_{k=1}^N Y_k \right) \right]}{\sum_{i=1}^M \exp \left[-2\lambda_1 \left(\sum_{k=1}^N \varepsilon_k L_{ki} \right) / \left(\sum_{k=1}^N Y_k \right) \right]} \quad (i = 1, 2, \dots, M). \quad (11)$$

可以证明, 以 (4) 式的熵表达式按同样的最大熵法直接求得的 $\{p_i\}$ 具有与 (11) 式完

全相同的形式。

三、计算实例

我们构造了一个具有实验统计误差的理论计算谱,其超精细场分布为

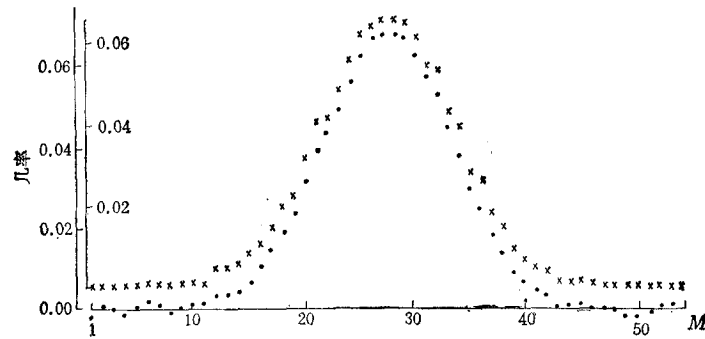


图1 理论计算谱拟合的超精细场分布 $50\text{kOe} \leq H_{hf} \leq 450\text{kOe}$, $M = 54$; 上面(整个曲线上升了一段距离)的由 x 点组成的分布曲线为最大熵结果, $\alpha = 1.2$; 下面的由圆点组成的分布曲线为 Hesse-Rübarsch 结果, 光滑因子为 1.0

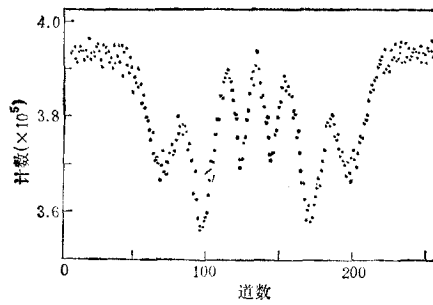


图2 $\text{Fe}_{39}\text{Ni}_{39}\text{B}_{12}\text{Si}_8\text{Mn}_7$ 非晶条带的室温穆斯堡尔谱

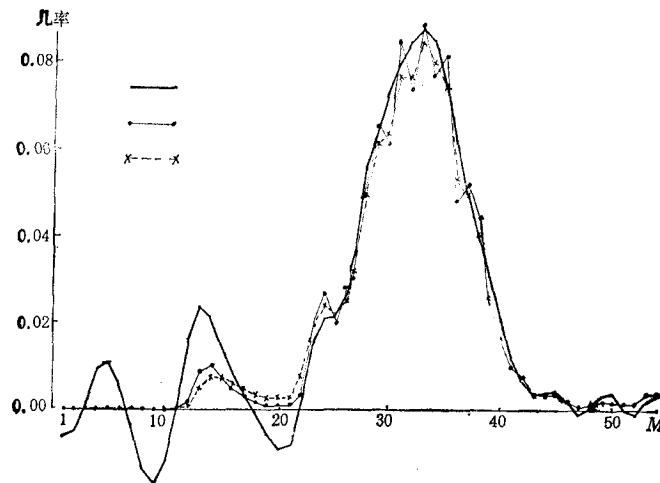


图3 图2谱线拟合的超精细场分布 $0 \leq H_{hf} \leq 400\text{kOe}$, $M = 54$; 粗实线为 Hesse-Rübarsch 结果, 光滑因子 $G = 20$, $\chi^2 = 1.63$; 细实线为 $\alpha_0 = 1.9$ 的最大熵结果; 虚线为 $\alpha = 2.07$ 的最大熵结果

$$P(H_{\text{hf}}) = c \cdot \exp \left[- \frac{(H_{\text{hf}} - 250)^2}{4000} \right] \quad (50\text{kOe} \leq H_{\text{hf}} \leq 450\text{kOe}), \quad (12)$$

平均同质异能移位和四极分裂均为零, 子谱各峰的半高处全宽为 0.19126 mm/s . 拟合时的 H_{hf} 取值范围为 $50-450 \text{ kOe}$, $M = 54$. 图 1 绘出了用 Hesse-Rübartsch 法和用本文最大熵法拟合的超精细场分布 $P(H_{\text{hf}})$, 最大熵法拟合时选取 $\alpha = 1.2$. 由图 1 可知, 两种方法拟合的结果十分接近, 与 (12) 式的差异亦很小.

图 2 为铁磁性非晶态合金 $\text{Fe}_{39}\text{Ni}_{39}\text{B}_{12}\text{Si}_8\text{Mn}_2$ (径 200°C 下 1 h 退火处理) 的室温透射 ^{57}Fe 穆斯堡尔谱, 图 3 绘出了用改进的 Hesse 法^[8]和最大熵法拟合结果的对比. 由图 3 可知, 两种拟合结果有比较明显的差别, 显然, 最大熵法的结果更为可信.

四、讨 论

(2b) 和 (6b) 式包含了实测谱的全部已知信息. 若在满足 (2b) 式条件下, $\{A_i\}$ 与其他已知信息完全无关, 即完全不考虑 (6b) 式的约束, 则根据最大熵条件, $\{A_i\}$ 必为均匀分布. 这正是拟合时的初值状态. 随着迭代拟合的 χ^2 值接近于 α 值, $\{A_i\}$ 也随之变化, 以便与已知的实测谱信息相对应, 同时又最大限度地与未知信息无关, 从而使拟合结果具有客观性. 从图 1 和图 3 可知, 尽管没有任何光滑处理, 但最大熵法消除了 $\{A_i\}$ 的负值和不合理的振荡, 使拟合结果更加可信. 例如, 图 3 中低场端小峰成为不容置疑的客观结果.

显然, α 取值表征了 $\{A_i\}$ 与已知实验谱相对应的程度, 即 α 表征了已知信息的约束程度. 小 α 值的过强约束是不合理的, 将使拟合过程难于收敛到一个完全满足最大熵条件的结果. 例如, 在图 1 中取 $\alpha = 1.0$ 时, 经多次迭代只能得到 $\chi^2 = 1.2$ 的拟合结果. 客观上存在一个最小的合理 α 值(以下用 α_0 表示). 显然, 由于 $\{A_i\}$ 的非负限制, 使最大熵法拟合的 χ^2 要大于 Hesse-Rübartsch 法的 χ^2 . 虽然在理论上, α_0 值应接近 1, 但对于实测穆斯堡尔谱, 由于测量的非线性等因素以及拟谱的理论模型和假设的不尽完善性, 使得 α_0 值往往比 1 大得多(见图 3 的例子). 显然, α_0 值越偏离于 1, 则表明上述的不完善性越大. 由于实验数据的统计误差, 实际上选取 $\alpha = \alpha_0$ 并不是最好的选择. 这将在下面的讨论中进一步得到阐述.

从图 1 可知, 对于理论计算谱, Hesse-Rübartsch 的结果在分布的两端均有围绕着零值的微小起伏, 而在主峰处比较光滑; 最大熵结果则完全消除了分布两端的不合理振荡, 但在主峰处的局域光滑性稍差一点. 这表明最大熵法具有使 $\{A_i\}$ 取正值和整体光滑作用^[6], 从而消除了 $\{A_i\}$ 的不合理振荡. 因此, 最大熵法是一种可以信赖的拟谱方法. 上述对比亦表明, 对于一个理想的重迭谱, 只要拟合理论模型和参数选取正确, 则最大熵法和 Hesse-Rübartsch 法的拟合结果十分接近.

由图 3 可知, 实测谱的 Hesse-Rübartsch 结果有明显的不合理性, 这主要表现在低场端的振荡太大. 低场端的负值峰高已达主峰高的六分之一. 显然, 低场端的正值峰亦不能确认为可靠的拟合结果. 这种振荡不能靠扩大光滑因子来消除(图 3 中光滑因子 $G = 20$, 已足够大了). 若选取拟合范围为 $H_{\text{hf}} = 83-400\text{kOe}$, 则因为实际上已令 $0-83\text{kOe}$

范围内的 $\{A_i\} \equiv 0$, 拟合的负值峰和正值小峰均明显减小而接近于最大熵结果. 出现上述问题的根本原因在于拟谱基本假设的近似性. 与此相比, 尽管最大熵法拟合的 χ^2 较 Hesse-Rübartsch 法的 χ^2 大, 但其结果更为可靠, 特别是可以确认低场端小峰的存在是十分合理的. 在图 3 中, 相应于 $\alpha_0 = 1.9$ 的拟合结果在主峰处有较大的起伏. 这表明最大熵法虽有整体光滑作用, 但对应于 α_0 的结果的局域光滑性较差. 这种起伏有一部分来源于实验数据的统计误差. 由于相应于 α_0 的拟合结果既满足了已知的全部实测信息, 又最大限度地与未知信息无关, 故它是最客观的结果. 因此, 如果能对同一样品重复多次测量, 对每个测量谱用最大熵法求解, 则多个最大熵求解结果的平均值将能消除统计误差的影响, 给出最客观又最接近于真实分布的估计.

实际上, 人们往往先验地认定超精细参数的分布是光滑的. 对于这种定性的先验信息, 我们可以通过放松 (6b) 式的约束程度, 即增大 $\alpha (> \alpha_0)$ 来改善最大熵法拟合结果的局域光滑性. 图 3 中 $\alpha = 2.07$ 的拟合结果表明, 主峰处的局域光滑性得到改善, 而低场端和高场端的分布以及主峰处的整体分布仅有很小的变化. 若取 $\alpha = 2.3$, 则主峰处的局域光滑性又进一步改善. 显然, 过大的 α 值, 尽管拟合结果的局域光滑性很好, 但由于计算谱与实测谱的符合程度太差, 使拟合结果的可信度大大下降. 考虑到 α 取值较大时, 迭代求解过程收敛较快, 故实际拟合时选取比 α_0 适当大的 α 值似乎更为合适. 这样, 既可改善局域光滑性, 又能保证两者的差异均在实验误差范围内. 采用其他局域光滑处理亦能改善最大熵法结果在主峰处的局域起伏. 但上述这些改善局域光滑性的措施, 或多或少地使某些参数 (例如 α) 依赖于更多的主观因素, 从而使拟合结果的客观性下降.

总之, 无论是对理论谱的十分理想的拟合, 还是对于实测谱的不尽完善的拟合, 最大熵法能给出更为客观和可信的结果. 此外, 正是因为最大熵法的局域光滑作用较差, 故其拟合结果应更有利于反映超精细参数分布的精细结构. 例如, 图 3 中主峰处左侧的伴线峰和其他精细结构、以及两端 (特别是低场端) 的分布细节, 在放松约束 ($\alpha = 2.07$) 的最大熵结果中仍十分明显. 这就为由超精细参数分布中获取非晶态合金的结构信息, 提供了更为有利的实验结果.

五、结 束 语

对于重迭的穆斯堡尔谱来说, 采用弱约束条件的最大熵法拟谱结果, 具有取正值和整体光滑性的优点, 从而消除了不合理的振荡, 使拟合结果更为客观和可信. 这些优点使最大熵方法成为一种具有实用价值的更为可靠的拟谱方法.

- [1] B. Window, *J. Phys. E*, **4**(1971), 401.
- [2] J. Hesse and A. Rübartsch, *J. Phys. E*, **7**(1974), 526.
- [3] I. Vincze, *Nucl. Instrum. and Methods*, **199**(1982), 247.
- [4] G. Le Caer and J. M. Dubois, *Phys. Stat. Sol. (a)*, **64**(1981), 275.
- [5] E. T. Jaynes, *Phys. Rev.*, **106**(1957), 620.
- [5] S. W. Wilkins, J. N. Varghese and M. S. Lehman, *Acta Cryst.*, **A39**(1983), 47.
- [7] 魏文铎, *物理学报*, **35**(1986), 171.
- [8] 徐祖雄、马如璋、李士, *科学通报*, **30**(1985), 710.

MAXIMUM ENTROPY EVALUATION OF OVERLAPPED MÖSSBAUER SPECTRA

XU ZU-XIONG MA RU-ZHANG

(Department of Metall Physics, Beijing University of Science Technology)

ABSTRACT

A calculation scheme for the maximum entropy evaluation of the distribution of hyperfine parameters from overlapped Mössbauer spectra is proposed, which is subjected to the weak constraint condition. Maximum entropy evaluation is consistent with all given informations and has little correlation with the indefinite informations. The fitted distribution is characteristic of its positivity and bulk smoothness. Its local smoothness can be improved with relaxation of the constraint. The maximum entropy evaluation proved a more reliable method for fitting overlapped Mössbauer spectra.