第1卷 第6期 光 学 学 报 Vol. 1, No. 6
1981年11月 ACTA OPTICA SINICA November 1981

# 关于光学显微术和电子显微术中的相位恢复问题

顾本源 杨国桢
(中国科学院物理研究所)

提 要

本文通过较严格的数学推导，得到一组由已知成像系统的两个平面上的光波强度分布去恢复相位的 Gerchberg-Saxton 算法[1,2,4]和 Misell 算法方程[3]。本文指出，只要成像系统的任何某两个平面上的波函数之间的变换是通过一个幺正算符实现的，那末，由已知它们的振幅，采用 GS 算法便可以实现相位复原。

## 一、引 言

在物理学的许多领域中都存在着一个经典问题，即如何由测得的光(电子)波的强度分布或散射微分截面去找回丢失了的相位信息，重建波函数，这就是熟知的所谓相位恢复问题。例如，在X-射线结构分析中，只能测定结构因子的绝对值大小，相位信息明显地丢失了；在散射问题中，由散射强度分布(微分截面)只能给出散射振幅的绝对值，但是其相位资料在确定散射物体结构时，是必不可少的；当使用光学显微术和电子显微术确定物体的结构时，出现类似的问题，人们只能直接地测得成像系统中像平面或某些其它平面上的强度分布，从而只知道这些平面上光(电子)波函数的振幅。为了确定物体的结构，必须获得波函数的相位信息。由此可见，相位恢复问题是很有意义和十分重要的问题。

关于光学显微术和电子显微术中的相位恢复问题，许多学者提出了各种复原方案和进行了详细讨论[1~13]。1972年 Gerchberg-Saxton 提出一种计算机算法[1,2,4]，由已知像平面和衍射平面(或称作出射光瞳)上的强度分布计算出它们的相位(以后简称为GS算法)。1973年 Misell 仿照 GS 算法，由已知两个具有不同离焦值的离焦像平面上的强度分布计算波函数的相位(Misell 算法)，详细地讨论了计算中所遇到的各种问题。这些算法在许多情形下都是切实可行的。根据这些算法，可以由已知成像系统的某两个平面上的强度分布有效地计算出波函数的相位。但是，他们都没有给出这些算法的严格的数学依据。本文从更一般情形下，采用严格的数学推导，指出只要成像系统的任何某两个平面上的波函数之间的变换是通过一个幺正算符实现的，那末，由已知它们的振幅分布，就可以应用 GS 算法求得相位分布。

第二节给出有关相位恢复问题的基本方程；第三节通过严格的数学推导，得到 GS 算法方程式，指出 GS 算法的普遍适用性；最后一节还讨论了算法存在的问题。

收稿日期：1980年3月19日

## 二、相位恢复问题中的基本方程

我们考虑一个物体,用相干辐射进行照明,它的一般成像过程如图1所示。由于成像过程是线性的以及任意波形可表示成平面波的相干迭加形式,因此我们可以假定入射波为单色平面波(波长为 $\lambda_0$),沿光轴入射,这样做不会失去问题的普遍性。为简化符号起见,我们只讨论一维情形,所有结果可直接推广到二维。

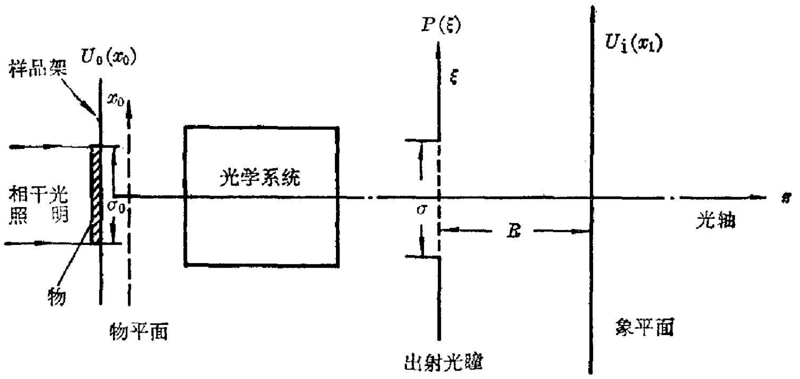


图1 成像过程的示意图($R$ 表示高斯参考球的半径)

Fig. 1 Diagram of imaging procedure ($R$—radius of Gaussian reference sphaere)

我们将直接紧靠物体的一个平面称作"物平面",其座标记作 $x_0$,这个平面上的波函数 $U_0(x_0)$ 称作物波函数。在光学和电子显微术中,所研究的物体的结构信息都包含在物波函数中,它是辐射场与组成物体的原子或分子相互作用的直接反映。一般说来,$U_0(x_0)$ 还不能直接被测得,通常要通过成像系统来观测,所以实际上人们测得的是成像系统的另一些平面上的强度分布,这些平面通常是出射光瞳面(或称衍射平面,座标记作 $\xi$,光阑孔径 $\sigma$),像平面(座标记作 $x_1$)和离焦像平面(座标记作 $x_2$,离焦量 $\Delta f$);与上述各平面相应的波函数分别记作 $P(\xi)$,$U_i(x_1)$ 和 $U_i'(x_2)$。根据显微镜的成像理论[14,15],在傍轴近似下,物波函数和像波函数之间通过脉冲响应 $K(x_1, x_0)$ 联系起来:

$$U_i(x_1)=\int_{\sigma_0} dx_0 K(x_1, x_0) U_0(x_0)=\mathbf{K}U_0, \tag{1}$$

式中 $\sigma_0$ 代表物平面上的有效孔径。脉冲响应是

$$K(x_1, x_0)=\int_{\sigma} d\xi \exp\{2\pi j[\phi(x_0, \xi)-(x_1-x_0)\xi]\}, \tag{2}$$

式中 $\phi(x_0, \xi)$ 是出射光瞳像差函数(测量单位为弧度),$\xi$ 的度量单位是 $f$(焦距),$x_0$ 用辐射波长 $\lambda_0$ 度量,$x_1$ 的度量单位为 $\lambda_0 M$,$M$ 表示仪器的放大率。

出射光瞳波函数 $P(\xi)$ 与像波函数 $U_i(x_1)$ 之间互为傅里叶变换关系:

$$U_i(x_1)=\int d\xi P(\xi)\exp\{-2\pi j\xi x_1\}=\mathbf{F}P, \tag{3a}$$

$$P(\xi)=\int dx_1 U_i(x_1)\exp\{2\pi j\xi x_1\}=\mathbf{F}^{-1}U_i, \tag{3b}$$

这里的 $P(\xi)$ 可写成:

$$P(\xi)=\int_{\sigma_0} dx_0 U_0(x_0)\exp\{2\pi j[\phi(x_0,\ \xi)+x_0\xi]\}。\tag{4}$$

现在考虑离焦像情形，此时出射光瞳上出现一个附加的波象差：

$$\Delta f\xi^2/2\lambda_0,$$

其中 $\Delta f$ 为离焦量。于是，出射光瞳波函数变成

$$P'(\xi)=P(\xi)\exp\{j\pi\Delta f\xi^2/\lambda_0\},\tag{5a}$$

离焦像的波函数 $U'_i(x_2)$ 可表示成

$$U'_i(x_2)=\mathbf{F}P'。\tag{5b}$$

由(3a)，(3b)，(5a)和(5b)可得到 $U'_i(x_2)$ 与 $U_i(x_1)$ 之间关系式：

$$U'_i=\mathbf{G}U_i,\tag{6a}$$

$$U_i=\mathbf{G}^{-1}U'_i,\tag{6b}$$

其中 $\mathbf{G}$ 和 $\mathbf{G}^{-1}$ 算符为

$$\mathbf{G}=\mathbf{F}[\exp(j\pi\Delta f\xi^2/\lambda_0)\mathbf{F}^{-1}],\tag{7a}$$

$$\mathbf{G}^{-1}=\mathbf{F}[\exp(-j\pi\Delta f\xi^2/\lambda_0)\mathbf{F}^{-1}],\tag{7b}$$

容易证明(6a)～(7b)公式完全等价于 Misell 所导出的方程[3]。这里，我们并不要求脉冲响应具有空间平移不变性。

在光学显微术和电子显微术中的相位恢复问题可分为两个步骤：第一步，由测量一张或两张图像的强度分布去求复波函数 $P(\xi)$；第二步，由 $P(\xi)$ 计算 $U_0(x_0)$。这里，第一步工作最为困难，本文着重讨论，而第二步工作已有许多学者作过讨论[13]，本文不准备涉及。

## 三、GS 算法和 Misell 算法的数学依据

1972 年 Gerchberg-Saxton 提出由已知像平面和出射光瞳上的强度分布，用计算机迭代求解算法求出相位分布[1,2,4]。1973 年 Misell 仿照 GS 算法提出由已知两个具有不同离焦值的离焦象的强度分布计算其相位的计算机算法[3]。

GS 算法的迭代方程式是

$$U_i^{(n)}(x_1)=\int d\xi P^{(n-1)}(\xi)\exp\{-2\pi jx_1\xi\},\tag{8a}$$

$$P^{(n)}(\xi)=\int dx_1 U_i^{(n)}(x_1)\exp\{2\pi jx_1\xi\}。\tag{8b}$$

我们把波函数写成

$$P(\xi)=|P(\xi)|\exp\{j\phi_p(\xi)\},$$

$$U_i(x_1)=|U_i(x_1)|\exp\{j\phi_i(x_1)\}。$$

GS 算法的步骤是这样的：最初任意给定 $\phi_p^{(0)}(\xi)$ 初值(例如，在 $-\pi$ 和 $+\pi$ 之间无规取值)，利用观测值 $|P(\xi)|$ 构成试探函数 $P^{(0)}(\xi)=|P(\xi)|\exp\{j\phi_p^{(0)}(\xi)\}$，代入 (8a) 式，算得 $U_i^{(1)}(x_1)$，只保留它的相位 $\phi_i^{(1)}(x_1)$。令它与测量得到的 $|U_i(x_1)|$ 组成一级近似波函数 $U_i^{(1)}=|U_i(x_1)|\exp\{j\phi_i^{(1)}(x_1)\}$，代入(8b)式计算得到 $P^{(1)}(\xi)$，同样地只保留相位 $\phi_p^{(1)}(\xi)$，再与 $|P(\xi)|$ 组成 $P^{(1)}(\xi)=|P(\xi)|\exp\{j\phi_p^{(1)}(\xi)\}$，代入(8a)去计算 $U_i^{(2)}$；如此迭代下去，直到 $|U_i^{(n)}(x_1)|$ 与观测值 $|U_i(x_1)|$ 之间平均绝对平方误差小到某一给定值为止。

Misell 算法则是利用(6a)和(6b)，算法步骤完全与 GS 算法类似。他们都没有给出算

法的严格的数学依据。

现在我们从更普遍的观点来讨论成像过程[16,17]，并给出GS算法和Misell算法较严格的数学推导。成像系统中任意两个平面上的波函数$U_1$和$U_2$是通过线性变换$\mathbf{H}$相联系的：

$$U_2=\mathbf{H}U_1。\tag{9}$$

如果辐射源总功率是有限的，那末$U_1$是绝对平方可积的，线性变换$\mathbf{H}$将$U_1$变为另一绝对平方可积函数$U_2$。在成像系统中，通常感兴趣的是无损耗过程，即

$$\int |U_1|^2dx_1=\int |\mathbf{H}U_1|^2dx_2。\tag{10}$$

由于上式对于所有$U_1\in L^2$（函数空间）都成立，这就要求

$$\mathbf{H}\cdot\mathbf{H}^+=\mathbf{H}^+\cdot\mathbf{H}=\mathbf{I},\tag{11}$$

上式表明$\mathbf{H}$必须是幺正算符，即$\mathbf{H}^+=\mathbf{H}^{-1}$。

我们认为在光学和电子光学中，由已知两个平面上强度分布去恢复相位的问题，可以采用更一般的提法：如果成像系统中任意的某两个平面上的波函数

$$U_1(x_1)=|U_1(x_1)|\exp\{j\phi_1(x_1)\} \text{ 和 } U_2(x_2)=|U_2(x_2)|\exp\{j\phi_2(x_2)\}$$

是通过一个幺正算符$\mathbf{H}$联系起来的，即

$$U_2=\mathbf{H}U_1 \quad \text{或} \quad |U_2|=\exp\{-j\phi_2\}\mathbf{H}U_1,\tag{12a}$$

$$U_1=\mathbf{H}^{-1}U_2 \text{ 或 } |U_1|=\exp\{-j\phi_1\}\mathbf{H}^{-1}U_2,\tag{12b}$$

那末由已知其振幅$|U_1|$和$|U_2|$，可以用迭代算法恢复相位分布。

现在来证明上述结论。假定$U_1(x_1)$和$U_2(x_2)$都是带限函数（在物理上这是合理的），那末可以取分立值表示。在选定的座标系（表像）下，(12a)和(12b)可表示成有限维的矩阵形式。具体地说，如果$U_1(x_1)$和$U_2(x_2)$的取样点数目均为$N$，则$|U_1(x_1)|$和$|U_2(x_2)|$相应于$N$的单列矩阵，$\mathbf{H}$和$\mathbf{H}^{-1}$为$N\times N$的矩阵，$\exp(-j\phi_1)$和$\exp(-j\phi_2)$为$N\times N$的对角矩阵。所谓恢复相位问题，就是要求找到这样的$\phi_1$和$\phi_2$，使得$\exp\{-j\phi_2\}\mathbf{H}\cdot[|U_1|\exp\{j\phi_1\}]$尽可能地逼近$|U_2|$。两个矩阵$\mathbf{A}$, $\mathbf{B}$的逼近程度，用“距离”$D$这个量来描述，$D$的定义是[16,17]：

$$\begin{aligned}D^2&=\|\mathbf{A}-\mathbf{B}\|^2=\frac{1}{N}\sum_{l,m=1}^{N}|(\mathbf{A})_{lm}-(\mathbf{B})_{lm}|^2\\&=\frac{1}{N}T_r\{\mathbf{AA}^++\mathbf{BB}^+\}-\frac{2}{N}\operatorname{Re}T_r(\mathbf{AB}^+)。\end{aligned}\tag{13}$$

$T_r$表示矩阵的阵迹。如果$D=0$，则$\mathbf{A}$与$\mathbf{B}$的所有矩阵元都相同；如果$D$不为零，则$\mathbf{A}$与$\mathbf{B}$的矩阵元之间平均差别为$D/\sqrt{N}$。在相位恢复问题中，

$$\mathbf{A}=|U_2|,\quad \mathbf{B}=\exp\{-j\phi_2\}\mathbf{H}[|U_1|\exp\{j\phi_1\}]$$

均为单列矩阵，要求$\mathbf{B}$逼近$\mathbf{A}$，距离$D$为

$$\begin{aligned}D^2&=\||U_2|-e^{-j\phi_2}\mathbf{H}U_1\|^2\\&=\frac{1}{N}\{T_r[|U_2|^2]+T_r[|U_1|^2]-2\operatorname{Re}T_r[e^{-j\phi_2}\mathbf{H}U_1|U_2|^+]\}\\&=\frac{1}{N}\{T_r[|U_2|^2]+T_r[|U_1|^2]-2\operatorname{Re}T_r[e^{-j\phi_1}|U_1|^+\mathbf{H}U_2]\},\end{aligned}\tag{14}$$

在这里,我们利用了矩阵阵迹的轮换不变性。真正的相位分布,原则上可以由解下面的变分方程得到,即

$$\delta_{\phi_1,\phi_2}\| |U_2| - e^{-j\phi_2}\mathbf{H}U_1 \| = 0, \tag{15}$$

式中 $\delta_{\phi_1,\phi_2}$ 包含对 $\phi_1$ 和 $\phi_2$ 求微分,因此(15)式实际上是 $2N$ 个方程组,当 $\phi_1(x_1)$ 和 $\phi_2(x_2)$ 为连续函数时就是变分方程。我们通过以下讨论,可以比较简单地得到(15)式的解。要做到最佳逼近,则要求 $D^2$ 最小,即要求

$$Q = \frac{2}{N}\mathrm{Re}\sum_{l=1}^{N}(e^{-j\phi_2})_u(\mathbf{H}U_1|U_2|^+)_u = \frac{2}{N}\mathrm{Re}\sum_{l=1}^{N}(e^{-j\phi_1})_u(|U_1|^+\mathbf{H}^+U_2)_u \tag{16}$$

最大。只有当求和之中每一项都取绝对值时, $Q$ 才取最大值,于是有:

$$e^{j\phi_2} = \mathbf{H}|U_1|e^{j\phi_1}/ABS(\mathbf{H}|U_1|e^{j\phi_1}), \tag{17a}$$

$$e^{j\phi_1} = \mathbf{H}^+|U_2|e^{j\phi_2}/ABS(\mathbf{H}^+|U_2|e^{j\phi_2}), \tag{17b}$$

此外, $\mathbf{H}$ 为幺正算符:

$$\mathbf{H}^+ = \mathbf{H}^{-1}, \tag{17c}$$

其中 $ABS$ 表示取绝对值。方程组 (17a)～(17c) 正是 GS 算法的最普遍表示式。如果令 $U_1=P$, $U_2=U_i$, $\mathbf{H}=\mathbf{F}$, (17a)～(17c) 正是 GS 算法, 而令 $U_1=U_i$, $U_2=U_i'$, $\mathbf{H}=\mathbf{G}$, 则为 Misell 算法。

## 四、讨　　论

本文从更普遍的角度,讨论了成像过程和相位恢复问题,通过较严格的数学推导,得出 GS 算法和 Misell 算法方程。我们指出,如果成像系统的任意某两个平面上的波函数是由一个幺正算符联系起来的,那末由已知这两个波函数的振幅分布可以应用 GS 算法恢复其相位。GS 算法和 Misell 算法仅仅是上述结论的两个特例。

使用计算机迭代算法求相位分布有两个关键性的问题:其一是这种迭代算法是否收敛?收敛的速度如何?其二是所求得的解是否唯一,与初始值有否依赖关系?关于由已知两个强度分布求相位的解的唯一性问题,已有不少论文做了详细的讨论[6,7,10],在已知相位解唯一存在的条件下, GS 算法是十分有用的。

Gerchberg-Saxton-Misell 曾经讨论了这种迭代算法的收敛问题[1~4]。但是,从本文讨论看来, GS 和 Misell 算法只能做到 $D$ 最小,并不能保证任何情况下都能够使 $D\to 0$,对于某些波函数, $D$ 可能在取某一较小值之后,再继续迭代下去时,不再减小,这时只好采用各种数学计算技巧来继续逼近下去。

假如系统不满足能量守恒条件(10),那末就不必要求 $\mathbf{H}$ 是幺正算符。例如,当系统中存在着衍射损耗,光学元件存在着吸收时,只要 $\mathbf{H}$ 的复共轭算符 $\mathbf{H}^+$ 是存在的,那末仍然可以应用 GS 算法来恢复相位。

### 参　考　文　献

[1] R. W. Gerchberg, W. O. Saxton; *Optik*, 1972, **35**, No. 2 (Apr), 237.
[2] R. W. Gerchberg, W. O. Saxton; *J. Phys.* (*D*), 1973, **6**, No. 5 (20 Mar), L31.
[3] D. L. Misell; *J. Phys.* (*D*), 1973, **6**, No. 1 (Jan), L6.

D. L. Misell; *J. Phys.* (*D*), 1973, **6**, No. 18 (5 Dec), 2200.
D. L. Misell; *J. Phys.* (*D*), 1973, **6**, No. 18 (5 Dec), 2217.

[4] R. W. Gerchberg, W. O. Saxton; «*Image Processing and Computer-aided Design in Electron Optics*», 66, edited by P. W. Hawkes (1973).

[5] R. W. Gerchberg, W. O. Saxton; *Optik*, 1971, **34**, No. 3 (Dec), 275.

[6] A. M. J. Huiser, A. J. J. Drenth *et al.*; *Optik*, 1976, **45**, No. 4 (Jul), 303.
A. J. J. Drenth, A. M. J. Huiser *et al.*; *Optica Acta*, 1975, **22**, No. 7 (Jul) 615.

[7] A. M. J. Huiser, H. A. Ferwerda; *Optik*, 1976, **46**, No. 4 (Dec), 407.
A. M. J. Huiser, H. A. Ferwerda; *Optica, Acta* 1976, **23**, No. 6 (Jun), 445.

[8] P. Van Toorn, H. A. Ferwerda; *Optik*, 1977, **47**, No. 2 (Feb), 123.
P. Van Toorn, H. A. Ferwerda; *Optica Acta*, 1976, **23**, No. 6 (Jan), 457, 469.

[9] A. M. J. Huiser, P. Van Toorn *et al.*; *Optik*, 1977, **47**, No. 1 (Jan), 1.

[10] H. A. Ferwerda, B. J. Hoenders *et al.*; *Photogra. Sci. & Engng*, 1977, **21**, No. 5 (Sep-Oct), 282.

[11] H. A. Ferwerda; Topics in Current Physics Vol. 9: «*Inverse Source Problems in Optics, 13, edited by H. P. Baltes*» (Berlin, Springer-Verlag, 1978).

[12] S. R. Robinson; *J. O. S. A.*, 1978, **68**, No. 1 (Jan), 87.

[13] H. A. Ferwerda, B. J. Hoenders; *Optik*, 1974, **39**, No. 4 (Feb), 317.

[14] M. Born, E. Wolf; «*Principles of Optics*» (Oxford, Pregamon Press., 1975).

[15] J. W. Goodman; «*Introduction to Fourier Optics*» (San Francisco, Mc Graw-Hill, 1968); 中译本(科学出版社, 1976); § 5–3, § 6–4.

[16] 霍裕平, 杨国桢, 顾本源, «物理学报», 1975, **24**, No. 6 (Nov), 438.

[17] 霍裕平, 杨国桢, 顾本源, «物理学报», 1976, **25**, No. 1 (Jan), 31.

# On the phase retrieval problem in optical and electronic microscopy

GU BENYUAN AND YANG GUOZHEN
(*Institute of Physics, Academia Sinica*)

(Received 19 March 1980)

## Abstract

In this paper, a set of equations for retrieving phase from two intensity distributions in the image plane or some other plane of an imaging system was obtained by the exactly mathematical derivation. The equations are completely analogous to the algorithm suggested by Gerchberg-Saxton [1,2,4] and Misell [3]. We point out that, if the transformation between two wave functions is performed by unitary operator and one of the wave functions is absolutely square integrable, the phase reconstruction from the known intensity distributions in any two planes of an imaging system can be achieved using GS algorithm.