
《高等光电系统与信号处理》讲义

(2022 版)

曹德忠

欧世峰 任 承 兰瑞君

隋晨红 申英杰

烟台大学 物理与电子信息学院

目录

第一章 信息光学的数学基础	1
§ 1.1 常用数学函数	1
§ 1.2 卷积与相关	4
§ 1.3 傅立叶变换	6
§ 1.4 线性系统的数学表示	9
第二章 标量衍射理论	14
§ 2.1 光波的数学描述	14
§ 2.2 基尔霍夫衍射理论	20
§ 2.3 衍射的角谱理论	25
§ 2.4 菲涅尔衍射	29
§ 2.5 弗朗禾费衍射	35
§ 2.6 巴比涅原理	40
§ 2.7 衍射光栅	41
§ 2.8 分数傅立叶变换与菲涅尔衍射	46
第三章 透镜的傅立叶变换性质	49
§ 3.1 透镜的位相调制	49
§ 3.2 透镜的傅立叶变换性质	53
§ 3.3 光学频谱分析	59
第四章 光学成像系统的频率响应	60
§ 4.1 相干照明衍射受限系统的点扩散函数	60
§ 4.2 相干照射下衍射受限系统的成像规律	65
§ 4.3 衍射受限系统的相干传递函数	67
§ 4.4 衍射受限非相干成像系统的传递函数	75
§ 4.5 有像差系统的传递函数	87
§ 4.6 相干与非相干成像系统的比较	90
第五章 部分相干理论	94
§ 5.1 多色光场解析信号表示	94
§ 5.2 互相干函数	98
§ 5.3 时间相干性	103
§ 5.4 空间相干性	109
§ 5.5 在准单色条件下的干涉	110
§ 5.6 互相干的传播	112
§ 5.7 范西特-泽尼克定理	115

第六章 光电器件的物理基础	119
6.1 光谱与光子能量	119
6.2 光度学(Photometry)与辐射度学(Radiometry)	120
6.3 辐通量的光谱分布与接收器的光谱响应	121
6.4 半导体基础知识	122
6.5 光电效应	131
第七章 光电探测器	149
7.1 真空光电管	149
7.2 光电倍增管	149
7.3 光电导器件	155
7.4 光伏器件	160
第八章 光电成像器件	172
8.1 象管	172
8.2 摄像管	176
8.3 CCD 固体摄像器件	183
第九章 非相干光电信号变换	189
9.1 时变光信号的直接测量	190
9.2 时变光信号的调制和解调	197
9.3 简单光学目标的空间定位	204
9.4 光学图象的扫描	211
9.5 几何参量的光电检测	215
第十章 相干光电信号变换	220
10.1 光学干涉和相干光信息分类	220
10.2 非相干探测和相干探测	223
10.3 相干光的相位调制和检测	228
10.4 相干光的频率调制和检测	237
10.5 运动参量的频率调制和零差检测	237
10.6 固定频移的频率调制和外差检测	242
10.7 直接光频调制和外差检测	245

第一章 信息光学的数学基础

信息光学的主要研究内容是光学信号在线性光学系统中的演化规律，必需借助于一些数学工具对光学信号和光学系统进行抽象描述。信息光学的数学基础介绍一些常用的数学函数，引入二维卷积、相关和傅里叶变换的数学运算，并简单介绍线性光学系统的性质。

§ 1.1 常用数学函数

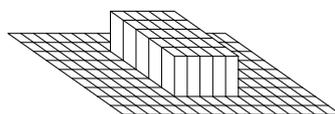
信息光学经常研究某一平面上的光场及其强度的分布，所讨论的光学信号和系统大多在二维空间进行描述，所涉及到的函数如下。

§ 1.1.1 常用二维函数

常用二维函数包括矩形函数、圆域函数、三角函数、sinc 函数和高斯函数。其中矩形函数和圆域函数常用于光阑形状的描述，而三角函数、sinc 函数和高斯函数是描述光斑强度分布的常用函数

1. 二维矩形(rectangle)函数

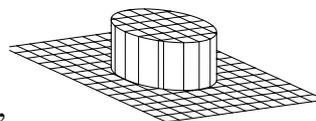
$$\text{rect}\left(\frac{x}{a}\right)\text{rect}\left(\frac{y}{b}\right) = \begin{cases} 1, & |x| \leq \frac{a}{2}, |y| \leq \frac{b}{2}, \\ 0, & \text{others,} \end{cases}$$



其中 a 、 b 分别为矩形函数在 x 、 y 方向上的分布范围。

2. 圆域(circle)函数

$$\text{circ}\left(\frac{\sqrt{x^2 + y^2}}{r_0}\right) = \begin{cases} 1 & \sqrt{x^2 + y^2} \leq r_0, \\ 0 & \text{others,} \end{cases}$$



其中 r_0 为圆域函数的分布半径。

3. 二维三角(triangle)函数

$$\text{tri}\left(\frac{x}{a}\right)\text{tri}\left(\frac{y}{b}\right) = \begin{cases} \left(1 - \left|\frac{x}{a}\right|\right)\left(1 - \left|\frac{y}{b}\right|\right), & |x| \leq a, |y| \leq b, \\ 0, & \text{others,} \end{cases}$$



其中 a 、 b 分别为三角函数的底在 x 、 y 方向上的分布范围。

4. 二维 sinc 函数

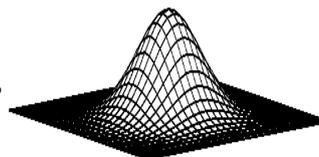
$$\text{sinc}\left(\frac{x}{a}\right)\text{sinc}\left(\frac{y}{b}\right) = \frac{\sin(\pi x/a)}{\pi x/a} \frac{\sin(\pi y/b)}{\pi y/b},$$



其中 a 、 b 分别为 sinc 函数在 x 、 y 方向上的特征分布范围。

5. 二维高斯(Gauss)函数

$$\text{Gaus}\left(\frac{x}{a}\right)\text{Gaus}\left(\frac{y}{b}\right) = \exp\left[-\pi\left(\frac{x}{a}\right)^2 - \pi\left(\frac{y}{b}\right)^2\right],$$



其中 a 、 b 分别为高斯函数在 x 、 y 方向上的特征布范围。

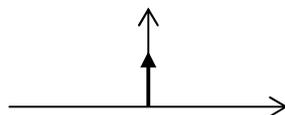
§ 1.1.2 脉冲函数

脉冲函数可用于描述点脉冲，如点光源或会聚光束中心的强度。由于任何的光学信号都可以看作是大量的连续或分离的点信号组成的，因此脉冲函数在信息光学中具有重要的地位。

1. δ (Dirac Delta) 函数

δ 函数是量子力学创始人之一狄拉克为解决物理问题而引入数学的，在物理中应用广泛，如它可以表示质点的密度，点电荷的电荷密度等物理量。数学上， δ 函数是指具有以下性质的函数

$$\delta(x) = \begin{cases} 0 & x \neq 0, \\ \infty & x = 0, \end{cases} \quad \text{并且积分 } \int_{-\infty}^{\infty} \delta(x) dx = 1.$$



δ 函数经常用一个向上的箭头表示（见上图）。 δ 函数具有如下性质

(1) 筛选性质： $\int_{-\infty}^{\infty} f(x)\delta(x-x_0)dx = f(x_0)$.

(2) δ 函数是偶函数： $\delta(-x) = \delta(x)$.

(3) 比例变换性质： $\delta(ax) = \frac{1}{|a|} \delta(x)$.

(4) 导数关系：

$$\begin{aligned} \textcircled{1} \int_{-\infty}^{\infty} f(x)\delta'(x)dx &= -\int_{-\infty}^{\infty} f'(x)\delta(x)dx \\ &= -f'(0). \end{aligned}$$

$$\begin{aligned} \textcircled{2} \int_{-\infty}^{\infty} f(x)\delta^{(n)}(x)dx &= (-1)^n \int_{-\infty}^{\infty} f^{(n)}(x)\delta(x)dx \\ &= (-1)^n f^{(n)}(0). \end{aligned}$$

$$\textcircled{3} \delta^{(n)}(-x) = (-1)^n \delta^{(n)}(x).$$

证明①： $\int_{-\infty}^{\infty} f(x)\delta'(x)dx = \int_{-\infty}^{\infty} f(x)d\delta(x)$

$$= f(x)\delta(x)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \delta(x)d[f(x)]$$

$$= 0 - \int_{-\infty}^{\infty} \delta(x)f'(x)dx = -f'(0)$$

证明②：可应用数学归纳法证明，并注意到极限条件

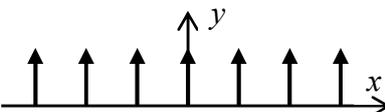
$$\delta^{(n)}(\infty) = \delta^{(n)}(-\infty) = 0.$$

证明③: 根据 $\int_{-\infty}^{\infty} f(x)\delta(x-x_0)dx=f(x_0)$, 并结合导数关系①②有

$$\begin{aligned} \delta^{(n)}(-x) &= \int_{-\infty}^{\infty} \delta^{(n)}(\alpha)\delta(\alpha+x)d\alpha \\ &= \int_{-\infty}^{\infty} \delta(\alpha+x)\delta^{(n)}(\alpha)d\alpha \\ &= (-1)^{(n)} \int_{-\infty}^{\infty} \delta^{(n)}(\alpha+x)\delta(\alpha)d\alpha \\ &= (-1)^{(n)} \delta^{(n)}(x). \end{aligned}$$

2. 梳状(comb)函数

梳状函数是指沿 x 轴分布, 间隔等于 1 的无穷多脉冲函数

$$\text{comb}(x) = \sum_{n=-\infty}^{\infty} \delta(x-n),$$


式中 n 取整数。

梳状函数可用于等间距抽样。间隔为 τ 的等间距脉冲序列可表示为如下的梳状函数

$$\sum_{n=-\infty}^{\infty} \delta(x-n\tau) = \frac{1}{\tau} \sum_{n=-\infty}^{\infty} \delta\left(\frac{x}{\tau}-n\right) = \frac{1}{\tau} \text{comb}\left(\frac{x}{\tau}\right).$$

而梳状函数和普通函数的乘积为

$$f(x) \times \frac{1}{\tau} \text{comb}\left(\frac{x}{\tau}\right) = \sum_{n=-\infty}^{\infty} f(n\tau)\delta(x-n\tau),$$

所以可以用梳状函数对其它函数进行等间距抽样。光学上常用梳状函数表示点光源的列阵, 或小孔列阵(如光栅等)的透过率函数。

§ 1.2 卷积与相关

§ 1.2.1 卷积

1. 定义

函数 $f(x,y)$ 与 $h(x,y)$ 的卷积记作

$$g(x, y) = \int \int_{-\infty}^{\infty} f(\xi, \eta) h(x - \xi, y - \eta) d\xi d\eta = f(x, y) \otimes h(x, y)$$

两函数卷积的计算步骤为：翻转、平移、相乘、积分。具体分析如下，将 $h(\xi, \eta)$ 翻转写为 $h(-\xi, -\eta)$ ，再平移 $h(x-\xi, y-\eta)$ ，最后与 $f(\xi, \eta)$ 相乘积分。可见：求两函数的卷积运算 $h(x,y) \otimes f(x,y)$ 是通过函数的相对移动（卷）并相乘（积）而得到的。

2. 卷积的运算定律

(1) 线性运算

设 a, b 为任意常数，则

$$\begin{aligned} [af_1(x, y) + bf_2(x, y)] \otimes h(x, y) &= a[f_1(x, y) \otimes h(x, y)] \\ &\quad + b[f_2(x, y) \otimes h(x, y)]. \end{aligned}$$

(2) 交换律： $f(x, y) \otimes h(x, y) = h(x, y) \otimes f(x, y)$.

(3) 有平移不变性

$$\begin{aligned} f(x - x_0, y - y_0) \otimes h(x, y) &= f(x, y) \otimes h(x - x_0, y - y_0) \\ &= g(x - x_0, y - y_0). \end{aligned}$$

(4) 结合律

$$[f(x, y) \otimes h_1(x, y)] \otimes h_2(x, y) = f(x, y) \otimes [h_1(x, y) \otimes h_2(x, y)].$$

(5) 坐标缩放性

$$f(ax, by) \otimes h(ax, by) = \frac{1}{|ab|} g(ax, by).$$

3. δ 函数的卷积

$$f(x, y) \otimes \delta(x, y) = \int \int_{-\infty}^{\infty} f(\xi, \eta) \delta(x - \xi, y - \eta) d\xi d\eta = f(x, y),$$

$$f(x, y) \otimes \delta(x - x_0, y - y_0) = f(x - x_0, y - y_0).$$

4. 卷积的效应

卷积对信号具有展宽和平滑化两种作用。(1)展宽：一般说来，卷积后函数的宽度等于被两个卷积函数的宽度之和。(2)平滑化：卷积后的函数本身的起伏振荡变得平缓圆滑。

§ 1.2.2 相关

1. 互相关的定义

设 $f(x, y)$ 和 $g(x, y)$ 为复函数，则两函数的互相关函数定义为

$$r_{fg} = f(x, y) \star g(x, y) \\ = \int \int_{-\infty}^{\infty} f(\xi, \eta) g^*(\xi - x, \eta - y) d\xi d\eta.$$

如果令 $\xi - x = \xi', \eta - y = \eta'$ ，则代入上式可得互相关的另一种定义形式

$$r_{fg} = f(x, y) \star g(x, y) \\ = \int \int_{-\infty}^{\infty} f(\xi' + x, \eta' + y) g^*(\xi', \eta') d\xi' d\eta'.$$

互相关运算与卷积运算不同，其一，不需翻转，其二，第二个函数取复共轭。根据互相关的定义可知，一般情况下互相关不满足交换律

$$f(x, y) \star g(x, y) \neq g(x, y) \star f(x, y)。$$

2. 互相关与卷积的关系

根据卷积和相关的定义可得： $f(x, y) \star g(x, y) = f(x, y) \otimes g^*(-x, -y)$ 。

只有当 $g(x, y)$ 为实的偶函数时，互相关与卷积运算才等价，

$$f(x, y) \star g(x, y) = f(x, y) \otimes g(x, y)。$$

计算互相关时要注意两个函数的顺序。互相关运算是两个信号之间存在多少相似性的量度。

3. 自相关

$$r_{ff} = f(x, y) \star f(x, y) \\ = \int \int_{-\infty}^{\infty} f(\xi, \eta) f^*(\xi - x, \eta - y) d\xi d\eta$$

或

$$r_{ff} = \int \int_{-\infty}^{\infty} f(\xi + x, \eta + y) f^*(\xi, \eta) d\xi d\eta$$

4. 自相关运算的性质

(1) 厄米性 $r_{ff}(x, y) = r_{ff}^*(-x, -y)$ 。

(2) 原点相关极大 $|r_{ff}(x, y)| \leq r_{ff}(0, 0)$ 。

-----*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-----

§ 1.3 傅立叶变换

§ 1.3.1 傅里叶变换

1. 傅里叶变换

一维连续变量的傅里叶变换及其逆变换为

$$G(f) = \int_{-\infty}^{\infty} g(x) e^{-i2\pi fx} dx,$$

$$g(x) = \int_{-\infty}^{\infty} G(f) e^{i2\pi fx} df.$$

在光学中，常用二维的傅里叶变换对为

$$\begin{aligned} G(f_x, f_y) &= \mathbf{F}\{g(x, y)\} & g(x, y) &= \mathbf{F}^{-1}\{G(f_x, f_y)\} \\ &= \iint_{-\infty}^{\infty} g(x, y) e^{-i2\pi(f_x x + f_y y)} dx dy, & &= \iint_{-\infty}^{\infty} G(f_x, f_y) e^{i2\pi(f_x x + f_y y)} df_x df_y. \end{aligned}$$

其中黑体字母 \mathbf{F} 代表傅里叶变换。

傅里叶变换对函数有严格的要求，本文不作过多讨论。在应用中，虽然物理过程大多是非周期的，可以将非周期函数延拓成周期函数，或者认为周期为无穷大，就可以进行傅里叶展开，称之为广义傅里叶变换。

2. 傅里叶变换的性质定理

(1) 线性定理

$$\mathbf{F}\{ag(x) + bh(x)\} = aG(f) + bH(f)$$

(2) 伸缩定理

$$\mathbf{F}\{g(ax)\} = \frac{1}{|a|} G\left(\frac{f}{a}\right), \mathbf{F}\left\{g\left(\frac{x}{a}\right)\right\} = |a|G(af)$$

(3) 平移定理

$$\mathbf{F}\{g(x \pm x_0)\} = \exp[\pm i2\pi f x_0]G(f)$$

(4) 对称定理 $g(x) \begin{cases} \text{奇} \\ \text{偶} \end{cases} \rightarrow G(f) \begin{cases} \text{奇} \\ \text{偶} \end{cases}$

(5) 积分定理

$$G(0) = \int_{-\infty}^{\infty} g(x) dx, \quad g(0) = \int_{-\infty}^{\infty} G(f) df.$$

(6) 卷积定理：如果 $\mathbf{F}\{g(x)\} = G(f)$ ， $\mathbf{F}\{h(x)\} = H(f)$ ，则

$$\mathbf{F}\{g(x) \otimes h(x)\} = G(f)H(f),$$

$$\mathbf{F}\{g(x) \cdot h(x)\} = G(f) \otimes H(f)$$

3. 可分离变量函数的傅里叶变换

若函数在直角坐标系中是可分离的， $g(x, y) = g_x(x)g_y(y)$ ，则其傅里叶变换也是可分离的

$$G(f_x, f_y) = G_x(f_x)G_y(f_y)。$$

在极坐标中，若函数是可分离变量的， $g(x, y) = g_R(r)g_\theta(\theta)$ ，则可通过汉克尔（Hankel）变换的一维运算来进行。下面研究轴对称函数的傅里叶--贝塞尔变换。

轴对称函数 $g_\theta(\theta) = 1$ ，则 $g(r, \theta) = g_R(r)$ 。 $g(x, y)$ 在直角坐标中的傅里叶变换式为：

$$G(f_x, f_y) = \int \int_{-\infty}^{\infty} dx dy \times g(x, y) e^{-i2\pi(f_x x + f_y y)}$$

利用变换式

$$r = \sqrt{x^2 + y^2}, \quad \theta = \arctg\left(\frac{y}{x}\right),$$

$$x = r \cos \theta, \quad y = r \sin \theta,$$

$$\rho = \sqrt{f_x^2 + f_y^2}, \quad \varphi = \arctg\left(\frac{f_y}{f_x}\right),$$

$$f_x = \rho \cos \varphi, \quad f_y = \rho \sin \varphi,$$

以及 $dx dy = r d\theta dr$ ，代入傅里叶变换 G 式得

$$\begin{aligned} G(\rho \cos \varphi, \rho \sin \varphi) &= \int_0^{\infty} dr \int_0^{2\pi} d\theta \times r g(r, \theta) e^{-i2\pi r \rho (\cos \theta \cos \varphi + \sin \theta \sin \varphi)} \\ &= \int_0^{2\pi} d\theta \int_0^{\infty} dr \cdot r g_R(r) e^{-i2\pi r \rho \cos(\theta - \varphi)} \end{aligned}$$

利用贝塞尔恒等式

$$J_0(a) = \frac{1}{2\pi} \int_0^{2\pi} \exp[-ia \cos(\theta - \varphi)] d\theta$$

G 式改写为：

$$G(\rho \cos \varphi, \rho \sin \varphi) = 2\pi \int_0^{\infty} g_R(r) J_0(2\pi r \rho) r dr$$

式中 J_0 为零阶贝塞尔函数，上式因只与 ρ 径向有关，故可以改写为

$$G(\rho) = 2\pi \int_0^{\infty} g_R(r) J_0(2\pi r \rho) r dr$$

同时，逆付氏变换为：

$$g(r) = 2\pi \int_0^{\infty} G(\rho) J_0(2\pi r \rho) \rho d\rho$$

这种圆对称函数的傅里叶变换在信息光学中会频繁出现，被称之为傅里叶--贝塞尔变换，即

$$\begin{aligned} \mathbf{B}\{g(r)\} &= G(\rho) \\ &= 2\pi \int_0^\infty g_R(r) J_0(2\pi r\rho) r dr \end{aligned}$$

$$\begin{aligned} \mathbf{B}^{-1}\{G(\rho)\} &= g(r) \\ &= 2\pi \int_0^\infty G(\rho) J_0(2\pi r\rho) \rho d\rho \end{aligned}$$

直角坐标中傅里叶变换的所有性质，在极角坐标的傅里叶--贝塞尔变换中有对应的结果。

§ 1.3.2 常用的傅里叶变换对

1. δ 函数的变换: $\mathbf{F}\{A\delta(x)\} = A\mathbf{F}\{\delta(x)\} = A$

2. 常数的变换: $\mathbf{F}\{1\} = \int_{-\infty}^\infty 1 \cdot \exp[-i2\pi fx] dx = \delta(f)$,

$$\mathbf{F}\{A\} = A\mathbf{F}\{1\} = A\delta(f)$$

3. $\delta(x + x_0)$ 的变换: $\mathbf{F}\{\delta(x + x_0)\} = \exp[i2\pi fx_0]$

4. $\exp[i2\pi f(x_0)]$ 的变换: $\mathbf{F}\{\exp[i2\pi f x_0]\} = \delta(f - f_0)$

5. $\cos(2\pi f_0 x)$ 的变换

$$\mathbf{F}\{\cos(2\pi f_0 x)\} = \frac{1}{2}[\delta(f - f_0) + \delta(f + f_0)]$$

6. $\text{rect}(x)$ 的变换: $\mathbf{F}\{\text{rect}(x)\} = \text{sinc}(f)$

7. $\text{sinc}(x)$ 的变换: $\mathbf{F}\{\text{sinc}(x)\} = \text{rect}(-f) = \text{rect}(f)$

8. 三角函数 $\text{tri}(x)$ 的变换: $\mathbf{F}\{\text{tri}(x)\} = \text{sinc}^2(f)$

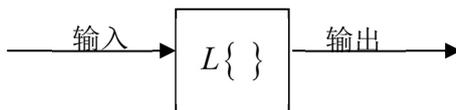
-----* * * * *-----

§ 1.4 线性系统的数学表示

一个稳定的物理系统，输入一个函数，输出一个确定的函数。系统的作用可用一个算符来表示 $L\{\dots\}$

$$g(x, y) = L\{f(x, y)\},$$

其中 $f(x, y)$ 和 $g(x, y)$ 分别为系统的输入和输出函数。从数学角度，我们可以不考虑系统内部的具体结构，将系统看作具有特定变换作用的黑箱子，如下图所示，只要变换的结果相同，则系统对外界输入的作用是相同的。



一个实际的系统，其输入输出函数的关系是极为复杂的，具体地指出变换算符 $L\{\dots\}$ 的具体形式和性质是十分困难的。我们主要研究相对简单的线性不变系统。

§ 1.4.1 线性系统

1. 定义

若对所有的输入函数 $f_1(x, y)$ 和 $f_2(x, y)$ 和复常数 a_1, a_2 ，输出满足下列关系式：

$$L\{a_1 f_1(x, y) + a_2 f_2(x, y)\} = a_1 L\{f_1(x, y)\} + a_2 L\{f_2(x, y)\}$$

则称系统为线性系统。线性系统的作用是把组合的变换转化为变换的组合。

线性系统表明，如果能把总输入函数分解成各个“基元函数” f_1, f_2 与相应的复常数 a_1, a_2 的乘积的总和，那么线性系统的总输出函数就等于各个基元函数的输出函数与相应的复常数的乘积的总和。系统对基元函数的响应显然反映了系统的性质。在光学中，基元函数的选择有两种，一为指数函数 $\exp[i\varphi]$ ，即平面波，二为 δ 函数，即点基元。

线性系统具有如下的特点：(1) 叠加性：系统中一个输入并不影响其它输入的响应；(2) 均匀性：系统能够保持对输入信号的缩放因子不变。

2. 线性系统的脉冲响应

设系统的输入和输出平面的坐标为 (x_1, y_1) 和 (x_2, y_2) ， $f(x_1, y_1)$ 为一线性系统的输入函数，可以将其看作为 (x_1, y_1) 平面上不同位置处的许多 δ 函数的线性组合，

$$f(x_1, y_1) = \int \int_{-\infty}^{\infty} f(\xi, \eta) \delta(x_1 - \xi, y_1 - \eta) d\xi d\eta.$$

通过线性系统后，其输出函数为：

$$\begin{aligned} g(x_2, y_2) &= L\{f(x_1, y_1)\} \\ &= L\left\{\int \int_{-\infty}^{\infty} f(\xi, \eta) \delta(x_1 - \xi, y_1 - \eta) d\xi d\eta\right\} \\ &= \int \int_{-\infty}^{\infty} f(\xi, \eta) L\{\delta(x_1 - \xi, y_1 - \eta)\} d\xi d\eta \\ &= \int \int_{-\infty}^{\infty} f(\xi, \eta) h(x_2, y_2; \xi, \eta) d\xi d\eta, \end{aligned}$$

式中

$$h(x_2, y_2; \xi, \eta) = L\{\delta(x_1 - \xi, y_1 - \eta)\}$$

称为系统的脉冲响应。表示在系统的输出平面 (x_2, y_2) 点处，由系统的输入平面上坐标为 (ξ, η) 点的 δ 函数所激励的响应。上式表明，线性系统的性质完全由脉冲响应函数来决定，对于 $h(x_2, y_2; \xi, \eta)$ 已知的系统，任何输入函数所对应的输出函数都可以用上述积分求出。

对于线性光学成像系统而言，上式的物理意义在于：对于一个线性成像系统，只要知道了物场中各点的像，则任何物的像便可求出。

§ 1.4.2 线性不变系统：时间不变系统、空间不变系统

时间不变系统：不同时间输入同一信号，其输出信号（函数）形式不变。即对于相同的输入信号，其输出信号不随输入时间的改变而改变。

空间不变系统：不因系统所在空间的改变而改变输入输出关系。

1. 线性不变系统的定义

输入 $f(x_1, y_1)$ ，通过系统后，其输出为 $g(x_2, y_2) = L\{f(x_1, y_1)\}$ 。如果 $f(x_1, y_1)$ 有一位移 (ξ, η) ，其输出的函数形式不变，

$$g(x_2 - \xi, y_2 - \eta) = L\{f(x_1 - \xi, y_1 - \eta)\},$$

则该系统称为不变系统。

同样地，对于位于原点的 δ 函数，其响应为 $h(x_2, y_2) = L\{\delta(x_1, y_1)\}$ ，偏离原点时，即 $\delta(x_1 - \xi, y_1 - \eta)$ ，其响应为

$$h(x_2 - \xi, y_2 - \eta) = L\{\delta(x_1 - \xi, y_1 - \eta)\}。$$

比较以上两个公式得

$$h(x_2, y_2, \xi, \eta) = h(x_2 - \xi, y_2 - \eta)。$$

系统既是线性系统又是空间不变系统，称为线性空不变系统。对于空间不变系统，它的输入和输出的变换关系是不随空间位置的变化而变化的。实际上，对于点光源的 δ 函数，其响应 h 是有一定分布的，叫做弥散，与日晕和月晕相似，但只要不随空间变化这种特性就称为等晕性。

2. 线性不变系统的叠加积分——卷积

由线性系统积分可知

$$\begin{aligned} g(x_2, y_2) &= \int \int_{-\infty}^{\infty} f(\xi, \eta) h(x_2, y_2; \xi, \eta) d\xi d\eta \\ &= \int \int_{-\infty}^{\infty} f(\xi, \eta) h(x_2 - \xi, y_2 - \eta) d\xi d\eta \\ &= f(x_2, y_2) \otimes h(x_2, y_2) \end{aligned}$$

说明线性系统的输出函数 $g(x, y)$ 等于输入函数 $f(x, y)$ 与系统脉冲响应的卷积。这种卷积运算体现的物理

意思是：将输入函数 $f(x, y)$ 分解为许多不同位置的 δ 函数的线性组合，每个脉冲按其位置不同分别加权然后叠加起来，就得出系统对输入函数的整体响应。需要注意的是，与线性叠加的意义相似，不同的是它不随位置变化而变化——线性空不变。

§ 1.4.3 线性不变系统的传递函数

对卷积 $g(x, y) = f(x, y) \otimes h(x, y)$ 取傅里叶变换

$$F \{g(x, y)\} = F \{f(x, y) \otimes h(x, y)\},$$

变换后可得

$$G(f_x, f_y) = F(f_x, f_y)H(f_x, f_y),$$

式中

$$\begin{aligned} H(f_x, f_y) &= F \{h(x, y)\} \\ &= \int \int_{-\infty}^{\infty} h(x, y) e^{-i2\pi(f_x x + f_y y)} dx dy \end{aligned}$$

称为系统的传递函数，亦称为频率响应函数。系统的传输关系可表述为：输出频谱 $G(f_x, f_y)$ 等于输入频谱 $F(f_x, f_y)$ 与传递函数 $H(f_x, f_y)$ 的乘积。

线性不变系统的频率响应关系把繁复的卷积运算变成简单的乘积运算。输出函数可通过求 $G(f_x, f_y)$ 的逆变换得出

$$g(x, y) = F^{-1} \{G(f_x, f_y)\}。$$

§ 1.4.4 线性不变系统的本征函数

1. 本征函数

设 $f(x, y, f_a, f_b)$ 为输入函数，若输出与输入的函数形式不变

$$L \{f(x, y, f_a, f_b)\} = H(f_a, f_b) f(x, y, f_a, f_b)$$

则 $f(x, y, f_a, f_b)$ 叫做系统的本征函数， $H(f_a, f_b)$ 叫做本征函数的本征值。一个线性不变系统的本征函数，通过系统后不改变函数形式，只是被衰减或放大和产生相移。线性空间不变系统的本征函数一般有以下两种：复指数函数和余弦函数。

2. 复指数函数是线性不变系统的本征函数

证明：设输入函数为

$$f(x, y) = e^{i2\pi(f_a x + f_b y)},$$

其频谱为

$$F(f_x, f_y) = F \{e^{i2\pi(f_a x + f_b y)}\} = \delta(f_x - f_a, f_y - f_b)$$

4. 证明 $F\{\exp[i2\pi fx]\} = \delta(f - f_0)$

5. 证明 $F\{\text{rect}(x)\} = \text{sinc}(f)$

第二章 标量衍射理论

光是特定频段的电磁波，应该采用矢量理论进行描述。在某些情形下，采用近似的标量衍射理论描述光场，数学处理简单，物理图像直观，而且所给出的结果与实际观察非常符合。例如当衍射孔径远远大于波长，且观察平面到衍射孔径的距离也远远大于波长时。本章首先介绍光波的数学描述，然后讨论基尔霍夫衍射理论和角谱理论两种标量衍射理论，并利用此理论处理常见的衍射问题，最后介绍分数傅里叶变换和菲涅尔衍射之间的联系。基尔霍夫衍射理论和角谱理论分别从空间域和频率域讨论衍射现象，它们在本质上是统一的。

§ 2.1 光波的数学描述

一、单色光波场的复振幅表示

定态光场（稳定的光场分布）要求光波的波列无限长，这实际上是无法实现的。当光波波列的持续时间比光扰动的周期长得多时，可以将这种光波波列当作无限长单色波列处理。单色光波场中 P 点在 t 时刻的振动可以用标量函数表示为

$$u(P, t) = a(P) \cos[\omega t - \varphi(P)],$$

根据欧拉公式，把上式写成指数形式，取其实部，写为

$$\begin{aligned} u(P, t) &= \operatorname{Re}\{a(P) \exp[-i\omega t + i\varphi(P)]\} \\ &= \operatorname{Re}\{a(P) \exp[i\varphi(P)] \exp[-i\omega t]\}. \end{aligned}$$

将上式中和空间有关的光场部分定义为单色光波场中 P 点的复数振幅

$$U(P) = a(P) \exp[i\varphi(x, y, z)],$$

则光振动可以表示为

$$u(P, t) = \operatorname{Re}\{U(P) \exp[-i\omega t]\}.$$

在单色波光场的计算中，可以直接利用复振幅进行。指数函数形式的复振幅在线性运算（加减乘除）中更为方便，只要将计算结果取实部即可得到实数表达式，尤其对于强度计算

$$I = |U|^2 = U^* U,$$

复振幅的计算更为方便。

二、球面波的表达式

1. 单色球面波的复振幅

从点光源发出的光，其波面为球面。一个复杂的光源可以看作是许多点光源的集合，因而球面波是基本的波面形式。单色的发散球面波 \vec{k} 与 \vec{r} 一致，在光场中任一 P 点所产生的复振幅为

$$U(P) = \frac{a_0}{r} \exp[i\vec{k} \cdot \vec{r}] = \frac{a_0}{r} \exp[ikr],$$

式中 \vec{k} 为波矢量，其大小 $k = 2\pi / \lambda$ ，方向沿光的传播方向， \vec{r} 表示观察点 $P(x, y, z)$ 的位置向量， a_0 表示距离光源单位距离处的振幅。

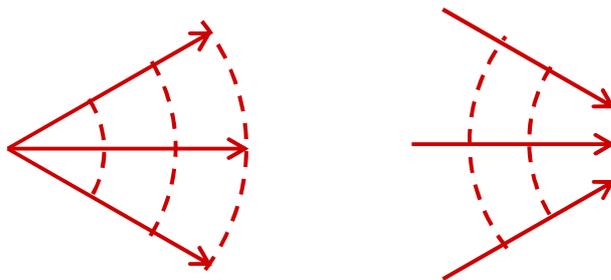


图 2.1 左图为会聚球面波，右图为发散球面波。

会聚球面波 \vec{k} 与 \vec{r} 反向，其复振幅为

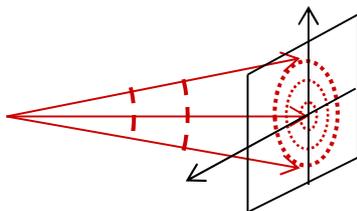
$$U(P) = \frac{a_0}{r} \exp[-ikr]$$

上述各式中，光源到考察点的距离为

$$r = \sqrt{(x - x_0)^2 + (y + y_0)^2 + (z - z_0)^2}$$

其中 (x_0, y_0, z_0) 为点光源的坐标。

2. 球面波光场中任一平面上的复振幅分布



设球面波中心与坐标原点重合，则 x, y 平面上的复振幅为

$$U(x, y, z) = \frac{a}{r} \exp[ikr]$$

其中 $r = (x^2 + y^2 + z^2)^{\frac{1}{2}} = z \left(1 + \frac{x^2 + y^2}{z^2} \right)^{\frac{1}{2}}$

在近轴情况下，即 $z^2 \gg x^2 + y^2$ ，可按牛顿二项式展开得

$$r = z \left[1 + \frac{1}{2} \frac{x^2 + y^2}{z^2} - \frac{1}{8} \left(\frac{x^2 + y^2}{z^2} \right)^2 + \dots \right]$$

在近轴近似中，对于 $U(x, y, z)$ 的振幅部分，取 $r \approx z$ ；而对于 $U(x, y, z)$ 的位相部分，取

$$r \approx z \left(1 + \frac{x^2 + y^2}{2z} \right)$$

代入 $U(x, y, z)$ 式得 z 处 (x, y) 平面上光场的分布

$$\begin{aligned}
 U(x, y, z) &\approx \frac{a}{z} \exp(ikz) \cdot \exp\left(ik \frac{x^2 + y^2}{2z}\right) \\
 &= U_0 \exp\left(ik \frac{x^2 + y^2}{2z}\right).
 \end{aligned}$$

若球面波中心不在坐标原点，上式改为

$$U(x, y, z) = U_0 \exp\left[ik \frac{(x - x_0)^2 + (y + y_0)^2}{2z}\right],$$

其中 $U_0 = \frac{a}{z} \exp(ikz)$ 。

在 (x, y) 平面上位相相同点的轨迹满足等相位线方程

$$(x - x_0)^2 + (y + y_0)^2 = C,$$

式中 C 为一常数。不同的 C 值对应的等相位线构成同心圆族。注意相位差 2π 的同心圆之间的间隔并不相等，而是由中心向外越来越密集。

三、平面波

1. 单色平面波

平面波也是光波的一种最简单的形式，对平面波的分析是角谱理论和光学傅里叶变换的基础。沿 \vec{k} 方向传播的单色平面波，在光场中 $P(x, y, z)$ 点产生的复振幅可以表示为

$$\begin{aligned}
 U(P) &= a \exp[i\vec{k} \cdot \vec{r}] \\
 &= a \exp[ik(x \cos \alpha + y \cos \beta + z \cos \gamma)]
 \end{aligned}$$

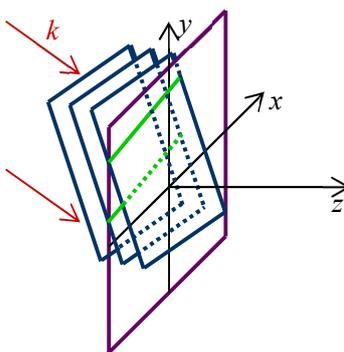
式中 a 表示常量振幅， $\cos \alpha$ 、 $\cos \beta$ 、 $\cos \gamma$ 为传播方向的方向余弦，它们之间满足

$$\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1。$$

单色平面波的等相位面满足

$$x \cos \alpha + y \cos \beta + z \cos \gamma = c。$$

可见，等相面是一些平行平面。



2. 任一平面上的平面波

沿某一方向传播的平面波，在 $z = z_1$ 处的 (x, y) 平面上的复振幅表示为

$$U(x, y, z_1) = a \exp(ikz_1 \cos \gamma) \exp[ik(x \cos \alpha + y \cos \beta)]$$

$$= U_0 \exp[ik(x \cos \alpha + y \cos \beta)]$$

当 z_1 一定时， U_0 为一常数，等相位线满足

$$x \cos \alpha + y \cos \beta = c,$$

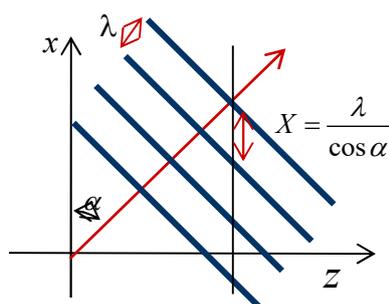
可知等位线是一些平行线 $y = -\frac{\cos \alpha}{\cos \beta}x + c$ 。相位差为 2π 的等相位线是一组等间距的平行线。

四、平面波的空间频率

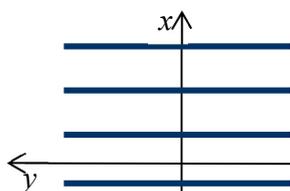
1. 光波矢量位于 (x, z) 平面内

当平面光波矢量位于 (x, z) 平面内， $\beta = \pi/2$ ， (x, y) 平面上的复振幅分布

$$U(x, y) = A \exp[ikx \cos \alpha +iky \cos \beta]。$$



等相位线方程为 $x \cos \alpha = C$ ，在 (x, y) 平面内，这些不同 C 值的等相位线是垂直于 x 轴的平行线，如图所示。



相位差依次相差 2π 的等相线，其间隔为 X ，则

$$kX \cos \alpha = 2\pi$$

可得在 x 方向等相线的间距

$$X = \frac{2\pi}{k \cos \alpha} = \frac{\lambda}{\cos \alpha}。$$

定义空间频率为单位长度内变化的周期数，则

$$f_x = \frac{1}{X} = \frac{\cos \alpha}{\lambda}$$

称为 x 方向的空间频率，其单位为 $1/\text{m}$ 。同理可知 y 方向等相线间距 $Y=\infty$ ， y 方向空间频率 $f_y=0$ 。

根据以上讨论，传播方向余弦为 $(\cos \alpha, 0)$ 的单色平面波在 (x, y) 平面上的复振幅的周期分布可以用空间频率 $(f_x = \frac{\cos \alpha}{\lambda}, f_y = 0)$ 来描述。因此可以改写为

$$U(x, y) = A \exp[i2\pi f_x x],$$

公式直接通过空间频率表示 (x, y) 平面上的复振幅分布，空间频率与传播方向余弦存在一一对应。

2. 任意情形的空间频率

光波是时间和空间的函数，具有时间周期性与空间周期性。单色光波的一般表示式

$$U(x, y, z) = U_0(x, y, z) \cdot \exp[i\vec{k} \cdot \vec{r}],$$

显然对于单色光波的时间量和空间量有如下的对应关系

$$\text{时间角频率 } \omega = 2\pi\nu = \frac{2\pi}{T}$$

$$\text{空间角频率 } K = 2\pi f = \frac{2\pi}{\lambda}$$

其中： ν ——时间频率， T ——时间周期

f ——空间频率， λ ——空间周期。

根据波矢量 $\vec{k} = k\{\cos \alpha, \cos \beta, \cos \gamma\}$ 和位置矢量 $\vec{r} = \{x, y, z\}$ 可得

$$\vec{k} \cdot \vec{r} = k(x \cos \alpha + y \cos \beta + z \cos \gamma),$$

代入复振幅表达式：

$$\begin{aligned} U(x, y, z) &= U_0(x, y, z) e^{ik(x \cos \alpha + y \cos \beta + z \cos \gamma)} \\ &= U_0(x, y, z) \times e^{i2\pi(\frac{\cos \alpha}{\lambda}x + \frac{\cos \beta}{\lambda}y + \frac{\cos \gamma}{\lambda}z)} \\ &= U_0(x, y, z) e^{i2\pi(f_x x + f_y y + f_z z)} \end{aligned}$$

式中 $f_x = \frac{\cos \alpha}{\lambda}$ ， $f_y = \frac{\cos \beta}{\lambda}$ ， $f_z = \frac{\cos \gamma}{\lambda}$ 。当 $\alpha, \beta, \gamma < 90^\circ$ 时， $f_x, f_y, f_z > 0$ ，表示 \vec{k} 沿正方向传播；

当 $\alpha, \beta, \gamma > 90^\circ$ 时 $f_x, f_y, f_z < 0$ ，表示 \vec{k} 沿负方向传播。

3. 空间频率表示的平面波

由空间频率的定义 $f_x = \frac{\cos \alpha}{\lambda}$ ， $f_y = \frac{\cos \beta}{\lambda}$ ， $f_z = \frac{\cos \gamma}{\lambda}$ ，平面波可表示为

$$U(x, y, z) = U_0 e^{i2\pi(f_x x + f_y y + f_z z)}。$$

而二维的波函数为：

§ 2.2 基尔霍夫衍射理论

1678年，惠更斯为解释波的传播提出子波的假设，认为波面上每一点都可以作为次级子波的波源，后一时刻的波阵面则可看作是这些子波的包络面。1818年，菲涅耳引入干涉概念对惠更斯原理进行了补充，认为子波源应当是相干的，后空间光场是子波干涉的结果。

惠更斯作图法加上干涉原理，就称为惠更斯—菲涅耳原理。设 P_0 点是点源 P' 发出的一个球面单色波阵面的瞬时位置， P 点是光扰动待定的一点。按惠更斯—菲涅耳原理，得 P_0 点面元对 P 点扰动的贡献为：

$$dU(P) = C \cdot K(\theta) \frac{A e^{ikr'}}{r'} \cdot \frac{e^{ikr}}{r} ds$$

P 点的总扰动为：

$$U(P) = C \iint_{\Sigma} \frac{A}{r'} e^{ikr'} \cdot \frac{K(\theta)}{r} \cdot e^{ikr} ds$$

式中： C 为常数， $K(\theta)$ 为倾斜因子

一、亥姆霍兹方程

单色光波场中任意一点 P 的光振动 u 满足标量波动方程

$$\nabla^2 u - \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = 0,$$

其中 ∇^2 是拉普拉斯算符，在直角坐标下

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}。$$

实数的光振动可以表示成时空分离的形式

$$u(P, t) = \text{Re}\{U(P) \exp[-i\omega t]\},$$

将其代入到波动方程中得

$$(\nabla^2 + k^2)U(P) = 0,$$

其中 $k = \frac{2\pi\nu}{c} = \frac{2\pi}{\lambda}$ 为波数。这就是亥姆霍兹方程，对于平面波和球面波均满足此方程。

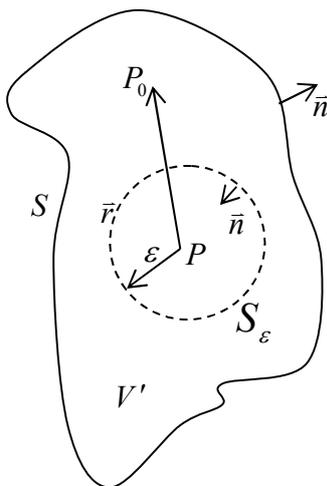
二、亥姆霍兹-基尔霍夫积分定理

为了寻求光场中的任意一点 P 的复振幅用光场其它各点的复振幅表示出来，需要用到格林定理，它可表述如下：

当 $U(P)$ 和 $G(P)$ 是空间位置坐标的任意两个复数函数, S 为包围空间某体积 V 的封闭曲面。若在 S 面内和 S 面上, $U(P)$ 和 $G(P)$ 均单值连续, 且具有单值连续的一阶和二阶导数, 则

$$\iiint_V (G\nabla^2 U - U\nabla^2 G) dV = \oiint_S (U \frac{\partial G}{\partial n} - G \frac{\partial U}{\partial n}) dS$$

式中 $\frac{\partial}{\partial n}$ 表示 S 上每一点沿向外法线方向的偏导数。1882 年基尔霍夫利巧妙选择格林函数 G 和封闭曲面, 用格林定理求解波动方程。



选择包围观察点 P 的任意封闭曲面 S , 如图所示。选择由 P 点向外发散的球面波为格林函数。在任意 P_0 上 G 的值为

$$G(P_0) = \frac{\exp[ikr]}{r}$$

式中 r 表示 P 指向 P_0 点的矢量 \vec{r} 的长度。

函数 G 及其一阶、二阶偏导数在被包围的体积 V 内必须是连续的。为了排除 P 的不连续性, 用半径为 ϵ 的小球面 S_ϵ 镶嵌在 P 点周围。积分的体积 V' 介于 S 和 S_ϵ 之间的那部分空间, 积分曲面是复合曲面 $S' = S + S_\epsilon$ 。注意图中曲面外法线方向的标定。

在体积 V' 内, 球面波 G 满足亥姆霍兹方程

$$(\nabla^2 + k^2)G = 0。$$

将其代入格林定理表达式的左端

$$\iiint_{V'} (G\nabla^2 U - U\nabla^2 G) dV = -\iiint_{V'} k^2 (GU - UG) dV = 0，$$

因此格林定理简化为

$$\oiint_{S+S_\epsilon} (U \frac{\partial G}{\partial n} - G \frac{\partial U}{\partial n}) dS = 0$$

或者

$$\oiint_S (U \frac{\partial G}{\partial n} - G \frac{\partial U}{\partial n}) dS = -\oiint_{S_\epsilon} (U \frac{\partial G}{\partial n} - G \frac{\partial U}{\partial n}) dS$$

对于 S 上的任意一点 P0, 满足公式 (SSS),

$$\frac{\partial G(P_0)}{\partial n} = \frac{\partial G(P_0)}{\partial r} \cdot \frac{\partial r}{\partial n} = \left(ik - \frac{1}{r}\right) \cdot \frac{\exp[ikr]}{r} \cdot \cos(\vec{n}, \vec{r}),$$

式中 $\cos(\vec{n}, \vec{r})$ 代表外向法线 \vec{n} 与矢量 \vec{r} 之间夹角的余弦。对于 S_ε 上的点 P0, $\cos(\vec{n}, \vec{r}) = -1$, 此时

$$G(P_0) = \frac{\exp[ik\varepsilon]}{\varepsilon}$$

$$\frac{\partial G(P_0)}{\partial n} = \left(\frac{1}{\varepsilon} - ik\right) \frac{\exp[ik\varepsilon]}{\varepsilon}。$$

令 $\varepsilon \rightarrow 0$, 根据 U 及其导数在 P 点的连续性得

$$\lim_{x \rightarrow 0} \iint_{S_\varepsilon} \left(G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n}\right) dS = \lim_{x \rightarrow 0} 4\pi\varepsilon^2 \left(G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n}\right)$$

$$= -4\pi U(P)$$

把此结果代入格林定理得到亥姆霍兹-基尔霍夫积分定理

$$U(P) = \frac{1}{4\pi} \iint_S \left[\frac{\partial U}{\partial n} \left(\frac{e^{ikr}}{r}\right) - U \frac{\partial}{\partial n} \left(\frac{e^{ikr}}{r}\right) \right] dS。$$

亥姆霍兹-基尔霍夫积分定理的意义在于衍射场中任意点 P 的复振幅分布 U(P) 可以用包围在该点的任意封闭曲面 S 上各点扰动的边界值计算得出。

三、基尔霍夫衍射公式

现在讨论无限大不透明屏上一个孔径的衍射。假定光从左向右照射衍射屏, 选择的封闭曲面由两部分组成, 紧靠屏幕后面的平面 S_1 和中心在观察点 P 上, 半径为 R 的大球面罩 S_2 组成。由积分定理

$$U(P) = \frac{1}{4\pi} \iint_S \left[G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n} \right] dS,$$

选择 G 为球面波 $G = \frac{\exp[ikr]}{r}$ 。

当 R 增大时, S_2 趋近于大的半球壳, 则 S_2 面上 G 的分布为

$$G = \frac{\exp[ikR]}{R},$$

$$\frac{\partial G}{\partial n} = \left(ik - \frac{1}{R}\right) \frac{\exp[ikR]}{R} = ikG$$

S_2 面上的积分为

$$\iint_{S_2} \left[G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n} \right] dS = \iint_{\Omega} G \left(\frac{\partial U}{\partial n} - ikU \right) R^2 d\omega$$

式中的 Ω 是 S_2 对 P 点所张的立体角。只要光场分布满足索末菲辐射条件

$$\lim_{R \rightarrow \infty} R \left(\frac{\partial U}{\partial n} - ikU \right) = 0,$$

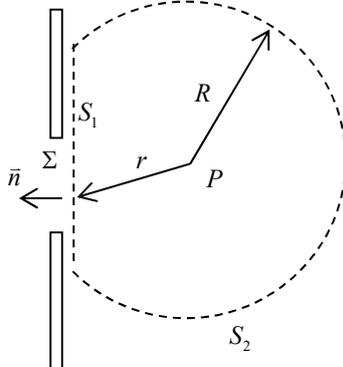
则 S_2 上的积分就称为零, 这一条件只要在 $R \rightarrow \infty$ 总会得到满足。

如此一来, P 点的扰动可以只用紧靠在屏后的平面 S_1 上的计算可以得到, 即

$$U(P) = \frac{1}{4\pi} \iint_{S_1} \left[G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n} \right] dS。$$

屏幕上除了透光的孔径之外，其余部分是不透明的。基尔霍夫采取如下的假定：

- (1) 在孔径 Σ 上，场分布 U 及其偏导数与没有屏幕时完全相同；
- (2) 在屏幕的几何阴影区，场分布 U 及其偏导数的值恒为零。



这样， P 点的扰动可以由孔径 Σ 之内的光场分布及其法向偏导求出

$$U(P) = \frac{1}{4\pi} \iint_{\Sigma_1} \left[G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n} \right] dS。$$

又由于从孔径到观察点的距离 r 通常远远大于波长，因而

$$\frac{\partial G(P_0)}{\partial n} = \left(ik - \frac{1}{r} \right) \frac{e^{ikr}}{r} \frac{\partial r}{\partial n} \approx ik \frac{e^{ikr}}{r} \cos(\bar{n}, \bar{r}),$$

代入上式可得

$$U(P) = \frac{1}{4\pi} \iint_{S_1} \frac{e^{ikr}}{r} \left[\frac{\partial U}{\partial n} - ikU \cos(\bar{n}, \bar{r}) \right] dS。$$

假设点光源发出球面波照明屏幕的孔径，点光源到孔径的距离为 r' ，则

$$U(P_0) = A \frac{e^{ikr'}}{r'}, \quad \frac{\partial U}{\partial n} = ik \frac{Ae^{ikr'}}{r'} \cos(\bar{n}', \bar{r}')$$

代入后得到基尔霍夫衍射公式

$$\begin{aligned} U(P) &= \frac{1}{i\lambda} \iint_{\Sigma} \frac{Ae^{ikr'}}{r'} \left[\frac{\cos(n, r) - \cos(n, r')}{2} \right] \cdot \frac{e^{ikr}}{r} ds \\ &= \frac{1}{i\lambda} \iint_{\Sigma} U(P_0) K(\theta) \cdot \frac{e^{ikr}}{r} ds, \end{aligned}$$

与菲涅耳原理一致。

四、光场传播的线性性质

令 $h(P, P_0) = \frac{1}{i\lambda} K(\theta) \frac{\exp[ikr]}{r}$ 代入基尔霍夫衍射公式得

§ 2.3 衍射的角谱理论

一、角谱的传播

根据衍射理论：对（孔径平面）物平面，复振幅 $U(x_0, y_0)$ 为

$$U(x_0, y_0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d \frac{\cos \alpha}{\lambda} d \frac{\cos \beta}{\lambda} \times A_0 \left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda} \right) e^{i2\pi \left(\frac{\cos \alpha}{\lambda} x_0 + \frac{\cos \beta}{\lambda} y_0 \right)}$$

而对于（观察平面）像平面，复振幅 $U(x, y)$ 可写为：

$$U(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d \frac{\cos \alpha}{\lambda} d \frac{\cos \beta}{\lambda} \times A \left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda} \right) e^{i2\pi \left(\frac{\cos \alpha}{\lambda} x + \frac{\cos \beta}{\lambda} y \right)}$$

将上式代入衍射的亥姆霍兹方程

$$(\nabla^2 + k^2)U(P) = 0$$

得

$$0 = (\nabla^2 + k^2) \left\{ A \left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda} \right) e^{i2\pi \left(\frac{\cos \alpha}{\lambda} x + \frac{\cos \beta}{\lambda} y \right)} \right\}$$

因为 A 仅为 Z 的函数

$$\frac{\partial}{\partial x} A \left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda} \right) = \frac{\partial}{\partial y} A \left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda} \right) = 0$$

$$\frac{\partial}{\partial z} A \left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda} \right) = \frac{d}{dz} A \left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda} \right)$$

对于指数函数，有

$$\frac{\partial}{\partial x} e^{i2\pi \left(\frac{\cos \alpha}{\lambda} x + \frac{\cos \beta}{\lambda} y \right)} = (i2\pi \frac{\cos \alpha}{\lambda}) \times e^{i2\pi \left(\frac{\cos \alpha}{\lambda} x + \frac{\cos \beta}{\lambda} y \right)}$$

$$\frac{\partial}{\partial y} e^{i2\pi \left(\frac{\cos \alpha}{\lambda} x + \frac{\cos \beta}{\lambda} y \right)} = (i2\pi \frac{\cos \beta}{\lambda}) \times e^{i2\pi \left(\frac{\cos \alpha}{\lambda} x + \frac{\cos \beta}{\lambda} y \right)}$$

$$\frac{\partial}{\partial z} e^{i2\pi \left(\frac{\cos \alpha}{\lambda} x + \frac{\cos \beta}{\lambda} y \right)} = 0$$

代入亥姆霍兹方程，得到方程的基本解

$$A \left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda} \right) = C \left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda} \right) \times e^{ikz \sqrt{1 - \cos^2 \alpha - \cos^2 \beta}}$$

其中常量 $C \left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda} \right)$ 由初始条件决定。

在孔径平面 $z=0$ ，角谱为 $A_0 \left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda} \right)$ ，则

$$C\left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda}\right) = A_0\left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda}\right)$$

得到方程的解

$$A\left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda}\right) = A_0\left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda}\right) \times e^{ikz\sqrt{1-\cos^2 \alpha - \cos^2 \beta}}$$

这就是角谱理论的衍射公式。

下面讨论方向余弦对角谱理论的影响。

1) 当 $\cos^2 \alpha + \cos^2 \beta < 1$ 角谱空间中，传播仅仅引入相位移动，振幅不受影响。

2) 当 $\cos^2 \alpha + \cos^2 \beta > 1$ ，光波为倏逝波。因为

$$\sqrt{1 - \cos^2 \alpha - \cos^2 \beta} = i\mu$$

则观察平面的角谱

$$A\left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda}\right) = A_0\left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda}\right) e^{-\mu z}$$

随这传播距离的增大迅速衰减。

3) 当 $\cos^2 \alpha + \cos^2 \beta = 1$ ，

$\cos \gamma = 0, \gamma = 90^\circ$ 。理论表明光将沿垂直z轴的方向传播，与事实不符，显示出标量理论的限制性。此时应

该采取矢量理论更为合适。

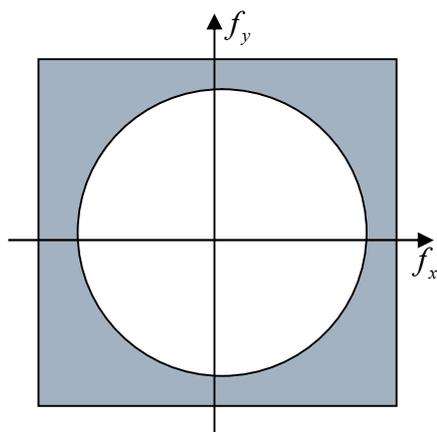
二、表征系统的脉冲响应

根据角谱与空间频率之间的关系，角谱的衍射理论可以写为

$$A(f_x, f_y) = A_0(f_x, f_y) e^{ikz\sqrt{1-(\lambda f_x)^2 - (\lambda f_y)^2}} = A_0(f_x, f_y) H(f_x, f_y),$$

其中传递函数定义为

$$H(f_x, f_y) = e^{ikz\sqrt{1-(\lambda f_x)^2 - (\lambda f_y)^2}}$$



忽略倏逝波，传递函数写为

$$H(f_x, f_y) = \begin{cases} e^{ikz\sqrt{1-(\lambda f_x)^2 - (\lambda f_y)^2}}, & f_x^2 + f_y^2 < 1 \\ 0, & \text{others} \end{cases}$$

公式表明，可以把光波的传播看成一个空间滤波器，它具有有限的带宽。

三、角谱理论与基尔霍夫衍射理论

基尔霍夫理论是衍射的球面波理论，考察点光场看作是无穷多球面子波叠加的结果；而角谱理论是衍射的平面波理论，考察点光场可以看作是无穷多平面波叠加的结果。

衍射的角谱理论表明物-像平面的空间频率之间满足

$$A(f_x, f_y) = A_0(f_x, f_y) e^{ikz\sqrt{1-(\lambda f_x)^2-(\lambda f_y)^2}},$$

而衍射屏处光场的频谱为

$$A_0(f_x, f_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U_0(x_0, y_0) e^{i2\pi[f_x x_0 + f_y y_0]} dx_0 dy_0$$

考察点处光场的傅里叶分解为

$$U(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(f_x, f_y) e^{i2\pi(f_x x + f_y y)} df_x df_y$$

将后两式代入角谱理论中可得

$$U(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U_0(x_0, y_0) dx_0 dy_0 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{ikz\sqrt{1-(\lambda f_x)^2-(\lambda f_y)^2}} e^{i2\pi[f_x(x-x_0)+f_y(y-y_0)]} df_x df_y$$

令 $h(x-x_0, y-y_0) = F^{-1}\{H(f_x, f_y)\}$ ，在近轴条件下，

$$\begin{aligned} h(x-x_0, y-y_0) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{ikz\sqrt{1-(\lambda f_x)^2-(\lambda f_y)^2}} \times e^{i2\pi[f_x(x-x_0)+f_y(y-y_0)]} df_x df_y \\ &= \frac{\exp[ik\sqrt{z^2+(x-x_0)^2+(y-y_0)^2}]}{i\lambda\sqrt{z^2+(x-x_0)^2+(y-y_0)^2}} \end{aligned}$$

代回前一式中，

$$U(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U_0(x_0, y_0) h(x-x_0, y-y_0) dx_0 dy_0 = U_0(x, y) \otimes h(x, y)$$

这一卷积的结果再一次证明，光的传播可看作线性空不变系统。基尔霍夫理论是在空间域讨论光的传播，把孔径平面看作点光源的集合，观察面上的光场分布是球面子波的叠加。角谱理论是在频率域讨论光的传播，把孔径平面看作平面波分量的组合，观察面的光场是这些平面波分量的叠加。两种衍射理论的一致性的根本原因在于标量的波动方程。

四、孔径对角谱的影响

衍射屏处的孔径用复振幅透过率函数表示

$$t(x_0, y_0) = \begin{cases} 1, & in \\ 0, & out \end{cases}$$

考虑基尔霍夫假设，屏后的光场分布

$$U_1(x_0, y_0) = U_0(x_0, y_0)t(x_0, y_0)$$

§ 2.4 菲涅尔衍射

菲涅尔衍射是指衍射屏到光源和观察点的距离为有限远的衍射现象。为了简化对衍射图样的数学计算，通常都要对衍射理论作某种程度的近似。

一、菲涅耳公式的导出

根据基尔霍夫衍射公式的卷积形式得观察平面的光场分布

$$U(x, y) = \iint_{-\infty}^{\infty} U(x_0, y_0) h(x - x_0, y - y_0) dx_0 dy_0$$

式中：

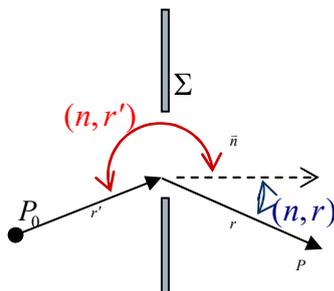
$$h(x - x_0, y - y_0) = \frac{1}{i\lambda r} e^{ikr},$$

$$r = [z^2 + (x - x_0)^2 + (y - y_0)^2]^{\frac{1}{2}}.$$

1. 基尔霍夫公式的简化

下面采用近轴条件对衍射公式进行简化。在近轴条件下，倾斜因子

$$K(\theta) = \frac{\cos(\vec{n}, \vec{r}) - \cos(\vec{n}, \vec{s})}{2} \approx \frac{1 - (-1)}{2} = 1.$$



将 r 按照级数展开：

$$\begin{aligned} r &= z \left[1 + \left(\frac{x - x_0}{z} \right)^2 + \left(\frac{y - y_0}{z} \right)^2 \right]^{\frac{1}{2}} \\ &= z \left\{ 1 + \frac{1}{2} \left[\left(\frac{x - x_0}{z} \right)^2 + \left(\frac{y - y_0}{z} \right)^2 \right] - \dots \right\} \end{aligned}$$

分母中的 r 可用 z 代替。分子中因 $k = \frac{2\pi}{\lambda}$ 值很大， r 需取更精确的近似，在展开式中略去第三项及其以后的项，得到

$$r = z \left[1 + \frac{1}{2} \left(\frac{x - x_0}{z} \right)^2 + \frac{1}{2} \left(\frac{y - y_0}{z} \right)^2 \right],$$

这种近似方法就是菲涅耳近似。将菲涅耳近似的结果代入 h 式得：

$$h(x-x_0, y-y_0) = \frac{e^{ikz}}{i\lambda z} e^{i\frac{k}{2z}[(x-x_0)^2+(y-y_0)^2]}.$$

代入 $U(x, y)$ 式得菲涅耳公式为

$$U(x, y) = \frac{e^{ikz}}{i\lambda z} \int \int_{-\infty}^{\infty} dx_0 dy_0 \cdot U(x_0, y_0) \times e^{i\frac{k}{2z}[(x-x_0)^2+(y-y_0)^2]}$$

2. 菲涅耳近似条件的适用范围

在 r 展开式的指数项中，被略去的第三项不应引起明显的位相误差，即

$$k \frac{[(x-x_0)^2+(y-y_0)^2]^2}{8z^3} \ll 1$$

移项得

$$z^3 \gg \frac{\pi}{4\lambda} [(x-x_0)^2+(y-y_0)^2]^2$$

或

$$\frac{[(x-x_0)^2+(y-y_0)^2]^2}{8z^3} \ll \frac{\lambda}{10}$$

上式称为菲涅耳衍射条件，满足该条件的区域称为菲涅耳区。

3. 菲涅耳衍射与傅里叶变换的关系

将菲涅耳衍射公式的指数项展开得

$$(x-x_0)^2+(y-y_0)^2 = x^2+y^2+x_0^2+y_0^2-2(xx_0+yy_0),$$

代入菲涅耳衍射公式

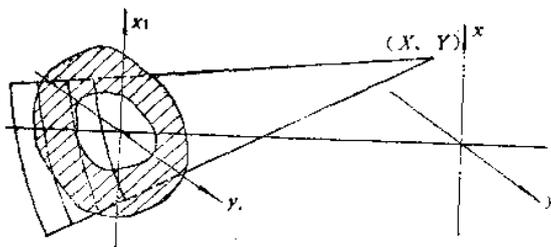
$$\begin{aligned} U(x, y) &= \frac{e^{ikz+ik\frac{x^2+y^2}{2z}}}{i\lambda z} \int \int_{-\infty}^{\infty} dx_0 dy_0 U(x_0, y_0) \times e^{i\frac{kx_0^2+y_0^2}{2z}-ik\frac{xx_0+yy_0}{z}} \\ &= \frac{e^{ikz}}{i\lambda z} e^{i\frac{k}{2z}(x^2+y^2)} \times F \left\{ U(x_0, y_0) e^{i\frac{k}{2z}(x_0^2+y_0^2)} \right\}, \end{aligned}$$

我们得到菲涅耳衍射公式的傅里叶变换表达形式。由此可见，观察平面上的复振幅分布正比于

$U(x_0, y_0) e^{i\frac{k}{2z}(x_0^2+y_0^2)}$ 的傅里叶变换，观察平面上光场的函数分布随着 z 增大会发生变化，沿 z 轴亮暗交替。

4. 会聚球面波照明的情况

为了消除菲涅耳衍射中 $e^{i\frac{k}{2z}(x_0^2+y_0^2)}$ 的影响，可以采用会聚球面波照明的方法，如图所示。



孔径平面 (x, y) 上光场的复振幅为

$$U_0(x_0, y_0) = \frac{a_0}{z} e^{-ik \frac{(x_0 - X)^2 + (y_0 - Y)^2}{2z}},$$

其中 (X, Y) 为球面波的会聚中心。则孔径后光场的复振幅分布为:

$$U'_0(x_0, y_0) = t(x_0, y_0) U_0(x_0, y_0)。$$

根据菲涅尔衍射公式，在观察平面 (x, y) 上的复振幅为:

$$\begin{aligned} U(x, y) &= \frac{e^{ikz + ik \frac{x^2 + y^2}{2z}}}{i\lambda z} F \left\{ U'_0(x_0, y_0) e^{ik \frac{x_0^2 + y_0^2}{2z}} \right\} \\ &= \frac{a_0 e^{ikz + ik \frac{x^2 + y^2}{2z}}}{i\lambda z^2} \times \int \int_{-\infty}^{\infty} dx_0 dy_0 \cdot t(x_0, y_0) e^{ik \frac{x_0^2 + y_0^2}{2z}} \\ &\quad \times e^{-ik \frac{(x_0 - X)^2 + (y_0 - Y)^2}{2z}} e^{-ik \frac{xx_0 + yy_0}{z}} \end{aligned}$$

整理得

$$\begin{aligned} U(x, y) &= \frac{a_0 \exp[ikz]}{i\lambda z^2} \exp\left[ik \frac{x^2 + y^2}{2z}\right] \\ &= \frac{a_0 e^{ikz + ik \frac{x^2 + y^2}{2z} - ik \frac{X^2 + Y^2}{2z}}}{i\lambda z^2} F \{t(x_0, y_0)\} \end{aligned}$$

式中傅里叶变换的变量对应于 $f_x = \frac{x - z}{\lambda z}$, $f_y = \frac{y - z}{\lambda z}$ 。观察屏上的菲涅耳衍射图样的复振幅分布与衍射屏上的复振幅透过率的傅里叶变换成正比。

二、菲涅尔衍射的角谱理论

衍射的角谱理论

$$A\left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda}\right) = A_0\left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda}\right) H\left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda}\right),$$

其中传输函数

$$H\left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda}\right) = e^{ikz\sqrt{1-\cos^2 \alpha - \cos^2 \beta}}$$

其相位部分按照级数展开，只取前两项

$$\sqrt{1-\cos^2 \alpha - \cos^2 \beta} \approx 1 - \frac{1}{2}(\cos^2 \alpha + \cos^2 \beta)$$

则

$$H\left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda}\right) = e^{ikz - i\frac{kz}{2}(\cos^2 \alpha + \cos^2 \beta)}$$

将传输函数作变量代换 $f_x = \frac{\cos \alpha}{\lambda}$, $f_y = \frac{\cos \beta}{\lambda}$ ，并改写为

$$H(f_x, f_y) = e^{ikz - i\pi\lambda z(f_x^2 + f_y^2)}。$$

对于观察平面的光场

$$U(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x_0, y_0) h(x-x_0, y-y_0) dx_0 dy_0$$

其中脉冲响应函数由上式传输函数的傅里叶逆变换得出

$$\begin{aligned} h(x-x_0, y-y_0) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} df_x df_y \cdot e^{ikz - i\pi\lambda z(f_x^2 + f_y^2)} e^{i2\pi[f_x(x-x_0) + f_y(y-y_0)]} \\ &= \frac{e^{ikz}}{i\lambda z} e^{i\frac{k}{2z}[(x-x_0)^2 + (y-y_0)^2]} \end{aligned}$$

代入观察平面的光场中得

$$\begin{aligned} U(x, y) &= \frac{e^{ikz}}{i\lambda z} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx_0 dy_0 U(x_0, y_0) e^{i\frac{k}{2z}[(x-x_0)^2 + (y-y_0)^2]} \\ &= \frac{e^{ikz + ik\frac{x^2+y^2}{2z}}}{i\lambda z} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx_0 dy_0 U(x_0, y_0) \times e^{ik\frac{x_0^2+y_0^2}{2z} - ik\frac{xx_0+yy_0}{z}} \\ &= \frac{e^{ikz + ik\frac{x^2+y^2}{2z}}}{i\lambda z} F \left\{ U(x_0, y_0) e^{ik\frac{x_0^2+y_0^2}{2z}} \right\} \end{aligned}$$

注意到傅里叶变换中 $f_x = \frac{x}{\lambda z}$, $f_y = \frac{y}{\lambda z}$ 。比较菲涅尔衍射公式和角谱理论发现，这两种理论等价！

三、塔尔伯特(Talbot)效应

塔尔伯特(Talbot)效应是菲涅尔衍射的一个令人感兴趣的效应，它是由塔尔伯特在 1830 年发现的。塔尔伯特(Talbot)效应是指周期性物体经菲涅尔衍射后，会产生自成像的现象。

考虑一维周期性物体

$$t(x) = \frac{1}{2}[1 + \beta \cos(2\pi x/d)],$$

我们利用角谱理论处理该物体的菲涅尔衍射。物体的角谱(傅里叶变换)为

$$T(f_x) = F\{t(x)\} = \int_{-\infty}^{\infty} t(x) e^{-i2\pi f_x x} dx = \frac{1}{2} \delta(f_x) + \frac{\beta}{4} \left[\delta\left(f_x - \frac{1}{d}\right) + \delta\left(f_x + \frac{1}{d}\right) \right].$$

菲涅尔衍射系统的传输函数为

$$H(f_x) = \exp[-i\pi\lambda z f_x^2] \exp[ikz].$$

根据以上分析得观察平面光场的角谱

$$\begin{aligned} T'(f_x) &= T(f_x)H(f_x) \\ &= e^{ikz - i\pi\lambda z f_x^2} \left[\frac{\delta(f_x)}{2} + \frac{\beta}{4} \delta\left(f_x - \frac{1}{d}\right) + \frac{\beta}{4} \delta\left(f_x + \frac{1}{d}\right) \right] \\ &= e^{ikz} \times \left[\frac{1}{2} \delta(f_x) + \frac{\beta}{4} \delta\left(f_x - \frac{1}{d}\right) e^{-i\frac{\pi\lambda z}{d^2}} + \frac{\beta}{4} \delta\left(f_x + \frac{1}{d}\right) e^{-i\frac{\pi\lambda z}{d^2}} \right] \end{aligned}$$

观察平面光场分布在空间域的形式

$$t'(x) = F^{-1}\{T'(f_x)\} = \frac{e^{ikz}}{2} + \frac{\beta}{2} e^{ikz - i\frac{\pi\lambda z}{d^2}} \cos\left(\frac{2\pi x}{d}\right)$$

因此衍射强度分布为

$$|t'(x)|^2 = \frac{1}{4} \left| 1 + \beta e^{-i\frac{\pi\lambda z}{d^2}} \cos(2\pi x/d) \right|^2.$$

下面讨论不同传播距离处的衍射强度分布情况。

(1) Talbot 像

$$\text{当 } \frac{\pi\lambda z}{d^2} = 2j\pi, \quad j=0,1,2,3,\dots \text{ 时, } z = \frac{2jd^2}{\lambda}, \text{ 强度分布为}$$

$$|t'(x)|^2 = \frac{1}{4} |1 + \beta \cos(2\pi x/d)|^2,$$

它与物体的强度分布是完全相同的。

(2) Talbot 反转像(副像)

$$\text{当 } \frac{\pi\lambda z}{d^2} = (2j+1)\pi \text{ 时, } z = \frac{(2j+1)d^2}{\lambda}, \text{ 强度分布为}$$

$$|t'(x)|^2 = \frac{1}{4} |1 - \beta \cos(2\pi x/d)|^2,$$

它与物体的强度分布正好相反。

(3) Talbot 子像,

$$\text{当 } \frac{\pi\lambda z}{d^2} = \left(j + \frac{1}{2}\right)\pi \text{ 时, } z = \left(j + \frac{1}{2}\right) \frac{d^2}{\lambda}, \text{ 强度分布为}$$

§ 2.5 夫琅禾费衍射

夫琅禾费衍射是指衍射屏到光源和观察点的距离为无穷远的衍射现象。在处理夫琅禾费衍射图样时，需要对衍射理论作更进一步的近似。

一、夫琅禾费衍射公式

1. 夫琅禾费近似

从基尔霍夫衍射定理的卷积形式出发

$$U(x, y) = \iint_{-\infty}^{\infty} U(x_0, y_0) h(x - x_0, y - y_0) dx_0 dy_0,$$

式中脉冲响应函数为

$$h(x - x_0, y - y_0) = \frac{1}{i\lambda r} \exp[ikr]$$

其中在 r 的展开式中忽略 xy 的三次幂和 $x_0 y_0$ 的二次幂以上的项，得到

$$r = \left[z^2 + (x - x_0)^2 + (y - y_0)^2 \right]^{\frac{1}{2}}$$

$$\approx z + \frac{x^2 + y^2}{2z} - \frac{xx_0 + yy_0}{z},$$

这就是夫琅和费近似。将之代入 $U(x, y)$ 式，得夫琅和费衍射公式为

$$U(x, y) = \frac{e^{ikz + ik\frac{x^2 + y^2}{2z}}}{i\lambda z} \iint_{-\infty}^{\infty} U(x_0, y_0) e^{-ik\frac{xx_0 + yy_0}{z}} dx_0 dy_0$$

2. 夫琅和费近似条件及范围

在夫琅禾费衍射近似中，忽略的 $x_0 y_0$ 的二次幂应满足 $k \frac{(x_0^2 + y_0^2)}{2z} \ll 1$ ，移项得 $\frac{(x_0^2 + y_0^2)}{2z} \ll \frac{\lambda}{2\pi}$ 。例

如采用波长为 $\lambda = 500\text{nm}$ ，边长为 1mm 的方形孔，其夫琅禾费衍射区域为：

$$\frac{2\pi (x_0^2 + y_0^2)}{\lambda \cdot 2z} \leq 1$$

可得衍射距离满足

$$z \geq \frac{\pi(x_0^2 + y_0^2)}{\lambda} = \frac{\pi[(0.5 \times 10^{-3})^2 + (0.5 \times 10^{-3})^2]}{500 \times 10^{-9}}$$

$$= \frac{3.14 \times 0.5 \times 10^{-6}}{500 \times 10^{-9}} = 3.14(\text{m})$$

二、夫琅和费衍射与傅里叶变换的关系

夫琅禾费衍射公式

$$U(x, y) = \frac{e^{ikz + ik\frac{x^2+y^2}{2z}}}{i\lambda z} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x_0, y_0) e^{-ik\frac{xx_0 + yy_0}{z}} dx_0 dy_0$$

将积分部分改写为：

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x_0, y_0) e^{-i2\pi\left(\frac{x \cdot x_0}{\lambda z} + \frac{y \cdot y_0}{\lambda z}\right)} dx_0 dy_0 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x_0, y_0) e^{-i2\pi(f_x x_0 + f_y y_0)} dx_0 dy_0 = F\{U(x_0, y_0)\}$$

代入弗朗禾费衍射公式 $U(x, y)$ 得：

$$U(x, y) = \frac{e^{ikz + i\frac{k}{2z}(x^2+y^2)}}{i\lambda z} F\{U(x_0, y_0)\}$$

上式表明，观察屏上光场复振幅分布正比于衍射屏上透射光复振幅分布的傅里叶变换。特殊地，当用单位振幅平面波垂直照明衍射物时衍射屏平面上透射光场的复振幅为

$$U(x_0, y_0) = 1 \cdot t(x_0, y_0) = t(x_0, y_0)$$

代入 $U(x_0, y_0)$ 得：

$$U(x, y) = \frac{e^{ikz + i\frac{k}{2z}(x^2+y^2)}}{i\lambda z} F\{t(x_0, y_0)\}$$

三、夫琅和费衍射举例

1. 矩形孔的衍射



矩形孔的复振幅透过率为：

$$t(x_0, y_0) = \text{rect}\left(\frac{x_0}{a}\right) \text{rect}\left(\frac{y_0}{b}\right)$$

由傅里叶变换得矩形孔的频谱为

$$\begin{aligned} F\{t(x_0, y_0)\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{rect}\left(\frac{x_0}{a}\right) \text{rect}\left(\frac{y_0}{b}\right) \times e^{-i2\pi(f_x x_0 + f_y y_0)} dx_0 dy_0 \\ &= F\left\{\text{rect}\left(\frac{x_0}{a}\right)\right\} \times F\left\{\text{rect}\left(\frac{y_0}{b}\right)\right\} \end{aligned}$$

其中

$$F\left\{\text{rect}\left(\frac{x_0}{a}\right)\right\} = \frac{a}{-i2\pi a f_x} e^{-i2\pi a f_x x} \Big|_{-1/2}^{1/2} = \text{sinc}(a f_x),$$

$$F \{ \text{rect}(y_0/b) \} = b \text{sinc}(bf_y),$$

所以 $F \{ t(x_0, y_0) \} = ab \text{sinc}(af_x) \text{sinc}(bf_y)$ 。

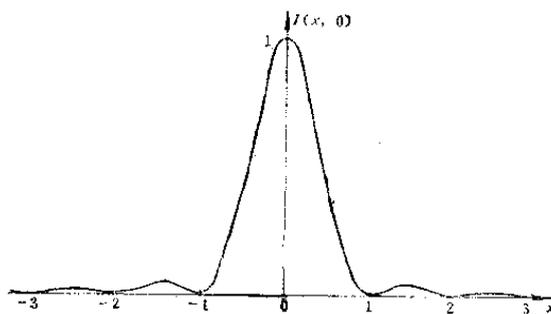
根据夫琅和费衍射公式得：

$$\begin{aligned} U(x, y) &= \frac{e^{ikz + ik \frac{x^2 + y^2}{2z}}}{i\lambda z} F \{ t(x_0, y_0) \} \\ &= \frac{e^{ikz + ik \frac{x^2 + y^2}{2z}}}{i\lambda z} \cdot ab \text{sinc}(af_x) \text{sinc}(bf_y) \\ &= \frac{ab e^{ikz + ik \frac{x^2 + y^2}{2z}}}{i\lambda z} \text{sinc}\left(\frac{ax}{\lambda z}\right) \text{sinc}\left(\frac{by}{\lambda z}\right) \end{aligned}$$

得衍射平面的光强度为：

$$I(x, y) = \left(\frac{ab}{\lambda z}\right)^2 \cdot \text{sinc}^2\left(\frac{ax}{\lambda z}\right) \text{sinc}^2\left(\frac{by}{\lambda z}\right)$$

光强在 x 方向的分布如图所示。



当 $x = 0$, $y = 0$ 时, $\text{sinc}(0) = 1$, 矩形孔衍射强度分布的中央 (0 级) 最大光强为

$$I(0,0) = \left(\frac{ab}{\lambda z}\right)^2$$

当 $\frac{xa}{\lambda z} = \pm 1$ 时, U, I 有零值, 故得中央亮斑的宽度为

$$\Delta x = 2 \frac{\lambda z}{a}, \quad \Delta y = 2 \frac{\lambda z}{b}.$$

2. 圆孔衍射

圆孔的复振透过率

$$t(r_0) = \text{circ}\left(\frac{r_0}{a}\right) = \begin{cases} 1 & |r_0| \leq a \\ 0 & \text{其它} \end{cases}$$

为圆对称函数, 其傅里叶氏变换用傅里叶-贝塞尔变换表示

$$B\{\text{circ}(r_0)\} = 2\pi \int_0^{\infty} \text{circ}(r) r_0 J_0(2\pi r_0 \rho) dr_0。$$

根据第一类贝塞尔函数的性质,

$$\begin{aligned} 2\pi \int_0^1 r_0 J_0(2\pi r_0 \rho) dr_0 &= 2\pi \int_0^{2\pi\rho} \frac{r'}{2\pi\rho} J_0(r') \frac{dr'}{2\pi\rho} \\ &= \frac{1}{2\pi\rho^2} \int_0^{2\pi\rho} r' J_0(r') dr' \\ &= \frac{1}{2\pi\rho^2} 2\pi\rho \cdot J_1(2\pi\rho) \\ &= \frac{J_1(2\pi\rho)}{\rho}。 \end{aligned}$$

由傅里叶-贝塞尔变换的缩放性质

$$B\left\{g_r\left(\frac{r}{a}\right)\right\} = |a|^2 G(a\rho)$$

得

$$B\left\{\text{circ}\left(\frac{r_0}{a}\right)\right\} = a^2 B\{\text{circ}(a\rho)\} = a^2 \left[\frac{J_1(2\pi a\rho)}{a\rho} \right] = \pi a^2 \frac{2J_1(2\pi a\rho)}{2\pi a\rho}$$

故得圆孔夫琅和费衍射为:

$$\begin{aligned} U(r) &= \frac{1}{i\lambda z} e^{ikz} e^{i\frac{k}{2z}r^2} B\{t(r_0)\} \\ &= \frac{1}{i\lambda z} e^{ikz+i\frac{kr^2}{2z}} \cdot \pi a^2 \left[\frac{2J_1(2\pi ar/\lambda z)}{2\pi ar/\lambda z} \right] \\ &= \frac{ka^2}{i2z} e^{ikz} e^{i\frac{k}{2z}r^2} \cdot \left[\frac{2J_1(kar/z)}{kar/z} \right] \end{aligned}$$

强度分布为:

$$I(r) = \left(\frac{ka^2}{2z} \right)^2 \left[\frac{2J_1(kar/z)}{kar/z} \right]^2。$$

一阶贝塞尔函数在当 $r=0$ 时存在极限

$$\lim_{r \rightarrow 0} \frac{J_1(kar/z)}{kar/z} = \frac{1}{2}。$$

因此, 衍射斑中心的强度分布为(光强)

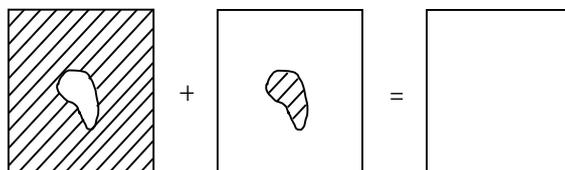
$$I(0) = \left(\frac{ka^2}{2z} \right)^2 \left[2 \times \frac{1}{2} \right]^2 = \left(\frac{ka^2}{2z} \right)^2，$$

故得：
$$I(r) = I(0) \left[\frac{2J_1(kar/z)}{kar/z} \right]。$$

-----*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-----

§ 2.6 巴比涅原理

一、互补屏



如图所示的两个衍射屏，其中一个衍射屏的开孔部分和另一个衍射屏的不透明部分恰好对应，反之亦然，则称这一对衍射屏为互补屏。互补屏的透过率函数分别用 $t_1(x)$ 和 $t_2(x)$ 表示，则

$$t_1(x, y) + t_2(x, y) = t(x, y) = 1, \quad t_2(x, y) = t(x, y) - t_1(x, y)。$$

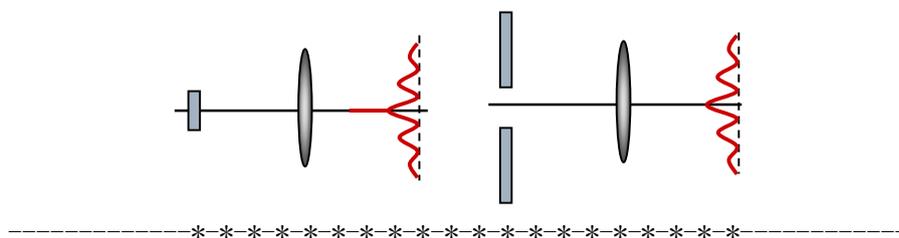
二、巴卑涅原理

互补屏的频谱满足

$$F \{t_1(x, y)\} + F \{t_2(x, y)\} = F \{t(x, y)\} = \delta(f_x, f_y),$$

上式表明，两个互补屏的空间频谱互补为 δ 函数。由弗朗禾费衍射可知，互补屏在远处观察点产生的衍射光场，其振幅之和等于光波自由传播时在该点的复振幅，这一结论称为巴卑涅原理。

由于函数 $\delta(f_x, f_y)$ 除了原点之外，其余位置的取值均为零，我们可以得到如下的结论：对于互补屏产生的弗朗禾费衍射分布，除了轴上点以外，其余各点的强度分布完全一样；在每一轴外点产生的光场复振幅相位相差 π 。

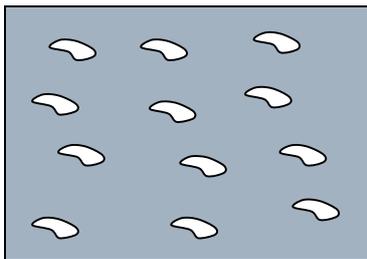


§ 2.7 衍射光栅

光栅具有周期性重复排列的结构，它是光学仪器和光学信息处理中重要的光学元件，可以对入射光场的振幅和相位施加周期性的空间调制。

一、列阵定理

假设如图所示的小孔的透过率函数为 $t_0(x_0, y_0)$ ，其对应的空间频谱为 $T_0(f_x, f_y)$ 。制作一衍射屏，上面开有 N 个与该小孔形状完全相同的小孔，保持这些孔径的空间取向一致。



这一小孔列阵构成的衍射屏，其透过率函数表示为 N 个空的组合

$$t(x_0, y_0) = \sum_{n=1}^N t_0(x_0 - \xi_n, y_0 - \eta_n)。$$

利用 δ 函数的筛选性质，可以将透过率函数写成卷积的形式

$$t(x_0, y_0) = t_0(x_0, y_0) \otimes \sum_{n=1}^N \delta(x_0 - \xi_n, y_0 - \eta_n)。$$

即整个衍射屏的透过率函数可以用单孔函数和 δ 函数的卷积表示。根据傅里叶变换的卷积定理，可以求出小孔列阵衍射屏的傅里叶变换频谱为

$$\begin{aligned} T(f_x, f_y) &= F \{t(x_0, y_0)\} \\ &= F \{t_0(x_0, y_0)\} \times F \left\{ \sum_{n=1}^N \delta(x_0 - \xi_n, y_0 - \eta_n) \right\} \\ &= T_0(f_x, f_y) \times \sum_{n=1}^N \exp[-i2\pi(f_x \xi_n + f_y \eta_n)]， \end{aligned}$$

上式表明：取向相同的同形孔径构成的列阵，其频谱等于单个基元孔径频谱与排列成同样组态的点光源列阵的频谱的乘积。这就是列阵定理。

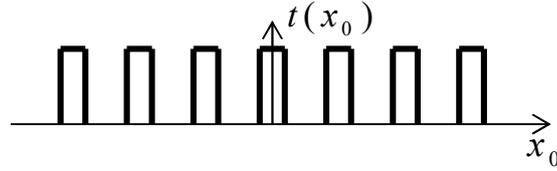
衍射屏的功率谱为

$$\left| T(f_x, f_y) \right|^2 = \left| T_0(f_x, f_y) \right|^2 \times \left| \sum_{n=1}^N \exp[-i2\pi(f_x \xi_n + f_y \eta_n)] \right|^2。$$

因此，当用单色平面波垂直照明衍射屏时，整个衍射屏的弗朗禾费衍射图样正是单个孔径衍射图样和 N 个点源列阵产生的多光束干涉的乘积。

二、线光栅

线光栅是最常见的光栅，其制作方法是在一块玻璃片上刻上大量宽度和间距均相等的平行线条，刻痕部位不透光，刻痕之间的光滑部分透光。线光栅是一种最简单的透射型振幅光栅。



线光栅可以认为由无穷多平行狭缝构成的，其透过率函数

$$t(x_0) = \text{rect}\left(\frac{x_0}{a}\right) \otimes \sum_{n=-\infty}^{\infty} \delta(x_0 - nd),$$

其中 a 为狭缝的宽度， d 为相邻狭缝的中心间距。根据列阵定理，光栅的频谱为

$$\begin{aligned} T(f_x) &= F \left\{ \text{rect}\left(\frac{x_0}{a}\right) \right\} \times F \left\{ \sum_{n=-\infty}^{\infty} \delta(x_0 - nd) \right\} \\ &= a \text{sinc}(a f_x) \times \text{comb}(d f_x) \end{aligned}$$

衍射图样中的各级条纹是分离的。

在实际当中，由于光栅或光束的尺寸总是有限的，假设光栅的整体孔径是边长为 L 的正方形，可以用矩形函数表示对透过率函数的限制

$$t(x_0, y_0) = \left[\text{rect}\left(\frac{x_0}{a}\right) \otimes \sum_{n=-\infty}^{\infty} \delta(x_0 - nd) \right] \times \text{rect}\left(\frac{x_0}{L}\right) \text{rect}\left(\frac{y_0}{L}\right)$$

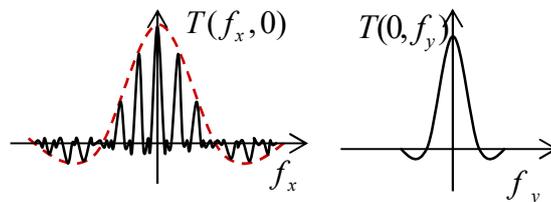
光栅的频谱为

$$\begin{aligned} T(f_x, f_y) &= [a \text{sinc}(a f_x) \times \text{comb}(d f_x)] \otimes L^2 \text{sinc}(L f_x) \text{sinc}(L f_y) \\ &= \frac{a}{d} \sum_{n=-\infty}^{\infty} \text{sinc}(a f_x) \delta\left(f_x - \frac{n}{d}\right) \otimes L^2 \text{sinc}(L f_x) \text{sinc}(L f_y) \\ &= \frac{a L^2}{d} \text{sinc}(L f_y) \sum_{n=-\infty}^{\infty} \text{sinc}\left(\frac{na}{d}\right) \text{sinc}\left[L\left(f_x - \frac{n}{d}\right)\right] \end{aligned}$$

其夫朗禾费衍射光场的复振幅分布为

$$\begin{aligned} U(x, y) &= \frac{e^{ikz + i\frac{k}{2z}(x^2 + y^2)}}{i\lambda z} \text{sinc}\left(\frac{Ly}{\lambda z}\right) \times T(f_x, f_y) \Big|_{f_x = \frac{x}{\lambda z}, f_y = \frac{y}{\lambda z}} \\ &= \frac{a L^2}{i\lambda z d} e^{ikz + i\frac{k}{2z}(x^2 + y^2)} \times \sum_{n=-\infty}^{\infty} \text{sinc}\left(\frac{na}{d}\right) \text{sinc}\left[L\left(\frac{x}{\lambda z} - \frac{n}{d}\right)\right]. \end{aligned}$$

由上式可以看出：谱线的轻度分布受到单缝衍射图样的调制；只要频谱间隔 $\frac{\lambda z}{d}$ 足够大，各谱线在空间中分离，光栅宽度 L 越大，谱线越窄，光栅的分辨本领越高。

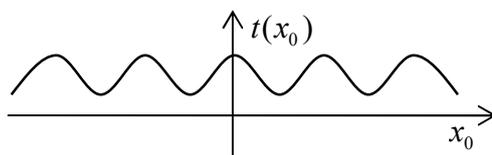


三、余弦振幅光栅

余弦型振幅光栅的透过率函数具有在 0~1 之间变化的余弦形式，透过率函数 (参见下图) 写成

$$t(x_0, y_0) = \left[\frac{1}{2} + \frac{m}{2} \cos(2\pi f_0 x_0) \right] \times \text{rect}\left(\frac{x_0}{L}\right) \text{rect}\left(\frac{y_0}{L}\right)$$

公式中考虑了光栅的有限尺寸， $m/2$ 是余弦变化的幅度， f_0 为光栅的频率。



一般而言， $f_0 \gg 2/L$ 。利用卷积定理计算光栅的频谱为

$$T(f_x, f_y) = F \left\{ \frac{1}{2} + \frac{m}{2} \cos(2\pi f_0 x_0) \right\} \otimes F \left\{ \text{rect}\left(\frac{x_0}{L}\right) \text{rect}\left(\frac{y_0}{L}\right) \right\},$$

其中

$$F \left\{ \frac{1}{2} + \frac{m}{2} \cos(2\pi f_0 x_0) \right\} = \frac{1}{2} \delta(f_x, f_y) + \frac{m}{4} \delta(f_x + f_0, f_y) + \frac{1}{4} \delta(f_x - f_0, f_y)$$

$$F \left\{ \text{rect}\left(\frac{x_0}{L}\right) \text{rect}\left(\frac{y_0}{L}\right) \right\} = L^2 \text{sinc}(Lf_x) \times \text{sinc}(Lf_y)$$

所以

$$T(f_x, f_y) = \frac{L^2}{2} \text{sinc}(Lf_y) \left\{ \text{sinc}(Lf_x) + \frac{m}{2} \text{sinc}[L(f_x + f_0)] + \frac{m}{2} \text{sinc}[L(f_x - f_0)] \right\}$$

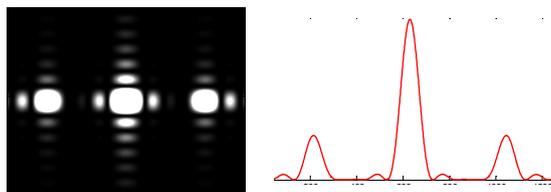
余弦型振幅光栅的 Fraunhofer 衍射光场分布为

$$U(x, y) = \frac{e^{ikz}}{i\lambda z} e^{i\frac{k}{2z}(x^2+y^2)} T(f_x, f_y)$$

注意到将 $f_x = \frac{x}{\lambda z}, f_y = \frac{y}{\lambda z}$ 代入上式得

$$U(x, y) = \frac{L^2 e^{ikz}}{i2\lambda z} e^{i\frac{k}{2z}(x^2+y^2)} \text{sinc}\left(\frac{Ly}{\lambda z}\right) \times \left\{ \text{sinc}\left(\frac{Lx}{\lambda z}\right) + \frac{m}{2} \text{sinc}\left[\frac{L}{\lambda z}(x + \lambda z f_0)\right] + \frac{m}{2} \text{sinc}\left[\frac{L}{\lambda z}(x - \lambda z f_0)\right] \right\},$$

显然衍射光场有三组间隔为 $\lambda z f_0$ 的 sinc 函数组成。下图给出衍射光场强度分布及截面图，供参考。



余弦振幅型光栅的分辨本领为 $P = jN$ ，此处 $j=1$ 。如果考虑波长为 λ 和 λ' 的光入射到余弦振幅型光栅，观察它们的一级衍射谱线。根据瑞利判据，一种波长谱线的强度极大值与另一种波长谱线的强度极小值重合，则认为可以分辨这两个波长，即

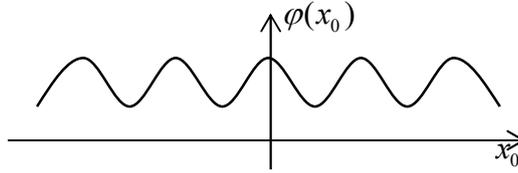
$$f_0 \lambda' z - f_0 \lambda z = \frac{\lambda z}{L}$$

整理后可得 $\frac{\lambda}{\lambda' - \lambda} = f_0 L$ ， $P = f_0 L = N$ ，余弦振幅型光栅的分辨本领与光栅上的条纹数目成正比！

四、正弦相位光栅

当光栅完全透明时，不能进行振幅调制。但是光栅的光学厚度的变化可以进行相位调制，这就是所谓的投射型相位光栅。我们研究薄的正弦型相位光栅，其复振幅透过率函数为

$$t(x_0, y_0) = \exp\left[i \frac{m}{2} \sin(2\pi f_0 x_0)\right] \times \text{rect}\left(\frac{x_0}{L}\right) \text{rect}\left(\frac{y_0}{L}\right)$$



式中考虑了光栅的有限尺寸， $m/2$ 是相位变化的幅度， f_0 为相位变化的频率。

利用卷积定理可得光栅的频谱为

$$T(f_x, f_y) = F\left\{e^{i \frac{m}{2} \sin(2\pi f_0 x_0)}\right\} \otimes F\left\{\text{rect}\left(\frac{x_0}{L}\right) \text{rect}\left(\frac{y_0}{L}\right)\right\},$$

利用贝塞尔函数的性质得到

$$F\left\{e^{i \frac{m}{2} \sin(2\pi f_0 x_0)}\right\} = \sum_{q=-\infty}^{\infty} J_q\left(\frac{m}{2}\right) \delta(f_x - qf_0, f_y)$$

其中 $J_q(x)$ 是 q 阶第一类贝塞尔函数。而

$$F\left\{\text{rect}\left(\frac{x_0}{L}\right) \text{rect}\left(\frac{y_0}{L}\right)\right\} = L^2 \text{sinc}(Lf_x) \times \text{sinc}(Lf_y),$$

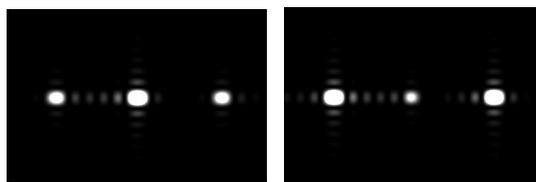
所以

$$T(f_x, f_y) = L^2 \text{sinc}(Lf_y) \times \sum_{q=-\infty}^{\infty} J_q\left(\frac{m}{2}\right) \text{sinc}[L(f_x - qf_0)].$$

正弦相位光栅的夫朗禾费衍射光场复振幅分布为

$$\begin{aligned} U(x, y) &= \frac{1}{i\lambda z} e^{ikz + i\frac{k}{2z}(x^2 + y^2)} T(f_x, f_y) \\ &= \frac{L^2}{i\lambda z} e^{ikz + i\frac{k}{2z}(x^2 + y^2)} \text{sinc}\left(\frac{Ly}{\lambda z}\right) \times \sum_{q=-\infty}^{\infty} J_q\left(\frac{m}{2}\right) \text{sinc}\left[\frac{L}{\lambda z}(x - q\lambda z f_0)\right] \end{aligned}$$

其中利用代换 $f_x = \frac{x}{\lambda z}, f_y = \frac{y}{\lambda z}$ 。由于一类贝赛尔函数可以是复数或零，因此参数 m 不同时，衍射图样的分布也不尽相同。下面给出了当 $m=1$ 和 2 时的衍射图样强度分布模拟图。



理论模拟 $m=1$ $m=2$.

-----*-----

§ 2.8 分数傅立叶变换与菲涅尔衍射

1980年, Namias建立了完整的分数傅里叶变换理论, 在二十世纪九十年代初期, Lohmann等人将分数傅里叶变换的方法和手段引入光学领域。分数傅里叶变换在光学信息处理方面有较为重要的应用。本节讨论分数傅里叶变换和菲涅尔衍射之间的联系。

一、分数傅里叶变换

1. 定义

一维函数的分数傅里叶变换定义为

$$G(\xi) = F_{\alpha}\{g(x)\} = \sqrt{\frac{e^{-i(\frac{\pi}{2}-\alpha)}}{2\pi \sin \alpha}} \int_{-\infty}^{\infty} e^{i\frac{(\xi^2+x^2)}{2 \tan \alpha} - i\frac{\xi x}{\sin \alpha}} g(x) dx$$

式中 $G(\xi)$ 称为 $g(x)$ 的 α 阶分数傅里叶变换, 或分数傅里叶谱 ($|\alpha| \leq \pi$)。由于阶数 α 可以是整数, 也可以是分数, 因此有人将分数傅里叶变换称之为广义傅里叶变换。

特别地, 当 $\alpha = \pi/2$ 或 $\alpha = -\pi/2$ 时, 分数傅里叶变换变成常规的傅里叶变换和傅里叶逆变换。而当 $\alpha = 0, \alpha = \pi$ 时, 分数傅里叶变换变成

$$F_0\{g(x)\} = g(\xi), \quad F_{\pi}\{g(x)\} = g(-\xi)。$$

以上两式表明, 0 阶分数傅里叶变化给出函数本身; π 阶分数傅里叶变换则给出它的倒像。

2. 分数傅里叶变换的性质

(1) 线性性质 $F_{\alpha}\{Ag(x) + Bh(x)\} = AF_{\alpha}\{g(x)\} + BF_{\alpha}\{h(x)\}$

(2) 位移性质

$$F_{\alpha}\{g(x+a)\} = e^{i\alpha \sin \alpha (\xi + \frac{\alpha \cos \alpha}{2})} \times G(\xi + \alpha \cos \alpha)$$

(3) 可加性

$$F_{\alpha}\{F_{\beta}\{g(x)\}\} = F_{\alpha}F_{\beta}\{g(x)\} = F_{\alpha+\beta}\{g(x)\}$$

(4) 可交换性

$$F_{\alpha}\{F_{\beta}\{g(x)\}\} = F_{\beta}F_{\alpha}\{g(x)\}$$

(5) 周期性 $F_{2n\pi+\alpha}\{\{g(x)\}\} = F_{\alpha}\{g(x)\}$

以上的分数傅里叶性质均可以推广到二维情形。文献中分数傅里叶变换的阶数有不同的表示方式, 设 $p = \frac{\alpha}{\pi}$, 则 α 阶的分数傅里叶变换可表示为 $F^p\{g(x)\}$, p 的变化范围 $|p| \leq 1$ 。

二、菲涅尔衍射与分数傅里叶变换

将一维形式的分数傅里叶变换改写为矢量表示的二维形式

$$G(\mathbf{s}) = \sqrt{\frac{\exp[-i(\frac{\pi}{2} - \alpha)]}{2\pi \sin \alpha}} \exp[\frac{i \times s^2}{2 \tan \alpha}] \int_{-\infty}^{\infty} g(\mathbf{r}) \exp[i \frac{r^2}{2 \tan \alpha} - i \frac{\mathbf{r} \cdot \mathbf{s}}{\sin \alpha}] d\mathbf{r}$$

用矢量 \mathbf{r} 表示衍射孔径坐标 (x_0, y_0) 用矢量 \mathbf{s} 表示考察平面坐标 (x, y) , 菲涅尔衍射公式也作同样形式的改写

$$U(\mathbf{s}) = \frac{e^{ikz + i\frac{k}{2z}s^2}}{i\lambda z} \times \int_{-\infty}^{\infty} U(\mathbf{r}) e^{i\frac{kr^2}{2z} - i\frac{2\pi}{\lambda z}\mathbf{r} \cdot \mathbf{s}} d\mathbf{r}$$

比较两公式发现它们之间有相类似的结构。为了找到菲涅尔衍射和分数傅里叶变换之间的对应关系，作如下变量代换：

$$\begin{aligned} \boldsymbol{\rho} &= \boldsymbol{\mu} \cdot \mathbf{r} = \sqrt{2\pi \tan \alpha / (\lambda z)} \cdot \mathbf{r}, \\ \boldsymbol{\sigma} &= \boldsymbol{\nu} \cdot \mathbf{s} = \sqrt{2\pi \sin \alpha \cos \alpha / (\lambda z)} \cdot \mathbf{s} \end{aligned}$$

代入菲涅尔衍射公式中，得到

$$U(\frac{\boldsymbol{\sigma}}{\boldsymbol{\nu}}) = C e^{i\frac{\tan \alpha}{2}\sigma^2} F_{\alpha}\{U(\frac{\boldsymbol{\rho}}{\boldsymbol{\mu}})\},$$

其中常数 $C = \frac{1}{i\lambda z \mu} e^{ikz}$ 。公式表明，在观察平面上的菲涅尔衍射分布等于孔径平面投射光场分布的 α 阶

分数傅里叶变换与一个二次相位因子的乘积。

需要注意的是，缩放因子 μ 和 ν 提供了满足变换关系的条件。对不同阶数 α 和距离 z , μ 和 ν 的值是不同的。其次，二次相位因子并不影响强度的探测。观察平面上菲涅尔衍射的光强分布正比于孔径平面投射场分布 $U(x_0, y_0)$ 的 α 阶分数傅里叶变换的模的平方。

三、透镜的分数傅里叶变换

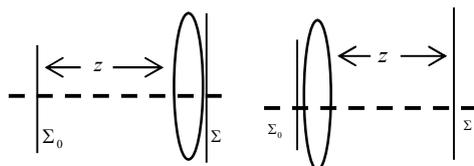
分数傅里叶变换除了利用菲涅尔衍射可以实现以外，还可以借助于透镜对光场的变换作用实现。如下图所示，在观察平面处放置一焦距为 f 的透镜，当孔径与透镜的距离满足

$$f = \frac{z}{\sin^2 \alpha}$$

时，在透镜后的观察平面 Σ 上将得到孔径平面透射光场的 α 阶分数傅里叶变换

$$U(\boldsymbol{\sigma}) = CF_{\alpha}\{U(\boldsymbol{\rho})\},$$

其中仍需要考虑变量代换和两个平面缩放因子的条件。



上面图中提供了准确实现 α 阶分数傅里叶变换的另一种方案。正透镜紧贴在孔径平面，此时

第三章 透镜的傅立叶变换性质

透镜是光学中最常用的光学元器件之一。在几何光学中，透镜被用以成像、聚焦和准直。信息光学中，对平面物体进行傅里叶变换的手段是实现它的弗朗禾费衍射。这一变换手段通常借助于透镜得到实现。本章所研究的透镜的傅里叶变换作用是光学信息处理方法的基础。

数学上处理透镜一般应用薄透镜近似。所谓的薄透镜是指：若一条光线从透镜的一面上坐标为 (x, y) 的点射入，而在相对的另一面也从近似相同的坐标处射出。即如果可以忽略光线在透镜内的平移，则此透镜是一个薄透镜。

§ 3.1 透镜的位相调制

一、透镜的复振幅透过率

为研究透镜对光波的作用，引入复振幅透过率，定义为

$$t_l(x, y) = \frac{U'_l(x, y)}{U_l(x, y)},$$

其中 $U_l = a_l(x, y)\exp[i\phi_l(x, y)]$ 是透镜前表面光场复振幅； $U'_l = a'_l(x, y)\exp[i\phi'_l(x, y)]$ 是透镜后表面光场的复振幅。将两光场的复振幅代入分振幅透过率函数，得

$$t_l(x, y) = \frac{a'_l(x, y)}{a_l(x, y)} e^{i[\phi'_l(x, y) - \phi_l(x, y)]} = A(x, y) e^{i\phi(x, y)}$$

若忽略透镜对入射光能量的损耗(忽略吸收和反射)， $A(x, y) = 1$ ，则

$$t_l(x, y) = e^{i\phi(x, y)}.$$

该式表明，透镜对入射光场的作用：光场位置不变，但有位相移动，薄透镜相当于一个位相变换器。

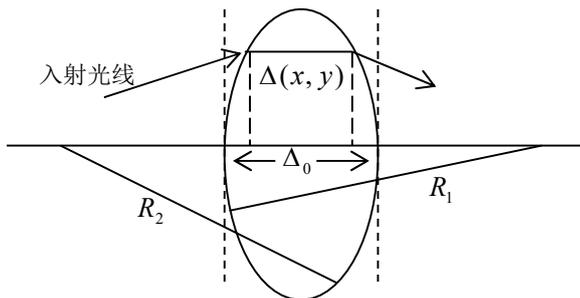
二、薄透镜的位相变换函数

在薄透镜的情况下，忽略由于折射所引起的传播距离的差值，可以近似认为光线在透镜上的入射点和出射点具有相同的坐标，以简化问题的分析。

1. 透镜厚度函数

设透镜的最大厚度为 Δ_0 ，在位置 (x, y) 处的厚度为 $\Delta(x, y)$ ，于是光线通过透镜 (x, y) 处产生的位相落后为：

$$\phi(x, y) = kr = kn\Delta(x, y) + k[\Delta_0 - \Delta(x, y)]$$



代入 $t(x, y)$ 式得:

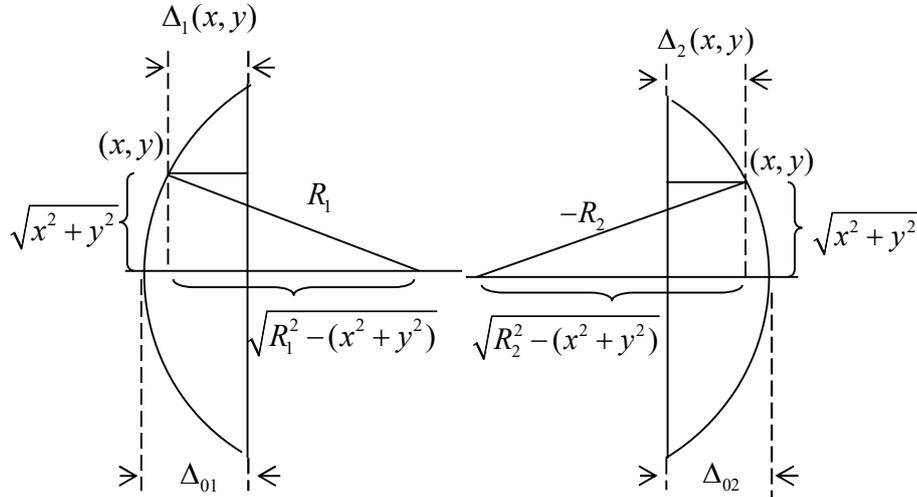
$$t_l(x, y) = e^{ik\Delta_0} e^{ik(n-1)\Delta(x, y)},$$

式中 n 为透镜材料的折射率。可见, 关键在于求出厚度函数 $\Delta(x, y)$ 。对于透镜而言, 还需要找出厚度函数与透镜的主要结构参数 (构成透镜的两个球面的曲率半径 R_1 和 R_2) 之间的关系, 才能最终确定透镜的位相调制。

采用笛卡尔符号法则: 由球面的顶点量到球心, 光线自左向右传播, 向左为负, 向右为正。在本例中, $R_1 > 0, R_2 < 0$ 。

为了求出 $\Delta(x, y)$, 将透镜剖成两半, 如下页图:

$$\Delta(x, y) = \Delta_1(x, y) + \Delta_2(x, y)$$



而:

$$\Delta_1(x, y) = \Delta_{01} - \left(R_1 - \sqrt{R_1^2 - (x^2 + y^2)} \right) = \Delta_{01} - R_1 \left[1 - \sqrt{1 - (x^2 + y^2)/R_1^2} \right]$$

同理有:

$$\Delta_2(x, y) = \Delta_{02} - \left(-R_2 - \sqrt{R_2^2 - (x^2 + y^2)} \right) = \Delta_{02} + R_2 \left[1 - \sqrt{1 - (x^2 + y^2)/R_2^2} \right]$$

注意 $\sqrt{R_2^2} = -R_2$ 。将表达式 $\Delta_1(x, y)$ 和 $\Delta_2(x, y)$ 代入 $\Delta(x, y)$ 式得:

$$\Delta(x, y) = \Delta_0 - R_1 \left[1 - \sqrt{1 - (x^2 + y^2)/R_1^2} \right] + R_2 \left[1 - \sqrt{1 - (x^2 + y^2)/R_2^2} \right]$$

式中: $\Delta_0 = \Delta_{01} + \Delta_{02}$ 。考虑近轴近似条件, 认为透镜中心区域 (x, y) 的范围远远小于两球面的半径, 将 $\Delta(x, y)$ 按照小量 $(x^2 + y^2)/R_1^2$ 和 $(x^2 + y^2)/R_2^2$ 进行泰勒级数展开, 忽略小量的高阶项, 得

$$\sqrt{1 - (x^2 + y^2)/R_1^2} \approx 1 - \frac{x^2 + y^2}{2R_1^2}$$

$$\sqrt{1 - (x^2 + y^2)/R_2^2} \approx 1 - \frac{x^2 + y^2}{2R_2^2}$$

这种近似的实质是利用抛物面来近似透镜近轴区域的球面。将之代入 $\Delta(x, y)$ 中,

$$\begin{aligned}\Delta(x, y) &= \Delta_0 - \frac{x^2 + y^2}{2R_1^2} + \frac{x^2 + y^2}{2R_2^2} \\ &= \Delta_0 - \frac{x^2 + y^2}{2} \left(\frac{1}{R_1} - \frac{1}{R_2} \right).\end{aligned}$$

2. 透镜的复振幅透过率

将透镜的厚度函数代入分振幅透过率函数 $t(x, y)$ 式, 可得近轴近似下光波在通过薄透镜时发生的相位延迟

$$\begin{aligned}t(x, y) &= e^{ik\Delta_0} e^{ik(n-1) \left[\Delta_0 - \frac{x^2 + y^2}{2} \left(\frac{1}{R_1} - \frac{1}{R_2} \right) \right]} \\ &= e^{ikn\Delta_0} e^{-ik(n-1) \frac{x^2 + y^2}{2} \left(\frac{1}{R_1} - \frac{1}{R_2} \right)}\end{aligned}$$

根据光学中的透镜焦距的定义,

$$\frac{1}{f} = (n-1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right)$$

则公式改写为

$$t_l(x, y) = \exp[ikn\Delta_0] \cdot \exp\left[-ik \frac{x^2 + y^2}{2f}\right].$$

上式就是透镜对入射光场的相位调制作用。其中第一项 $\exp[ikn\Delta_0]$ 表示透镜对入射光场的常数相位延迟, 并不会影响相位的空间分布, 可在讨论中略去不计。第二项的相位因子表明光波经过透镜的相位延迟与该点到透镜中心的距离的平方成正比, 而与透镜的焦距成反比, 出射光场体现处与球面波相类似的特点。

需要指明的是, 虽然以上结论是根据双凸透镜推得的, 但是只要根据符号法则正确地确定焦距的正负, 就可以推广到各种类型的透镜系统。对于焦距 $f > 0$ 的透镜, 如双凸透镜, 平凸透镜, 正弯月透镜等, 称之为正透镜; 反之, 对于焦距 $f < 0$ 的透镜, 如双凹透镜, 平凹透镜, 负弯月透镜等, 称之为负透镜。

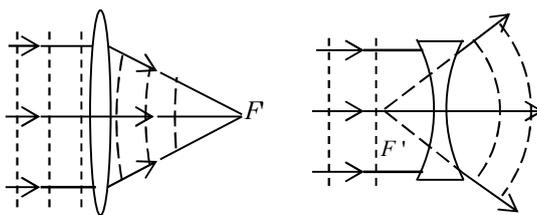
下面讨论平面波入射时出射光场的性质。假定单位振幅的平面波入射到透镜, 入射平面上的复振幅分布为

$$U_i(x, y) = 1.$$

略去常数相位因子后得到输出平面上光场的复振幅分布为

$$U'_i(x, y) = \exp\left[-ik \frac{x^2 + y^2}{2f}\right].$$

显然在近轴近似下, 得到球面波的表达式。对于正透镜, $f > 0$, 得到一个向透镜之后的焦点处会聚的球面波; 若 $f < 0$, 则得到从透镜前面焦点发散的球面波, 参见图。由此可见, 透镜具有二次型的变换因子, 在光路中起位相变换作用, 对入射波前施加相位调制。



如果考虑透镜本身的尺寸，可引入光瞳函数表示透镜的分振幅透过率

$$t_l(x, y) = \exp\left[-i \frac{k}{2f} (x^2 + y^2)\right] \times P(x, y)$$

其中 $P(x, y)$ 表示透镜的有限孔径

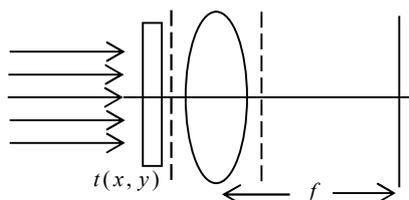
$$P(x, y) = \begin{cases} 1, & \text{in the pupil} \\ 0, & \text{out of the pupil} \end{cases}$$

-----*-----

§ 3.2 透镜的傅立叶变换性质

透镜不仅具有位相变换的作用，更重要的是还具有傅里叶变换的性质。一定方向的平面波经过透镜之后，能够会聚于后焦平面的一点，表明透镜后焦平面上的光场复振幅和入射波前的角谱之间存在一定的关系。透镜的这一作用的根本原因在于透镜对入射波前施加的位相调制能力。本节以衍射理论为基础，详细讨论透镜在三种情形下的傅里叶变换性质。

一、物体紧靠透镜放置



如图所示，物体紧靠透镜放置，采用振幅为 A 的单色平面波照射。假设物体的透过率函数为 $t(x, y)$ ，则透镜前表面的入射光场可表示为

$$U_i(x, y) = At(x, y)。$$

忽略透镜有限孔径的限制，它的复振幅透过率函数为

$$t_l(x, y) = \exp[-i \frac{k}{2f}(x^2 + y^2)]，$$

公式中已经略去常数相位延迟，则透镜出射平面的光场复振幅分布为

$$U'_l(x, y) = U_i(x, y) \cdot t_l(x, y) = At(x, y) \exp[-i \frac{k}{2f}(x^2 + y^2)]，$$

光波从透镜传播 f 距离，到达后焦平面上所产生的场分布可根据菲涅尔衍射公式进行计算

$$U_f(x_f, y_f) = \frac{1}{i\lambda f} \exp[i \frac{k}{2f}(x_f^2 + y_f^2)] \cdot F \left\{ U'_l(x, y) \exp[i \frac{k}{2f}(x^2 + y^2)] \right\}$$

式中已经舍去常数相位因子，且 $f_x = \frac{x}{\lambda f}$, $f_y = \frac{y}{\lambda f}$ 。综合以上各式可得

$$\begin{aligned} U_f(x_f, y_f) &= \frac{A}{i\lambda f} \exp[i \frac{k}{2f}(x_f^2 + x_f^2)] \times F \{t(x, y)\} \\ &= \frac{A}{i\lambda f} \exp[i \frac{k}{2f}(x_f^2 + x_f^2)] \cdot T(\frac{x_f}{\lambda f}, \frac{y_f}{\lambda f}) \end{aligned}$$

其中

$$T(\frac{x_f}{\lambda f}, \frac{y_f}{\lambda f}) = F \{t(x, y)\}。$$

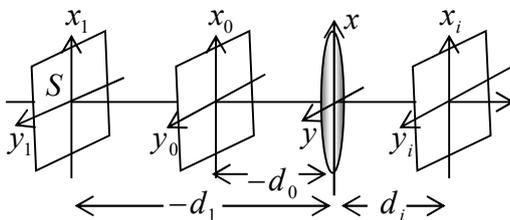
我们得到：透镜后焦平面上光场的分布正比于物体的傅里叶变换。或者说，后焦平面上点 (x_f, y_f) 的振幅和相位正比于物体频谱所包含的频率分量 $f_x = \frac{x}{\lambda f}$, $f_y = \frac{y}{\lambda f}$ 的振幅和相位。

事实上，这种傅里叶变换关系并不是准确的。由于变换式前存在位相因子 $\exp[i \frac{k}{2f}(x_f^2 + x_f^2)]$ ，后焦平面上的相位分布与物体频谱的位相分布并不相同。通常记录的是观察平面上的强度分布，相位变化并不会对强度分布造成影响，所以强度分布

$$I_f(x_f, y_f) = |U_f(x_f, y_f)|^2 = \left(\frac{A}{\lambda f} \right)^2 \cdot \left| T\left(\frac{x_f}{\lambda f}, \frac{y_f}{\lambda f} \right) \right|^2$$

与物体的功率谱成正比。

二、物体位于透镜前



如图所示，点光源 S 位于 (x_1, y_1) 平面的原点发光，照亮位于 (x_0, y_0) 平面上的物体，观察平面为 (x_i, y_i) 。其中光源到透镜的距离为 $-d_1$ ，物到透镜的距离为 $-d_0$ ，透镜到观察平面的距离为 d_i 。

在物面上的复振幅

$$U(x_0, y_0) = t(x_0, y_0) a_0 \exp[ikr] = t(x_0, y_0) a_0 \exp\left[ik \frac{x_0^2 + y_0^2}{2(d_1 - d_0)}\right]$$

在透镜前 (x 面)，(光波按菲涅耳衍射)

$$U_1(x, y) = \frac{1}{i\lambda d_0} \iint_{\Sigma_0} U(x_0, y_0) \cdot \exp\left[ik \frac{(x-x_0)^2 + (y-y_0)^2}{2d_0}\right] dx_0 dy_0$$

在透镜后 (x 面后)，光场的复振幅为：

$$U_2(x, y) = U_1(x, y) P(x, y) \exp\left[-ik \frac{x^2 + y^2}{2f}\right]$$

在象面上 (x_i, y_i 面)，光波又一次菲涅耳衍射

$$\begin{aligned} U(x_i, y_i) &= \frac{1}{i\lambda d_i} \iint_{\Sigma_p} dx dy \cdot U_2(x, y) \cdot \exp\left[ik \frac{(x_i - x)^2 + (y_i - y)^2}{2d_i}\right] \\ &= \frac{-a_0}{\lambda^2 d_0 d_i} \iint_{\Sigma_0} dx_0 dy_0 \cdot t(x_0, y_0) \times \iint_{\Sigma_p} dx dy \cdot \exp\left[i \frac{k}{2} (\Delta X + \Delta Y)\right] \end{aligned}$$

式中：

$$\begin{aligned} \Delta X &= \frac{x_0^2}{d_1 - d_0} + \frac{(x - x_0)^2}{d_0} - \frac{x^2}{f} + \frac{(x_i - x)^2}{d_i} \\ &= x_0^2 \left(\frac{1}{d_1 - d_0} + \frac{1}{d_0} \right) + x^2 \left(\frac{1}{d_0} + \frac{1}{d_i} - \frac{1}{f} \right) + x_i^2 \frac{1}{d_i} - 2x_0 x \frac{1}{d_0} - 2x_i x \frac{1}{d_i} \\ &= x_0^2 \frac{d_1}{(d_1 - d_0)d_0} + x^2 \frac{d_i f - d_i d_0 + d_0 f}{d_0 f d_i} + x_i^2 \frac{1}{d_i} - 2x_0 x \frac{1}{d_0} - 2x_i x \frac{1}{d_i} \\ &= x_0^2 \frac{f d_i}{d_0 [d_i (f - d_0) + f d_0]} + x^2 \frac{d_i (f - d_0) + f d_0}{d_0 f d_i} + \frac{x_i^2}{d_i} - \frac{2x_0 x}{d_0} - \frac{2x_i x}{d_i}, \end{aligned}$$

注意到 $\frac{1}{d_1} + \frac{1}{d_i} = \frac{1}{f}$ ， $d_1 = \frac{f d_i}{d_i - f}$ 所以

$$\frac{d_i}{(d_1 - d_0)d_0} = \frac{fd_i}{(d_i - f)\left(\frac{fd_i}{d_i - f} - d_0\right)d_0} = \frac{fd_i}{[fd_i - (d_i - f)d_0]d_0}$$

$$\Delta X = \left\{ x_0 \left[\frac{fd_i/d_0}{[d_i(f - d_0) + fd_0]} \right]^{\frac{1}{2}} - x \left[\frac{d_i(f - d_0) + fd_0}{d_0 fd_i} \right]^{\frac{1}{2}} + x_i \left[\frac{fd_0/d_i}{[d_i(f - d_0) + fd_0]} \right]^{\frac{1}{2}} \right\}^2$$

$$+ x_i^2 \frac{f - d_0}{d_i(f - d_0) + fd_0} - 2x_0 x_i \frac{f}{d_i(f - d_0) + fd_0}$$

利用代数公式：

$$(a - b + c)^2 = a^2 + b^2 + c^2 - 2ab - 2bc + 2ca$$

将大括号各项展开，整理即可得回上一式（反证法验证）。

令 $\alpha = d_i(f - d_0) + fd_0$ ， ΔX 式简写为：

$$\Delta X = \left\{ x_0 \sqrt{\frac{fd_i}{d_0 \alpha}} - x \sqrt{\frac{\alpha}{d_0 fd_i}} + x_i \sqrt{\frac{fd_0}{d_i \alpha}} \right\}^2 + x_i^2 \frac{f - d_0}{\alpha} - 2x_0 x_i \frac{f}{\alpha}$$

$$= x_0^2 \frac{fd_i}{d_0 \alpha} + x^2 \frac{\alpha}{d_0 fd_i} + x_i^2 \frac{f - d_0}{d_i \alpha} - 2x_0 \left[\frac{fd_i}{d_0 \alpha} \right]^{\frac{1}{2}} \cdot x \left[\frac{d}{d_0 fd_i} \right]^{\frac{1}{2}}$$

$$- 2x \left[\frac{\alpha}{d_0 fd_i} \right]^{\frac{1}{2}} \cdot x_i \left[\frac{fd_0}{d_i \alpha} \right]^{\frac{1}{2}} + 2x_i \left[\frac{fd_0}{d_i \alpha} \right]^{\frac{1}{2}} \cdot x_0 \left[\frac{fd_i}{d_0 \alpha} \right]^{\frac{1}{2}} + x_i^2 \frac{f - d_0}{\alpha} - 2x_0 x \frac{f}{\alpha}$$

$$= x_0^2 \frac{fd_i}{d_0 \alpha} + x^2 \frac{\alpha}{d_0 fd_i} + x_i^2 \left(\frac{fd_0}{d_i \alpha} + \frac{f - d_0}{\alpha} \right) - 2x_0 x \frac{1}{d_0} - 2xx_i \frac{1}{d_i}$$

$$= x_0^2 \frac{fd_i}{d_0 \alpha} + x^2 \frac{\alpha}{d_0 fd_i} + x_i^2 \frac{1}{d_i} - 2x_0 x \frac{1}{d_0} - 2xx_i \frac{1}{d_i}$$

同理：

$$\Delta Y = \left\{ y_0 \left[\frac{fd_i/d_0}{[d_i(f - d_0) + fd_0]} \right]^{\frac{1}{2}} - y \left[\frac{d_i(f - d_0) + fd_0}{d_0 fd_i} \right]^{\frac{1}{2}} + y_i \left[\frac{fd_0/d_i}{[d_i(f - d_0) + fd_0]} \right]^{\frac{1}{2}} \right\}^2$$

$$+ y_i^2 \frac{f - d_0}{d_i(f - d_0) + fd_0} - 2y_0 y_i \frac{f}{d_i(f - d_0) + fd_0}$$

$$= \left\{ y_0 \sqrt{\frac{fd_i}{d_0 \alpha}} - y \sqrt{\frac{\alpha}{d_0 fd_i}} + y_i \sqrt{\frac{fd_0}{d_i \alpha}} \right\}^2 + y_i^2 \frac{f - d_0}{\alpha} - 2y_0 y_i \frac{f}{\alpha}$$

根据积分公式： $\int_{-\infty}^{\infty} \exp[-\alpha x^2] dx = \sqrt{\frac{\pi}{\alpha}} = c$

可知：

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i\frac{k}{2}(x^2 + y^2)} dXdY = \int_{-\infty}^{\infty} e^{i\frac{k}{2}x^2} dx \int_{-\infty}^{\infty} e^{i\frac{k}{2}y^2} dy = \left[\int_{-\infty}^{\infty} e^{i\frac{k}{2}\theta^2} d\theta \right]^2 = c \quad (\text{常量})$$

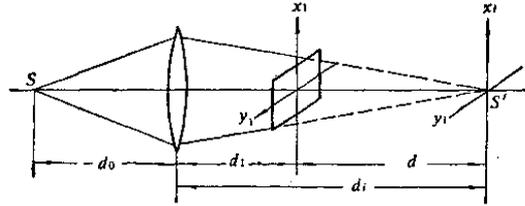
经过上述处理， $U(x_i, y_i)$ 式可写为：

$$\begin{aligned}
 U(x_i, y_i) &= c \cdot e^{i \frac{k(f-d_0)(x_i^2+y_i^2)}{2d_i(f-d_0)+fd_0}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t(x_0, y_0) e^{-ik \frac{f(x_0 x_0 + y_0 y_0)}{d_i(f-d_0)+fd_0}} dx_0 dy_0 \\
 &= c \cdot e^{i \frac{k(f-d_0)(x_i^2+y_i^2)}{2\alpha}} \times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t(x_0, y_0) e^{-ik \frac{f(x_0 x_0 + y_0 y_0)}{\alpha}} dx_0 dy_0
 \end{aligned}$$

注意到

$$\begin{aligned}
 \frac{f}{\alpha} &= \frac{f}{d_i(f-d_0)+fd_0} = \frac{f(f-d_0)}{\left(\frac{fd_1}{d_1-f}\right) + fd_0} = \frac{f}{\frac{f^2 d_1 - fd_1 d_0 + fd_0 d - f^2 d_0}{d_1 - f}} \\
 &= \frac{d_1 - f}{f(d_1 - d_i)} = \frac{f - d_1}{(d_0 - d_1)f} = \frac{1}{\sigma}
 \end{aligned}$$

三、物体位于透镜后



光学装置如图所示。我们通过以下五个步骤来计算观察平面上光场的复振幅分布。

1. 透镜前: $U_1(x, y) = a_0 \exp\left[ik \frac{x^2 + y^2}{2d'}\right]$

2. 透镜后: $U_2(x, y) = a_0 \exp\left[ik \frac{x^2 + y^2}{2d'} + ik \frac{x^2 + y^2}{2f}\right]$

3. 物前 (菲涅耳衍射)

$$O_1(x_0, y_0) = \frac{a_0}{i\lambda d_0} \iint_{\Sigma_p} dx dy \cdot \exp\left[ik \frac{x^2 + y^2}{2d'} - ik \frac{x^2 + y^2}{2f} - ik \frac{(x_0 - x)^2 + (y_0 - y)^2}{2d_0}\right]$$

4. 物后 $O_2(x_0, y_0) = O_1(x_0, y_0) \cdot \tilde{t}(x_0, y_0)$

5. 象面上 (又一次菲涅耳衍射)

$$\begin{aligned}
 U(x_i, y_i) &= \frac{1}{i\lambda(d_i - d_0)} \iint_{\Sigma} O_1(x_0, y_0) t(x_0, y_0) \\
 &\quad \times \exp\left[ik \frac{(x_i - x_0)^2 + (y_i - y_0)^2}{2(d_i - d_0)}\right] dx_0 dy_0
 \end{aligned}$$

全代入得:

$$\begin{aligned}
 U(x_i, y_i) &= \frac{-a_0}{\lambda^2 d_0 (d_i - d_0)} \iiint_{\Sigma} dx dy dx_0 dy_0 \\
 &\quad \times t(x_0, y_0) \exp\left[i \frac{k}{2} (\Delta X + \Delta Y)\right]
 \end{aligned}$$

式中

$$\begin{aligned}\Delta X &= \frac{x^2}{d'} - \frac{x^2}{f} + \frac{(x_0 - x)^2}{d_0} + \frac{(x_i - x_0)^2}{d_i - d_0} \\ &= \left\{ x \left(\frac{d_i - d_0}{d_0 d_i} \right)^{\frac{1}{2}} - x_0 \left[\frac{d_i}{d_0 (d_i - d_0)} \right]^{\frac{1}{2}} \right\}^2 + x_i^2 \frac{1}{d_i - d_0} - 2x_i x_0 \frac{1}{d_i - d_0}\end{aligned}$$

同理:

$$\Delta Y = \left\{ y \sqrt{\frac{d_i - d_0}{d_0 d_i}} - y_0 \sqrt{\frac{d_i}{d_0 (d_i - d_0)}} \right\}^2 + y_i^2 \frac{1}{d_i - d_0} - 2y_i y_0 \frac{1}{d_i - d_0}$$

代入得 (方法同上)

$$U(x_i, y_i) = C \cdot \exp \left[ik \frac{x_i^2 + y_i^2}{2(d_i - d_0)} \right] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t(x_0, y_0) \exp \left[-ik \frac{x_i x_0 + y_i y_0}{d_i - d_0} \right] dx_0 dy_0$$

考察上两式 $U(x_i, y_i)$ 可知, 不管物体置于透镜前或后, 只要观察面是光源共轭象面, 其衍射都属于夫琅和费型衍射, 即衍射物光场 (象面光场) 与衍射物之间存在傅里叶变换关系。

下面讨论一些特殊情况:

1) 物体置于透镜前焦面上, 有: $d_0 = f$ 由前一式得:

$$\begin{aligned}U(x_i, y_i) &= C \cdot \exp \left[i \frac{k (f - d_0)(x_i^2 + y_i^2)}{2 d_i (d_i - d_0) + f d_0} \right] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t(x_0, y_0) \\ &\quad \times \exp \left[-ik \frac{f(x_i x_0 + y_i y_0)}{d_i (d_i - d_0) + f d_0} \right] dx_0 dy_0 \\ &= C \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t(x_0, y_0) \exp \left[-ik \frac{(x_i x_0 + y_i y_0)}{f} \right] dx_0 dy_0 \\ &= C \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t(x_0, y_0) \exp \left[-i2\pi (f_x x_0 + f_y y_0) \right] dx_0 dy_0\end{aligned}$$

式中取 $f_x = \frac{x_i}{\lambda f}$, $f_y = \frac{y_i}{\lambda f}$, 在观察平面得到物平面的准确的傅里叶变换。

2) 物体紧靠透镜, 有: $d_0 = 0$ 由上二式得:

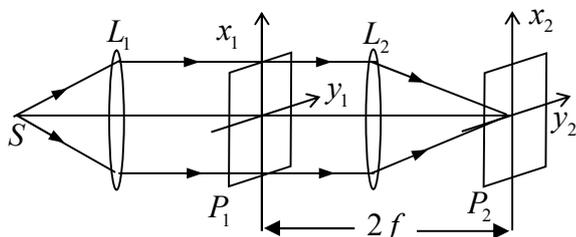
$$U(x_i, y_i) = C \cdot \exp \left[ik \frac{x_i^2 + y_i^2}{2d_i} \right] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t(x_0, y_0) \times \exp \left[-ik \frac{x_i x_0 + y_i y_0}{d_i} \right] dx_0 dy_0$$

若用平行光照明,

$$\begin{aligned}U(x_i, y_i) &= C \cdot \exp \left[ik \frac{x_i^2 + y_i^2}{2f} \right] \times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t(x_0, y_0) \cdot \exp \left[-ik \frac{(x_i x_0 + y_i y_0)}{f} \right] dx_0 dy_0 \\ &= C \cdot \exp \left[ik \frac{x_i^2 + y_i^2}{2f} \right] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t(x_0, y_0) \times \exp \left[-i2\pi (f_x x_0 + f_y y_0) \right] dx_0 dy_0\end{aligned}$$

§ 3.3 光学频谱分析

一、光学频谱分析的原理



光学频谱分析是利用透镜的傅里叶变换性质来产生物体的空间频谱，通过对它进行测量、分析物体的空间结构。频谱分析的光学系统如图所示， S 为相干点光源， L_1 为准直透镜， L_2 为傅里叶变换透镜。经准直的平面波照明位于 L_2 的前焦面的透明片，其复振幅透过率为 $t(x_1, y_1)$ 。在 L_2 的后焦面上，输出光场分布正比于物体的空间频谱 $T(f_x, f_y)$ ，即

$$U(f_x, f_y) = CF \{t(x_1, y_1)\} = CT(f_x, f_y)$$

或者

$$U(x_2, y_2) = CT\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right),$$

式中 C 是常量。强度记录得到物体的功率谱为

$$I(x_2, y_2) = C^2 \left| T\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \right|^2.$$

二、频谱分析的应用

用光学方法实现傅里叶变换的光学装置简单，处理信息容量大。

可以通过对物体的夫朗禾费衍射图样的测量来确定物体的形状尺寸，尤其对于尺寸较小的物体，直接测量常有困难，需要高精度的光学系统把它放大后测量。然而物体越小，其频谱展开越宽，频谱测量更为容易。

光学频谱分析系统可以用来对悬浮颗粒、粉尘作尺寸分析。粒子尺寸越小，频谱越扩展。由于傅里叶变换的位移性质，粒子在测量期间移动，不会影响衍射图样的位置和强度分布，为探测提供了方便。

光学频谱分析还适用于图像分析。特别适用于大量数据的快速、非精确测量。

第四章 光学成像系统的频率响应

光学成像系统是用来传递物的结构、灰度和色彩等信息传递系统。从物面到像面的质量完全取决于光学系统的传递特性。其传递能力，可以用来评价成像质量的好坏，通常把这称为像质评价。检验光学系统的成像质量，传统的方法通常是采用鉴别率板法和星点检验法。用鉴别率板法评价像质，简便易行，并能评价一个系统分辨率景物细微结构的能力，但不能对在可辨范围内的像质好坏作全面的评价，例如往往有这样的情况，鉴别率相同的物镜，粗细线条像的明晰程度可能大不一样。此外，鉴别率的等级完全由检验者主观判定，人为因素较大。对于较高的像质评价要求，可以用星点检验法。所谓星点检验，就是观测点源通过成像系统时所得像斑的形状。这个像斑就是成像系统的脉冲响应。当成像系统没有像差（或像差很小）时，像斑呈爱里圆；离焦或像差较大时，光强景往外分散或像斑不规则。星点检验法可以保证较高的成像质量，缺点是仍属主观检验，并且没有数据说明，只能作抽象的比较。

以上两种方法都是在空域中检验像质。随着空间频谱分析方法和线性系统理论用于光学系统成像分析获得成功，相应地产生了光学传递函数理论，从而使像质评价方法有了很大的改进。我们知道，光学系统是线性系统，而且在一定条件下还是线性空不变系统。因而可以用线性系统理论来研究它的性能。把输入信息分解成各种空间频率分量，然后考察这些空间频率分量在通过系统的传递过程中，丢失、衰减、相位移动等变化，也就是研究系统的空间频率传递特性即传递函数。这显然是一种全面评价光学系统成像质量的方法。传递函数可由光学系统的设计数据计算得出。虽然计算传递函数的步骤比较麻烦，检验传递函数的仪器也比较复杂。但是随着计算机和高精度光电探测技术的发展，光学传递函数的计算和测量方法日趋完善，并已实用化，成为像质评价的依据。这是现代光学中最重要的成就之一。

§ 4.1 相干照明衍射受限系统的点扩散函数

任何平面物分布看做是无数小面元的给合，而每个小面元都可看做一个加权的 δ 函数。对于一个透镜或一个成像系统，如果能清楚地了解物平面上任一小面元的光振动通过成像系统后，在像平面上所造成的光振动分布情况，通过线性叠加，原则上就能够求得任何物面光场分布通过系统后的像面光场分布，进而求得像面强度分布。这就是相干照射下的成像过程，所以，这里问题的关键是求出任意小面元的光振动所对应的像场分布。当该面元的光振动为单位脉冲即 δ 函数时，这个像场分布函数叫做点扩散函数或脉冲响应。通常用 $h(x_0, y_0; x_i, y_i)$ 表示，它表示物平面上 (x_0, y_0) 点的单位脉冲通过成像系统后在像平面上 (x_i, y_i) 点产生的光场分布。一般说来，它既然是 (x_0, y_0) 函数，也是 (x_i, y_i) 的函数。这样，脉冲函数可以表示为：

$$h(x_0, y_0; x_i, y_i) = F\{\delta(x - x_0, y - y_0)\} \quad (4.1.1)$$

对于任意的物函数 $f(x_0, y_0)$ ，可以把它看成是由物平面上许多面元组成，每个面元具有相应的脉冲响应，由于成像系统是线性系统，当用平面单色光照射时，其像平面上光场复振幅分布 $g(x, y)$ 可以用叠加积分表示成：

$$g(x_i, y_i) = \iint_{-\infty}^{\infty} f(x_0, y_0) h(x_0, y_0; x_i, y_i) dx_0 dy_0 \quad (4.1.2)$$

因此，只要能够确定成像系统的脉冲响应函数 $h(x_0, y_0; x_i, y_i)$ ，就能完备地描写该成像系统的性质。

4.1.1 透镜的点扩散函数

如图 4.1.1 所示，在相干照射下，一个消像差的正薄透镜对透明物成像的情况。物体放在透镜前距离为 d_0 的输入平面 $x_0 y_0$ 平面上，在透镜后为 d_i 的共轭面上 $x_i y_i$ 上观测成像情况。

假定紧靠物体后的复振幅分布为 $U(x'_0, y'_0)$ ，在点 (x'_0, y'_0) 处发出的单位脉冲为

$\delta(x_0 - x'_0, y_0 - y'_0)$ ，沿光波传播方向，逐面计算三个特定平面上的场分布：紧靠透镜前后

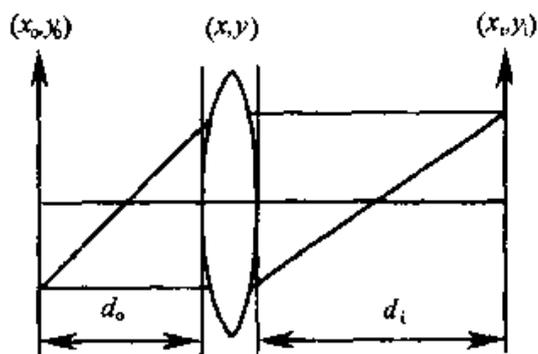


图 4.1.1 透镜点扩散函数简图

的两个平面上的场分布 dU_1 和 dU'_1 ，观测平面上的场分布 h 。这样就可以最终导出一个点源的输入输出关系。

$$\begin{aligned} dU_1(x'_0, y'_0; x, y) &= \frac{e^{ikd_0}}{i\lambda d_0} \iint_{-\infty}^{\infty} \delta(x_0 - x'_0, y_0 - y'_0) e^{\frac{ik}{2d_0}[(x-x_0)^2 + (y-y_0)^2]} dx_0 dy_0 \\ &= \frac{e^{ikd_0}}{i\lambda d_0} e^{\frac{ik}{2d_0}[(x-x'_0)^2 + (y-y'_0)^2]} \end{aligned} \quad (4.1.3)$$

考虑孔径效应，上述复振幅的光波通过孔径函数为 $P(x, y)$ ，焦距为 f 的透镜后，复振幅 $dU'(x_0, y_0; x, y)$ 为：

$$dU_1(x_0, y_0; x, y) = P(x, y) e^{-\frac{ik}{2f}(x^2 + y^2)} \times dU_1(x_0, y_0; x, y) \quad (4.1.4)$$

由于式(4.1.3)式中的 (x'_0, y'_0) 是任意的，上式可省去“撇”号。

由于透镜后表面到观察面，光场的传播满足菲涅耳衍射，于是物平面上的单位脉冲在观测面上引起的复振幅分布即点扩散函数可写作：

$$\begin{aligned} h(x_0, y_0; x_i, y_i) &= \frac{e^{ikd_i}}{i\lambda d_i} \iint_{-\infty}^{\infty} dU_1(x_0, y_0; x, y) e^{\frac{ik}{2d_i}[(x_i-x)^2 + (y_i-y)^2]} dx dy \\ &= \frac{e^{ik(d_0+d_i)}}{\lambda^2 d_0 d_i} \iint_{-\infty}^{\infty} dx dy P(x, y) e^{-\frac{ik}{2f}(x^2 + y^2)} e^{\frac{ik}{2d_0}[(x-x_0)^2 + (y-y_0)^2]} e^{\frac{ik}{2d_i}[(x_i-x)^2 + (y_i-y)^2]} \end{aligned} \quad (4.1.5)$$

由于物像平面的共轭关系满足高斯公式，故有：

$$\frac{1}{f} = \frac{1}{d_i} + \frac{1}{d_0} \quad (4.1.6)$$

这样，(4.1.5)式忽略常数相位因子并简化后得到：

$$h(x_0, y_0; x_i, y_i) = \frac{1}{\lambda^2 d_0 d_i} e^{\frac{ik}{2d_0}(x_0^2 + y_0^2)} e^{\frac{ik}{2d_i}(x_i^2 + y_i^2)} \iint_{-\infty}^{\infty} dx dy \cdot P(x, y) e^{-ik \left[\left(\frac{x_i + x_0}{d_i + d_0} \right) x + \left(\frac{y_i + y_0}{d_i + d_0} \right) y \right]} \quad (4.1.7)$$

上式所表示的点扩散函数依然显示的复杂，积分号前的 $e^{\frac{ik}{2d_i}(x_i^2 + y_i^2)}$ 这一相位因子，不影响最终探测的强度分布，可以略去。但由于在求物面上各点对像面光场的贡献时， $e^{\frac{ik}{2d_0}(x_0^2 + y_0^2)}$ 这一相位因子是要参与的，因而不能简单略去。

当透镜的孔径比较大时，物面上每一物点产生的脉冲响应是一个很小的像斑，那么能够对于像面上 (x_i, y_i) 点光场产生有意义贡献的，必定是物面上以几何成像所对应的以物点为中心的微小区域。在这个区域内，可近似地认为 x_0, y_0 不变，这样，其值就与 (x_i, y_i) 点的共轭物坐标：

$$x_0 = \frac{x_i}{M}, \quad y_0 = \frac{y_i}{M} \quad (4.1.8)$$

式中： $M = -\frac{d_i}{d_0}$ ，是成像透镜的横向放大率。这样有：

$$e^{\frac{ik}{2d_0}(x_0^2+y_0^2)} \approx e^{\frac{ik}{2d_0M^2}(x_i^2+y_i^2)} \quad (4.1.9)$$

通过近似后的相位因子不再依赖 (x_0, y_0) ，因此不会影响 x_i, y_i 平面上的强度分布，于是也可以略去。

这样，应用 $M = -\frac{d_i}{d_0}$ 和 $k = \frac{2\pi}{\lambda}$ ，(4.1.7) 式简化为：

$$h(x_0, y_0; x_i, y_i) = \frac{1}{\lambda^2 d_0 d_i} \iint_{-\infty}^{\infty} dx dy \cdot P(x, y) e^{-\frac{i2\pi}{\lambda d_i} [(x_i - Mx_0)x + (y_i - My_0)y]} \quad (4.1.10)$$

对物平面坐标 (x_0, y_0) 作坐标变换，令：

$$\tilde{x}_0 = Mx_0, \quad \tilde{y}_0 = My_0 \quad (4.1.11)$$

于是有：

$$h(\tilde{x}_0 / M, \tilde{y}_0 / M; x_i, y_i) = \frac{1}{\lambda^2 d_0 d_i} \iint_{-\infty}^{\infty} P(x, y) e^{-\frac{i2\pi}{\lambda d_i} [(x_i - \tilde{x}_0)x + (y_i - \tilde{y}_0)y]} dx dy \quad (4.1.12)$$

于是， $h(x_0, y_0; x_i, y_i)$ 可以写成 $h(x_i - \tilde{x}_0, y_i - \tilde{y}_0)$ 的形式，即，

$$h(x_i - \tilde{x}_0, y_i - \tilde{y}_0) = \frac{1}{\lambda^2 d_0 d_i} \iint_{-\infty}^{\infty} P(x, y) e^{-\frac{i2\pi}{\lambda d_i} [(x_i - \tilde{x}_0)x + (y_i - \tilde{y}_0)y]} dx dy \quad (4.1.13)$$

这表时，在近轴成像条件下，上式所表征的透镜成像是空不变的。而且，透镜的脉冲响应就等于透镜孔径的夫琅和费衍射图样，其中心位于几何光学理想像点 $(\tilde{x}_0, \tilde{y}_0)$ ，应用(4.1.7)，出就是点 (x_i, y_i) 。

透镜孔径的衍射作用明显显示与否，是由孔径线度相对于波长 λ 和像距 d_i 的比例决定的，为此，对孔径平面上的坐标 (x, y) ，也就是透镜所在的平面，作如下坐标变换：

$$\tilde{x} = \frac{x}{\lambda d_i}, \quad \tilde{y} = \frac{y}{\lambda d_i} \quad (4.1.14)$$

代入(4.1.13)式，可得：

$$h(x_i - \tilde{x}_0, y_i - \tilde{y}_0) = |M| \iint_{-\infty}^{\infty} d\tilde{x} d\tilde{y} \cdot P(\lambda d_i \tilde{x}, \lambda d_i \tilde{y}) e^{-i2\pi [(x_i - \tilde{x}_0)\tilde{x} + (y_i - \tilde{y}_0)\tilde{y}]} \quad (4.1.15)$$

上式就是透镜的点扩散函数。如果孔径大小(光瞳)相对于 λd_i 足够大，则在 \tilde{x}, \tilde{y} 的无限大区域内，都有 $P(\lambda d_i \tilde{x}, \lambda d_i \tilde{y}) = 1$ 。这样，上式就可以进一步简化为：

$$\begin{aligned}
 h(x_i - \tilde{x}_0, y_i - \tilde{y}_0) &= |M| \left| \int \int_{-\infty}^{\infty} e^{-i2\pi[(x_i - \tilde{x}_0)\tilde{x} + (y_i - \tilde{y}_0)\tilde{y}]} d\tilde{x}d\tilde{y} \right. \\
 &= |M| \delta(x_i - \tilde{x}_0, y_i - \tilde{y}_0)
 \end{aligned} \quad (4.1.16)$$

这时，物点成像为一个像点，即理想成像。即，当不考虑光瞳有限大小时，点脉冲通过成像系统后，其响应函数仍是点脉冲，其位置在： $x_i = Mx_0, y_i = My_0$ 。这便是几何光学中点物-点像对应的情况。

§ 4.1.2 衍射受限系统的点扩散函数

所谓衍射受限系统，是指不考虑系统的几何像差，仅仅考虑系统的衍射限制。我们知道，大多数光学系统，不象我们在基础光学中所熟知的单个凸透镜，通常是由若干个透镜（正透镜或负透镜）和其他光学元件（例如棱镜、光阑等）组合成的复合系统。而且，透镜也不一定是薄的。因此，在考察光学系统对成像的影响时，必然在若干个可能对光束起限制作用有通光孔径中，找到对光束起实际限制作用的那个孔径，该孔径可能某一透镜边框，也可能是光路中某一个特定光阑，决定系统光束大小的光阑称为孔径光阑 (Aperture stop) 或称为有效光阑 (Effective stop)。由基础光学知道，孔径光阑通过它前面的光学系统（物空间）所成的像，称为系统的入射光瞳 (Entrance pupil, 记为 En. P.)，它决定关进入系统的光束的大小。孔径光阑通过它后面的光学系统（像空间）所成的像，称为出射光瞳，简称出瞳 (Exit pupil, 记为 Ex. P.)，决定着从系统出射的光束的大小。并且，入射光瞳、孔径光阑和出射光瞳三才相互共轭。当轴的物点确定后，孔径光阑、入瞳和出瞳由系统元件参数及相对位置决定。由入射光瞳限制的物方光束必定能全部通过系统，成为被出射光瞳所限制的像方光束。

根据基础光学的讨论，一个成像系统的外部性质可以由入射光瞳或出射光瞳来描述。因此，不管成像系统的详细结构如何，都可以将它归结为下列普遍模型：光波由物平面变换到像平面，可以分为 3 个过程，即光由物平面到入瞳面，再由入瞳面到出瞳面，最后由出瞳面到像平面。如图 4.1.2 所示。当光波通过成像系统时，波面受到入瞳的限制，变换到空间就成为出瞳对出射波的限制。这两种限制是等价，是同一种限制在两个空间间的反映。这一结论称为光束限制的共轭原理。在考虑光波通过光学系统的衍射效应时，只须考虑其中任何一种限制，通常是考虑出瞳对光波的衍射作用。至于光波从入瞳到出瞳的传播，由于在此过程中波面已不再受到别的限制，故此段传播可以用几何光学很好描述。有了光瞳的概念，在研究光学成像系统的性质时，可以不去涉及系统的详细结构，而把整个系统的成像看成是一个“黑箱”的作用，只需知道黑箱边端（即入瞳和出瞳平面）的物理性质，就可以知道像平面上合乎实际的像场分布。

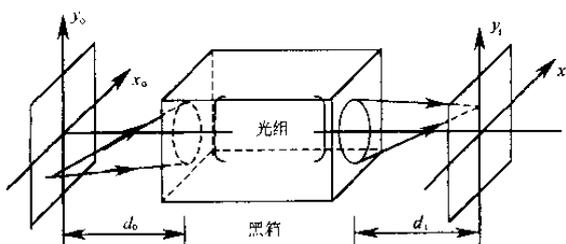


图 4.1.2 成像系统的普遍模型

为了确定系统的脉冲响应，首先需要知道这个“黑箱”对点光源发出的球面波的变换作用，即当入瞳平面上输入发散球面波时，出瞳平面透射的波场特性。对于实际光学系统，这一边端性质千差万别，但总的来说，可以分成两类，即衍射受限系统和有像差系统。

当像差很小或者系统的孔径和视场都不大，实际光学系统就可近似看做衍射受限系统。这时的边端性质变比较简单，也就是说，从物平面上任一点源，如果从该点发出的发散球面波通过成像系统后，因该系统受限制，转换成了新的理想球面波，并且在像平面上会聚成一个理想点，则称为该成像是衍射受限

的 (Diffraction-limited)。因此，衍射受限系统的作用，就是将投射到入瞳上的发散球面波变换成出瞳上的会聚球面波。至于有像差的系统，其边端条件是：点光源发出的发散球面波投射到入瞳上，在出瞳处的透射波场将明显偏离理想球面波，偏离程度由波像差决定。

阿贝尔认为衍射效应是由于有限的入瞳大小引起的，1896 年瑞利提出衍射效应来自有限大小的出瞳。由于一个光瞳只不过是另一个光瞳的几何像，这两种看法是等效的。衍射效应可以归结为入瞳或出瞳对于成像光波的限制。

由物点发出的球面波，在像方得到的将是一个被出射光瞳所限制的球面波，这个球面波是以理解像点为中心的。由于出射光瞳的限制作用，在像平面上将产生以理想像点为中心的出瞳孔径的夫琅和费衍射花样。这样，可以写出物面上以 (x_0, y_0) 点的单位脉冲通过衍射受限系统后在与物面共轭的像面上的复振幅分布，即点扩散函数为：

$$h(x_0, y_0; x_i, y_i) = K \iint_{-\infty}^{\infty} dx dy \cdot P(x, y) e^{-\frac{i2\pi}{\lambda d_i} [(x_i - Mx_0)x + (y_i - My_0)y]} \quad (4.1.17)$$

式中， K 是与 x_0, y_0 和 x_i, y_i 无关的复常数； $P(x, y)$ 是出瞳函数，也常称为光瞳函数，从前面知道，在光瞳内其值为 1，在光瞳外其值为 0。 d_i 是光瞳平面到像面的距离，已不是通常意义下的像距。需要注意的是：式 (4.1.17) 的推出，与式 (4.1.10) 一样，也是略去了关于 x_0, y_0 和 x_i, y_i 的二次相位因子，所以，与 (4.1.10) 一样，也是有条件。式 (4.1.17) 有明，若略去积分号前面的系数，脉冲响应函数的傅里叶变换。即衍射受限系统的脉冲响应是光学系统出瞳的夫琅和费衍射图样。其中心在几何光学的理想像点 (Mx_0, My_0) 处。

同样，对物平面上坐标 x_0, y_0 和光瞳平面上的坐标 x, y 做坐标变换，即：

$$\begin{aligned} \tilde{x}_0 = Mx_0 & , & \tilde{x} = \frac{x}{\lambda d_i} \\ \tilde{y}_0 = My_0 & , & \tilde{y} = \frac{y}{\lambda d_i} \end{aligned}$$

这样，便可得到：

$$h(x_i - \tilde{x}_0, y_i - \tilde{y}_0) = K \lambda^2 d_i^2 \iint_{-\infty}^{\infty} P(\lambda d_i \tilde{x}, \lambda d_i \tilde{y}) \times e^{-i2\pi [(x_i - \tilde{x}_0)\tilde{x} + (y_i - \tilde{y}_0)\tilde{y}]} d\tilde{x} d\tilde{y} \quad (4.1.18)$$

如果孔径大小(光瞳)相对于 λd_i 足够大，则在 \tilde{x}, \tilde{y} 的无限大区域内，都有 $P(\lambda d_i \tilde{x}, \lambda d_i \tilde{y}) = 1$ 。这样，上式就可以进一步简化为：

$$h(x_i - \tilde{x}_0, y_i - \tilde{y}_0) = K \lambda^2 d_i^2 \delta(x_i - \tilde{x}_0, y_i - \tilde{y}_0) \quad (4.1.19)$$

上式表明，当可以忽略光瞳的衍射时， (x_0, y_0) 点脉冲通过衍射受限系统后在像面上得到仍然是点脉冲，其位置在： $x_i = Mx_0, y_i = My_0$ ，这便是几何光学理想成像的情况。

-----*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-----

§ 4.2 相干照射下衍射受限系统的成像规律

如何求给定物分布通过衍射受限系统后,在像平面上的像分布,包括复振幅分布和光强分布。一个确定的物分布总可以很方便地分解成无数 δ 函数的线性组合,而每个 δ 函数可按(4.1.19)式求出其响应。然而,在像平面上将这些无数个脉冲响应合成的结果是和物面照射情况有关的,如果物面上某两个脉冲是相干的,则这两个脉冲在像平面上的响应便是相干叠加;如果这两个脉冲是非相干的,则这两个脉冲在像平面上的响应将是非相干叠加,即强度叠加。所以衍射受限系统的成像特性,对于相干照射和非相射是不同的。

设物的复振幅分布为 $U_0(x_0, y_0)$, 在相干照射下,物面上的各点是完全相干的。按(1.6.5),将物分布用 δ 函数表达为:

$$U_0(x_0, y_0) = \int \int_{-\infty}^{\infty} U_0(\xi, \eta) \delta(x_0 - \xi, y_0 - \eta) d\xi d\eta \quad (4.2.1)$$

物面上每一个脉冲通过系统后都形成一个复振幅分布,所有这些分布的相干叠加,便是物通过系统后所得的像的 $U_i(x_i, y_i)$, 即

$$\begin{aligned} U_i(x_i, y_i) &= F\{U_0(x_0, y_0)\} \\ &= F\left\{\int \int_{-\infty}^{\infty} U_0(\xi, \eta) \delta(x_0 - \xi, y_0 - \eta) d\xi d\eta\right\} \\ &= \int \int_{-\infty}^{\infty} U_0(\xi, \eta) F\{\delta(x_0 - \xi, y_0 - \eta)\} d\xi d\eta \quad (4.2.2) \\ &= \int \int_{-\infty}^{\infty} U_0(\xi, \eta) h(x_i - M\xi, y_i - M\eta) d\xi d\eta \\ &= \frac{1}{M^2} \int \int_{-\infty}^{\infty} U_0\left(\frac{\tilde{x}_0}{M}, \frac{\tilde{y}_0}{M}\right) h(x_i - \tilde{x}_0, y_i - \tilde{y}_0) d\tilde{x}_0 d\tilde{y}_0 \end{aligned}$$

式中的 h 为(4.1.18)式。对理解上式,称看看积分中第一项中 \tilde{x}_0, \tilde{y}_0 坐标的含义,我们知道,公式(4.1.19)式代表理想成像脉冲,代入上(4.2.2)式,所得到的像为 $U_i(x_i, y_i)$, 是理想的成像分布。用 $U_g(x_i, y_i)$ 表示,可得:

$$\begin{aligned} U_g(x_i, y_i) &= \frac{1}{M^2} \int \int_{-\infty}^{\infty} U_0\left(\frac{\tilde{x}_0}{M}, \frac{\tilde{y}_0}{M}\right) K \lambda^2 d_i^2 \times \delta(x_i - \tilde{x}_0, y_i - \tilde{y}_0) d\tilde{x}_0 d\tilde{y}_0 \\ &= \frac{K \lambda^2 d_i^2}{M^2} \int \int_{-\infty}^{\infty} U_0\left(\frac{\tilde{x}_0}{M}, \frac{\tilde{y}_0}{M}\right) \delta(x_i - \tilde{x}_0, y_i - \tilde{y}_0) d\tilde{x}_0 d\tilde{y}_0 \quad (4.2.3) \\ &= \frac{K \lambda^2 d_i^2}{M^2} U_0\left(\frac{x_i}{M}, \frac{y_i}{M}\right) \end{aligned}$$

上式表时,理想像 U_g 的分布形式与物的 U_0 的分布形式是一样的,只是在 x_i 和 y_i 方向放大了 M 倍。

$U_0(x_0, y_0)$ 与 $U_0\left(\frac{\tilde{x}_0}{M}, \frac{\tilde{y}_0}{M}\right)$ 的图形是一样的,只是由于 $\tilde{x}_0 = Mx_0, \tilde{y}_0 = My_0$, 所以图形在 \tilde{x}_0, \tilde{y}_0 坐标中的读

数比在 x_0, y_0 坐标中放大了 M 倍。所以 $U_0\left(\frac{\tilde{x}_0}{M}, \frac{\tilde{y}_0}{M}\right)$ 在 \tilde{x}_0, \tilde{y}_0 坐标中与 $U_g(x_i, y_i)$ 是一样的。因此,把

$U_0\left(\frac{\tilde{x}_0}{M}, \frac{\tilde{y}_0}{M}\right)$ 叫做 $U_0(x_0, y_0)$ 的理想像。令:

$$\tilde{h}(x_i - \tilde{x}_0, y_i - \tilde{y}_0) = \frac{1}{K \lambda^2 d_i^2} h(x_i - \tilde{x}_0, y_i - \tilde{y}_0) \quad (4.2.4)$$

§ 4.3 衍射受限系统的相干传递函数

从前面的讨论, 我们可以看到, 在相干照射下的衍射受限系统, 对复振幅的传递是线性空不变的。系统的成像特性在空域中是由点扩散函数 $\tilde{h}(x_i, y_i)$ 来表征的。由于空不变系统的变换特性在频域中描述起来更为方便。如果我们在频域中用 $\tilde{h}(x_i, y_i)$ 的频谱函数 $H(\xi, \eta)$ 来描述系统的成像特性, $H(\xi, \eta)$ 便称为相干成像系统衍射受限的相干传递函数 (Coherent transfer function, CTF)。

§ 4.3.1 相干传递函数

相干成像系统的物像关系由 (4.2.5) 中的卷积积分描述的, 即,

$$U_i(x_i, y_i) = \iint_{-\infty}^{\infty} U_g(\tilde{x}_0, \tilde{y}_0) \tilde{h}(x_i - \tilde{x}_0, y_i - \tilde{y}_0) d\tilde{x}_0 d\tilde{y}_0 \quad (4.3.1)$$

式中, $U_g(\tilde{x}_0, \tilde{y}_0)$ 是几何光学理想的复振幅分布, \tilde{h} 是系统的脉冲响应。卷积积分是把物点看做基元, 像点是物点产生的衍射图样的相干叠加。这是从空域的观点来看, 如果从频域角度来分析成像过程。我们选择复指数函数作为物的基元分布, 考查系统对各种频率成分的传递特性。定义系统的输入频谱 $G_{gc}(\xi, \eta)$ 和输出频谱 $G_{ic}(\xi, \eta)$ 分别为:

$$G_{gc}(\xi, \eta) = F\{U_g(\tilde{x}_0, \tilde{y}_0)\} \quad (4.3.2)$$

$$G_{ic}(\xi, \eta) = F\{U_i(x_i, y_i)\} \quad (4.3.3)$$

相干传递函数 CTF 为:

$$H(\xi, \eta) = F\{\tilde{h}(x_i, y_i)\} \quad (4.3.4)$$

在衍射受限系统中, \tilde{h} 由 (4.2.6) 式决定, 代入上式, 可得:

$$\begin{aligned} H(\xi, \eta) &= F\{F\{P(\lambda d_i \tilde{x}, \lambda d_i \tilde{y})\}\} \\ &= P(-\lambda d_i \xi, -\lambda d_i \eta) \end{aligned} \quad (4.3.5)$$

若将光瞳面上的坐标取反演形式, 则上式中的 P 所含的负号可以略去。实际上, 光瞳函数大多是对光轴呈中心对称的, 故舍去其中的负号不对产生实质性的影响。这样, 上式, 可以写成:

$$H(\xi, \eta) = P(\lambda d_i \xi, \lambda d_i \eta) \quad (4.3.6)$$

上式, 表明, 相干传递函数在数值上就等系统的光瞳函数。这样, 就把相干传递函数与表示系统物理属性的光瞳函数联系起来。

根据 2.4 节对光瞳函数的定义, 要以写成:

$$P(\lambda d_i \xi, \lambda d_i \eta) = \begin{cases} 1 & \text{在出瞳内} \\ 0 & \text{在出瞳外} \end{cases} \quad (4.3.7)$$

式中, 频域坐标 (ξ, η) 与其空域坐标 (x, y) 之间的关系为:

$$x = \lambda d_i \xi, \quad y = \lambda d_i \eta \quad (4.3.8)$$

由于出瞳的孔径沿 x 轴和 y 轴方向的线度是有限的, 因此, 沿 x 轴和 y 轴方向的空间频率的取值也是有限的, 其极大值定义为系统的截止频率 (Cut-off frequency), 记为 ρ_c , 并有:

$$\rho_{cx} = \frac{x_{\max}}{\lambda d_i}, \quad \rho_{cy} = \frac{y_{\max}}{\lambda d_i} \quad (4.3.9)$$

式中, x_{\max} 和 y_{\max} 分别是出瞳沿 x 轴和 y 轴的线度。由于出瞳的取值不是 1 就是 0, 故由式(4.3.6), 对于相干传递函数, 也有:

$$H(\xi, \eta) = \begin{cases} 1 & \text{在出瞳内} \\ 0 & \text{在出瞳外} \end{cases} \quad (4.3.10)$$

其取值也是 1 或 0。

假如不考虑孔径有限大小, 认为恒有 $P=1$, 则整个频谱面上都有 $H(\xi, \eta)=1$ 。这时像是物的准确复现, 没有任何信息丢失。这正是几何光学理想成像情况。

对于物分布中的某一空间频率分量 (ξ, η) , 系统能否将它传递到像面上, 取决于式(4.3.8)所确定的空域坐标值 (x, y) 是否在光瞳孔径之内, 若在光瞳内, 则此频率成分的平面波分量将毫无衰减(包括振幅和相位)地通过系统到达像面。如果在光瞳外, 则系统将完全不让这种频率成分的平面波分量通过, 在像平面上完全没有这种频率成分, 即系统对这种频率不予传递, 是截止的, 如图 4.1.1 所示。这就意味着, 对衍射受限相干成像系统, 存在一个有限通频带, 在通频带内, 系统允许每一频率分量无畸变地通过; 在通频带外, 频率响应突然变为零, 即通带以外的所有频率分量统统都被衰减掉。因此, 衍射受限相干成像系统对输入的各种频率成分的作用, 相当于一个低通滤波器。由此可见, 截止频率是检验光学成像系统质量优劣的重要参数之一。

由上面讨论可知, 为了求出相干传递函数 $H(\xi, \eta)$, 只须先求出光瞳函数 $P(x, y)$, 再把其中的 x, y 用 $\lambda d_i \xi, \lambda d_i \eta$ 替换。当系统的像差一定时, 相干传递函数直接由光瞳函数的形状、大小和位置确定。所以, 光瞳的选择, 对成像过程有重大的影响, 也是计算 $H(\xi, \eta)$ 的关键。下面以系统具有规则形状的光瞳的情形为例, 计算其相干传递函数。

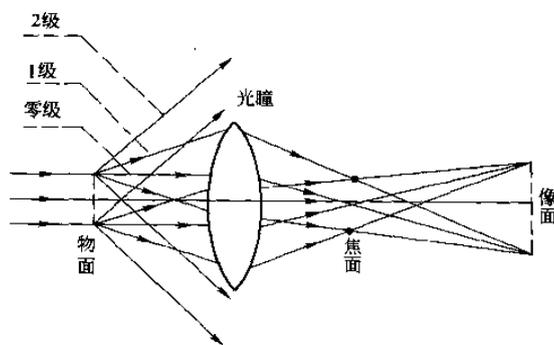


图 4.3.1 光瞳对高级衍射分量的限制

例 4.3.1 有一出射光瞳为正方形的衍射受限系统, 正方形的边长为 l , 试计算该系统的相干传递函数。

解 该系统出瞳的透射率函数 $P(x, y)$ 可以用一个二维矩形函数来描述, 如图 2.3.2(a)所示。

$$P(x, y) = \text{rect}\left(\frac{x}{l}\right)\text{rect}\left(\frac{y}{l}\right) = \begin{cases} 1 & \frac{|x|}{l}, \frac{|y|}{l} \leq \frac{1}{2} \\ 0 & \text{其他} \end{cases} \quad (4.3.11)$$

系统的相干传递函数是:

$$\begin{aligned}
 H(\xi, \eta) &= P(\lambda d_i \xi, \lambda d_i \eta) \\
 &= \text{rect}\left(\frac{\lambda d_i \xi}{l}\right) \text{rect}\left(\frac{\lambda d_i \eta}{l}\right) = \begin{cases} 1 & |\xi|, |\eta| \leq \frac{l}{2\lambda d_i} \\ 0 & \text{others} \end{cases} \quad (4.3.12)
 \end{aligned}$$

其函数图形如图 4.3.2(b)所示。显然， x 轴和 y 轴方向的空间截止频率分别为：

$$\rho_{cx} = \frac{l}{2\lambda d_i}, \quad \rho_{cy} = \frac{l}{2\lambda d_i} \quad (4.3.13)$$

系统的最大截止频率与 x 轴成 45 度然的方向上，为：

$$\rho_{c\max} = \frac{\sqrt{2}l}{2\lambda d_i} \quad (4.3.14)$$

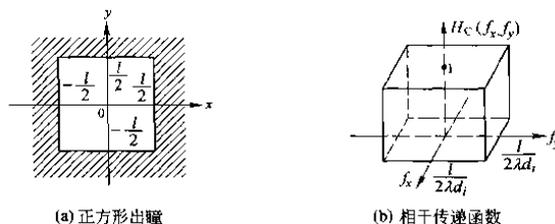


图 4.3.2 出瞳为正方形时系统的 CTF

例 4.3.2 设衍受限系统的出射光为一圆，其直径为 $D = 2r$ ，试计算该系统的相干传递函数。

解 该系统出瞳的透射率函数 $P(x, y)$ 可以用一个圆域函数来描述，如图 2.3.3(a)所示。

$$P(x, y) = \text{circ}\left(\frac{\sqrt{x^2 + y^2}}{D/2}\right) = \begin{cases} 1 & x^2 + y^2 \leq r^2 = \left(\frac{D}{2}\right)^2 \\ 0 & \text{others} \end{cases} \quad (4.3.15)$$

系统的相干传递函数是：

$$\begin{aligned}
 H(\xi, \eta) &= P(\lambda d_i \xi, \lambda d_i \eta) = \text{circ}\left(\frac{\lambda d_i \sqrt{\xi^2 + \eta^2}}{D/2}\right) \\
 &= \left(\frac{\sqrt{\xi^2 + \eta^2}}{D/(2\lambda d_i)}\right) = \begin{cases} 1 & \sqrt{\xi^2 + \eta^2} \leq \frac{D}{2\lambda d_i} \\ 0 & \text{others} \end{cases} \quad (4.3.16)
 \end{aligned}$$

其函数图形如图 4.3.3(b)所示。显然，根据出瞳的圆对称性，该系统在一切方向的空间截止频率分别为：

$$\rho_c = \sqrt{\xi^2 + \eta^2} = \frac{D}{2\lambda d_i} \quad (4.3.17)$$

如，当 $D = 1 \text{ cm}$ ， $d_i = 10 \text{ cm}$ ， $\lambda = 632.8 \text{ nm}$ ，用上式可以算得空间截止频率为 $\rho_c = 70 \text{ 线对/mm}$ 。

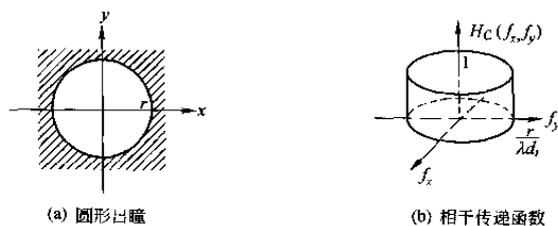


图 4.3.3 出瞳为圆形时系统的 CTF

例 4.3.3 图 4.3.4(a) 和图 4.3.4(b) 表示两个相干成像系统，各透镜的焦距为 f 。图 4.3.4(a) 中单透镜的光阑直径为 D 。为使图 4.3.4(b) 双透镜系统获得与图 4.3.4(a) 相同的截止频率，孔径光阑直径 a 应等于多大？

解 设物距为 d_0 ，像距为 d_i ，放大率为 M ，为使成实像时， M 为正，将像面坐标相对于物面坐标反演，使倒像的影响不反映在 M 上。于是 M 可表示成：

$$M = \frac{d_i}{d_0} = \frac{d_i - f}{f}$$

即有：

$$d_i = (1 + M)f$$

此系统的光瞳函数是直径为 D 的圆形孔径，其截止频率与物像空间截止频率 ρ_{0c} 的关系为：

$$\rho_c = \frac{D}{2\lambda d_i} = \frac{1}{M} \rho_{0c}$$

即有：

$$\rho_{0c} = \frac{MD}{2\lambda d_i} = \frac{MD}{2\lambda(1+M)f} \quad (4.3.18)$$

为求得当 ρ_{0c} 取最大值时的放大倍数 M ，将 ρ_{0c} 对 M 求导出并令其为零得：

$$\frac{d\rho_{0c}}{dM} = \frac{D}{2\lambda f} \frac{1}{(1+M)^2} = 0$$

因此，只有当放大倍数 M 为无穷大时，系统才有最大的空间截止频率，此截止频率为：

$$\rho_{0c \max} = \lim_{M \rightarrow \infty} \frac{D}{2\lambda f} \frac{M}{1+M} = \frac{D}{2\lambda f} \quad (4.3.19)$$

此时，物置于透镜前焦面，像在像方无穷远，在物空间的通频带为：

$$-\frac{D}{2\lambda f} < \rho < \frac{D}{2\lambda f} \quad (4.3.20)$$

这样，对单透镜系统，其截止频率为：

$$\rho_c = \frac{D}{4\lambda f} \quad (4.3.21)$$

对于图 4.3.4(b) 的双透镜经纬度，因其孔径光阑置于频谱面上，故入瞳和出瞳分别在物方和像方无限远处。又由于入瞳、孔径光阑与出瞳三者互为共轭，故对于这种放大率为 1 的系统，能通过孔径光阑的最高

空间频率，也必定能通过入瞳和出瞳。换言之，系统的截止可通过孔径光阑的尺寸来计算。

为保证图 4.3.4(b)物面上每一面元发出的低于某一空间频率的平面波，都毫无阻挡地通过光成像系统，则要法语相应的截止频率为：

$$\rho'_c = \frac{a/2}{\lambda f} = \frac{a}{2\lambda f} \quad (4.3.22)$$

当 $\rho_c = \rho'_c$ 时，可以得到：

$$a = \frac{D}{2}$$

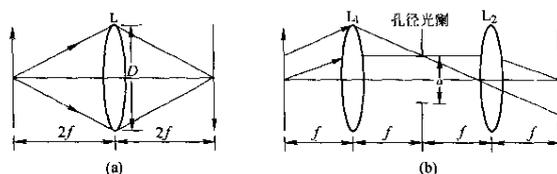


图 4.3.4 两个相干成像系统

§ 4.3.2 相干线扩散函数和边缘扩散函数

测量传递函数的方法，一是计算或测量出系统的点扩散函数，然后对它做傅里叶变换以求得传递函数。这种方法要求得到点扩散函数的精确表达式，有时难以得到时，这种方法就难以使用。另一种方法是把大量频率不同的本征函数逐入输入系统，并确定每个本征函数所受到的衰减及其相移，从而得到传递函数。这种方法较第一种方法直接，但测量数目大，有时实现起来困难。一种实用的方法是由线扩散函数确定传递函数。

1. 线扩散函数与边缘函数的概念

前面我们知道，对于相干照射成像系统，点物在像面上的响应，也就是点扩散函数是一种复振幅分布，所有点物响应的叠加即得像面上的复振幅分布，复振幅的绝对值的平方就是像面的强度分布。对于非相干照射成像系统，点物在像面上的响应即强度点扩散函数是一种强度分布，所有这些强度点扩散函数的叠加就得到像面强度分布。但无论是复振幅点扩散函数或是强度点扩散函数，就其叠加成线扩散函数的方式而言，是一样的。这样，我们可以以点物的强度响应为例讨论点扩散函数与线扩散函数的关系。

一个物点在像面上造成的强度分布，可以点扩散函数 $h(x_i, y_i)$ ， x_i, y_i 是像面上的位置坐标。校正好的镜，其轴上物点的像是圆对称的，如图 4.3.5(a)所示的像点。图上所示的是点扩散函数的轮廓形状，强度变化没有表示出来，但强度分布对像斑中心 O 呈圆对称。这样，对于从任何一方向，通过中心的一条狭缝，如图中虚线所示，从缝中看到的强度分布，以离 O 点远近表示，分布状况都是一样的。通常以沿 x_i 轴的狭缝的强度分布曲线 $h(x_i)$ 作为点扩散函数，其线形如图 4.3.5(b)所示。

一个狭缝通过光学系统后，光强分布依然是往两侧散开的，散开的情况取决于光学系统的点扩散函数。由于一根亮直线或一个亮狭缝，可以看成是由许多亮点的集合组成的，这许多沿直线排列的点源的像点的叠加就构成亮直线的光强分布。如果我们把直线像的长度方面取为 y_i 方向，那么沿 x_i 方向上的光强分布 $L(x_i)$ 就叫做线扩散函数，图 4.3.5(b)同样可以代表亮直线的成像情况，这就是实际成像的强度分布，也就是线扩散函数。点扩散函数的曲线形状和线扩散函数的曲线形状是不一样的。线扩散函数由点扩散函数叠加而所。

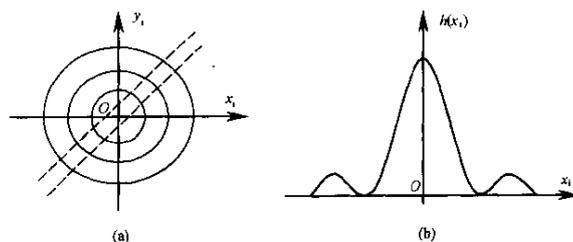


图 4.3.5 点扩散函数的分布

如果系统输入一线脉冲，设这线脉冲是平行于 y_0 轴的线光源，即：

$$U_0(x_0, y_0) = \delta(x_0) \quad (4.3.23)$$

这样，线性空不变系统的线扩散函数为：

$$\begin{aligned} L(x_i) &= \delta(x_i) \otimes h(x_i, y_i) = \iint_{-\infty}^{\infty} \delta(\xi) h(x_i - \xi, \eta) d\xi d\eta \\ &= \int_{-\infty}^{\infty} h(x_i, \eta) d\eta \end{aligned} \quad (4.3.24)$$

上式表明，线扩散函数仅依赖于 x_i ，其值等于点扩散函数沿 y_i 方向的线积分。

如果物为一狭像，实际上在像面上形成的分布就是线扩散函数。我们可以用一个与狭缝方向平行的刀片放置在像面上。开始时，刀片完全挡住狭缝像，刀片逐渐移动，也就逐渐放入狭缝像的光。在图 4.3.6 中所示的是狭缝线扩散函数 $L(x_i)$ ，刀片刃口移动到位置 x_i 时，放入的光通量与图中阴影面积成比例。这样一来，在刀片的整个移动过程中，进入探测器的光通量随刀口位置 x_i 的变化，得到一个函数 $E(x_i)$ ，这个函数称为边缘函数。

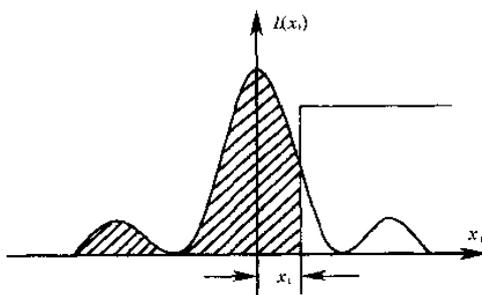


图 4.3.6 由线扩散函数产生的边缘扩散函数

边缘扩散函数 $E(x_i)$ 来源于线扩散函数 $L(x_i)$ ，它们的关系是：

$$E(x_i) = \int_{-\infty}^{x_i} L(\xi) d\xi \quad (4.3.25)$$

对上式微分，可得：

$$\frac{dE(x_i)}{dx_i} = L(x_i) \quad (4.3.26)$$

边缘扩散函数也可用下面方式导出，对系统输入一个阶跃函数，如均匀照射的直边或刀口形成的光分布。系统的输出叫阶跃响应或边缘扩散函数，即，

$$\begin{aligned}
 E(x_i) &= \text{step}(x_i) \otimes h(x_i, y_i) \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(\xi, \eta) \text{step}(x_i - \xi) d\xi d\eta \\
 &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} h(\xi, \eta) d\eta \right] \text{step}(x_i - \xi) d\xi \quad (4.3.27) \\
 &= \int_{-\infty}^{\infty} L(\xi) \text{step}(x_i - \xi) d\xi \\
 &= \int_{-\infty}^{x_i} L(\xi) d\xi
 \end{aligned}$$

2. 相干线性扩散函数和边缘扩散函数

在相干照射下的狭缝在像面上产生的复振幅就是相干线扩散函数，其一维傅里叶变换等于系统的传递函数沿 ξ 方向的截面分布，即有：

$$F\{L(x_i)\} = F\left\{\int_{-\infty}^{\infty} h(x_i, \eta) d\eta\right\} = H(\xi, 0) \quad (4.3.28)$$

于是改变相干照射的狭缝方向，分别对每一个方向测量线扩散函数，然后做一维傅里叶变换，就可确定相应各个方向的传递函数截面。点扩散为圆对称时，传递函数也是圆对称的，因而只需要一个截面就可能完全确定。如果点扩散函数是对 x_i, y_i 可分离变量的，那么传递函数也是可以分离变量的，因而只需要两个截面 $H(\xi, 0)$ 和 $H(\eta, 0)$ 就可以确定。

用线扩散函数的一维傅里叶变来确定传递函数，有时比由点扩散函数做二维傅里叶变换得到传递函数更为方便。由(4.3.28)式可得，一个平行于 y_0 轴的狭缝在像面上产生的相干线扩散函数为：

$$L(x_i) = F^{-1}\{H(\xi, 0)\} \quad (4.3.29)$$

上式是相干传递函数沿 ξ 轴截面的一维傅里叶逆变换。在衍射受限系统中的相干传递函数在通频带内为常数，无论孔径形状如何，相干传递函数的截面总是矩形函数，因而 $L(x_i)$ 呈 sinc 函数变化。对于衍射受限系统， $L(x_i)$ 可表示为：

$$L(x_i) = F^{-1}\{P(\lambda d_i \xi, 0)\} \quad (4.3.30)$$

如，直径为 D 的圆形光瞳，垂直于孔径的任意截面，都矩形函数，其光瞳函数为：

$$P(\lambda d_i \xi, 0) = \text{rect}\left(\frac{\lambda d_i \xi}{D}\right) \quad (4.3.31)$$

线扩散函数为：

$$L(x_i) = F^{-1}\left\{\text{rect}\left(\frac{\lambda d_i \xi}{D}\right)\right\} = \frac{D}{\lambda d_i} \text{sinc}\left(\frac{D x_i}{\lambda d_i}\right) \quad (4.3.32)$$

由式(4.3.25)，物面上放置一个刀口或直边，相干光均匀照射，像面上得到的相干边缘扩散函数由式(4.3.25)式确定，将(4.3.32)式代入(4.3.25)式得：

$$E(x_i) = \int_{-\infty}^{x_i} \frac{D}{\lambda d_i} \text{sinc}\left(\frac{D}{\lambda d_i} \xi\right) d\xi \quad (4.3.33)$$

上式展开，可以得到：

$$E(x_i) = \frac{1}{2} + \frac{1}{\pi} \left[\frac{\pi D x_i}{\lambda d_i} - \frac{1}{18} \left(\frac{\pi D x_i}{\lambda d_i} \right)^2 + \dots \right] \quad (4.3.34)$$

图 4.3.7 中给出衍射受限的相干线扩散函数与边缘扩散函数。可以看到，边缘扩散函数是振荡的，而直边的像也不再是亮暗严格分明的。

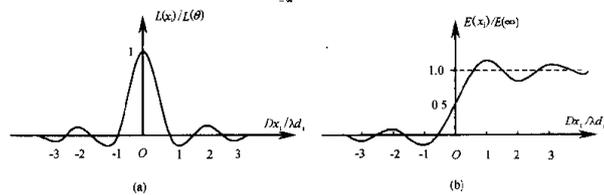


图 4.3.7 相干线扩散函数和边缘扩散函数

§ 4.4 衍射受限非相干成像系统的传递函数

当光学系统用非相干光照射时将会看到，系统的传递函数仍由出瞳确定，但二者之间的关系较为间接，不像相干传递函数那样简单，并直接由光瞳函数表示。非相干光照射的情形更复杂一些，并且更有趣。因此，非相干成像理论比起相干成像理论，内容更为丰富。在非相干照射下，物面上各点的振幅和相位随时间变化的方式是彼此独立、统计无关的。这样一来，虽然物面上的每一点通过系统仍可得到一个对应的复振幅分布，但由于物面的照射是非相干的，故不能通过对这些复振幅分布的相干叠加得到像的复振幅分布，而应该先由这些复振幅分别求出对就的强度分布，然后将这些强度分布叠加（非相干叠加）面得到像面强度分布。非相干成像是强度的线性系统，若成像是空不变的，则非相干成像是强度的线性空不变系统。

§ 4.4.1 非相干系成像系统的光学传递函数 (OTF)

实际的照射光源都不会是理想单色的，总具有一定的频带宽度，而成为非单色光。这时，由于不同频率的光波是独立进行传播的，光扰动的振幅和相位随时间发生各自的变化，而且这种变化具有统计无关的性质。所以，对非单色光照射情讨论要复杂的多。为了简化讨论起见，这里只限于分析准单色光情形。若照射光波的时间频带

宽度为 $\Delta\nu$ ，其中心频率为 ν_0 ，并且满足条件：

$$\frac{\Delta\nu}{\nu} \ll 1 \tag{4.4.1}$$

则称为光波准单色光(Quasi-monochromatic light)。当用准单色光照射时，可设物平面上光扰动的分布函数为 $f(x_0, y_0; t)$ 。要得到 $f(x_0, y_0; t)$ 在像平面上的响应 $g(x_i, y_i; t)$ ，可先采用傅里叶分析的方法，把 $f(x_0, y_0; t)$ 分解成一

系统单色波的线性组合，这样就可应用前段对单色光照射下获得的结果，求出一系统对每一单色波的响应，最后再把各个单色波的这些响应叠加

起来，就得到总的响应 $g(x_i, y_i; t)$ 。整个过程示

意如图 4.4.1 所示。下面，我们先分析一下准单色光照射时，光学成像系统的物像关系。

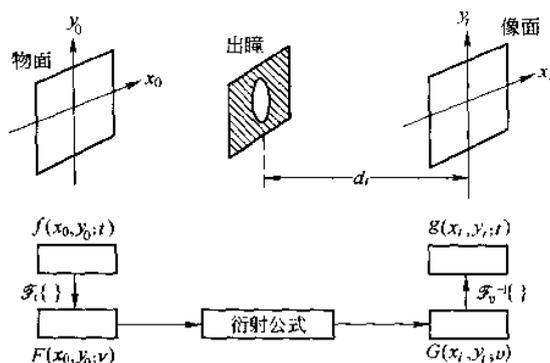


图 4.4.1 准单色光照射时物像关系图

先对 $f(x_0, y_0; t)$ 求关系变量 t 的傅里叶变换，即：

$$\begin{aligned} F(x_0, y_0; \nu) &= F\{f(x_0, y_0; t)\} \\ &= \int_{-\infty}^{\infty} f(x_0, y_0; t) e^{-i2\pi\nu t} dt \end{aligned} \tag{4.4.2}$$

$F(x_0, y_0; \nu)$ 是时间频率为 ν 的单色光波在物平面上的复振幅分布函数。按照前面对单色光情况的讨论, 对于衍射受限成像系统, 直接由叠加积分公式便可求得频率为 ν 的单色平面波在像平面上的响应 $G(x_i, y_i; \nu)$ 为:

$$G(x_i, y_i; \nu) = \int \int_{-\infty}^{\infty} F(x_0, y_0; \nu) \times h(x_i - x_0, y_i - y_0) dx_0 dy_0 \quad (4.4.3)$$

$G(x_i, y_i; \nu)$ 又可看成是实际输出像 $g(x_i, y_i; t)$ 的频谱函数, 从而有:

$$\begin{aligned} g(x_i, y_i; t) &= F^{-1} \{G(x_i, y_i; \nu)\} \\ &= \int_{-\infty}^{\infty} G(x_i, y_i; \nu) e^{i2\pi\nu t} d\nu \\ &= \int_{-\infty}^{\infty} d\nu e^{i2\pi\nu t} \times \int \int_{-\infty}^{\infty} F(x_0, y_0; \nu) \times h(x_i - x_0, y_i - y_0; \nu) dx_0 dy_0 \end{aligned} \quad (4.4.4)$$

假设系统的性态不随时间改变, 并用中心频率为 ν_0 的准单色光照射, 此时, $F(x_0, y_0; \nu)$ 只有在 $\nu = \nu_0$ 的窄带范围内不为零, 在此范围外可视为零。故在计算积分式(4.4.4)时, 可近似地将脉冲响应函数 h 中的 ν 用 ν_0 代替, 于是(4.4.4)可以写成:

$$\begin{aligned} g(x_i, y_i; t) &= \int \int_{-\infty}^{\infty} dx_0 dy_0 h(x_i - x_0, y_i - y_0; \nu_0) \times \left[\int_{-\infty}^{\infty} F(x_0, y_0; \nu) e^{i2\pi\nu t} d\nu \right] \\ &= \int \int_{-\infty}^{\infty} f(x_0, y_0; t) h(x_i - x_0, y_i - y_0; \nu_0) dx_0 dy_0 \end{aligned} \quad (4.4.5)$$

这就把叠加公式(4.1.2)推广到了准单色光情形。

由于光接收器(如肉眼、照相乳胶和光电探测器)都只能感知光的强度, 且其响应频率远小于光波频率, 故光接收器所感知到的像平面上的光强度为:

$$I(x_i, y_i) = \langle g(x_i, y_i; t) g^*(x_i, y_i; t) \rangle \quad (4.4.6)$$

式中尖括号表示对时间求平均。将式(4.4.5)代入上式后得:

$$\begin{aligned} I(x_i, y_i) &= \left\langle \int \int_{-\infty}^{\infty} f(x_0, y_0; t) h(x_i - x_0, y_i - y_0; \nu_0) dx_0 dy_0 \cdot \int \int_{-\infty}^{\infty} f^*(x, y; t) h^*(x_i - x, y_i - y; \nu_0) dx dy \right\rangle \\ &= \int \int_{-\infty}^{\infty} \int \int_{-\infty}^{\infty} \langle f(x_0, y_0; t) f^*(x, y; t) \rangle h(x_i - x_0, y_i - y_0; \nu_0) h^*(x_i - x, y_i - y; \nu_0) dx_0 dy_0 dx dy \end{aligned} \quad (4.4.7)$$

对于准单色光照射时, 在物平面上的复振幅分布函数中, 幅值随时间作缓慢变化, 而相位部分将因光波频率很高而随时间迅速变化。因此, 对于物面上的任意两点 (x_0, y_0) 和 (x, y) 处的光振动可分别写成:

$$f(x_0, y_0; t) = f(x_0, y_0) e^{i\phi(x_0, y_0; t)} \quad (4.4.8a)$$

$$f(x, y; t) = f(x, y) e^{i\phi(x, y; t)} \quad (4.4.8b)$$

将上式代入(4.4.7)式, 则有:

$$\langle f(x_0, y_0; t) f^*(x, y; t) \rangle = f(x_0, y_0) f^*(x, y) e^{i[\phi(x_0, y_0; t) - \phi(x, y; t)]} \quad (4.4.9)$$

而且, 随着照射方式的不同, 由式(4.3.7)和(4.3.9)会得出不同意义的结果。下面只讨论两类典型的照射, 即相干照射和非相干照射。

1. 相干照射(Coherent illumination)

在有激光器发出的光波、一个普通光源通过针孔后(点光源)出射的光波等这类光源照射下,物平面上任意两点光振动之间的相位差随时间的变化是恒定的,这种照射方式称为空间相干照射。此时,式(4.4.9)式中的相位差的平均值等于常数。因此,不失一般性地,可令其等于1。于是,由式(4.4.7)和(4.4.9),得到像面上的光强分布为:

$$I(x_i, y_i) = \left\langle \int \int_{-\infty}^{\infty} f(x_0, y_0; t) h(x_i - x_0, y_i - y_0; \nu_0) dx_0 dy_0 \cdot \int \int_{-\infty}^{\infty} f^*(x, y; t) h^*(x_i - x, y_i - y; \nu_0) dx dy \right\rangle$$

$$= |f(x_i, y_i) * h(x_i, y_i)|^2 = |g(x_i, y_i)|^2$$

(4.4.10)

上式表明,在相干照射下,衍射受限光学成像系统对光场复振幅变换而言是线性空不变系统;对于光强度的变换,则不是线性系统。

2. 非相干照射(Incoherent illumination)

如在漫射光源、扩展光源等这类光源照射下,物平面上各点的光振动随时间的变化都是统计无关的。这时,式(4.3.7)中的 $f(x_0, y_0; t)f^*(x, y; t)$ 除了在点 (x_0, y_0) 足够小的邻域内不为零外,在其余区域的值全为零。于是,对于物平面上靠得很近的两点的光振动,式(4.3.9)式可写成:

$$\langle f(x_0, y_0; t)f^*(x, y; t) \rangle$$

$$= \begin{cases} f(x_0, y_0)f^*(x, y), & (x - x_0)^2 + (y - y_0)^2 \leq \varepsilon^2 \\ 0, & \text{others} \end{cases} \quad (4.4.11)$$

式中, ε 为任意小的正数。或写成:

$$\langle f(x_0, y_0; t)f^*(x, y; t) \rangle = f(x_0, y_0)f^*(x, y)\delta(x - x_0, y - y_0) \quad (4.4.12)$$

将上式代入(4.3.7)式,得到非相干照射像面上的光强分布为:

$$I(x_i, y_i) = \int \int_{-\infty}^{\infty} |f(x_0, y_0)|^2 |h(x_i - x_0, y_i - y_0; \nu_0)|^2 dx_0 dy_0$$

$$= I_0(x_i, y_i) \otimes h_I(x_i, y_i) \quad (4.4.13)$$

式中, $I_0(x_0, y_0) = |f(x_0, y_0)|^2$ 是物平面上的强度分布; $I_I(x_i, y_i) = |h(x_i, y_i)|^2$ 称为系统的强度点扩散函数。

上式表明,在非相干照射方式下,衍射受限光学系统成像对光强度的变换是线性空不变的,而对复振幅的变换则不是线性的。

由上面的一般讨论,我们可看到,对于非相干线性空不变成像系统,物像关系满足下述卷积积分:

$$I_i(x_i, y_i) = k \int \int_{-\infty}^{\infty} I_g(\tilde{x}_0, \tilde{y}_0) h_I(x_i - \tilde{x}_0, y_i - \tilde{y}_0) d\tilde{x}_0 d\tilde{y}_0$$

$$= k I_g(x_i, y_i) * h_I(x_i, y_i) \quad (4.4.14)$$

式中, I_g 是几何光学理想像的强度分布,为像强度分布, k 是常数,由于它不影响 I_i 的分布形式,所以不用给出具体表达式。 h_I 为强度脉冲响应(或称为相干脉冲响应、点扩散函数)。它是点物产生的像斑的强度分布。从(4.4.13)式可以看出,也应该是复振幅点扩散函数绝对值的平方,即:

$$h_I(x_i, y_i) = |\tilde{h}(x_i, y_i)|^2 \quad (4.4.15)$$

式(4.4.14)和(4.4.15)表时,在非相干照射下,线性空不变成像系统的像强分布是理想像的强度分布与强度点扩散函数的卷积。系统的成像特性由 $h_I(x_i, y_i)$ 表示,而 $h_I(x_i, y_i)$ 又由 $\tilde{h}(x_i, y_i)$ 决定。

对于非相干照射下的强度线性空不变系统，在频域中来描写物像关系更加方便。将分式(4.4.14)两边进行傅里叶变换并略去无关紧要的常数后得：

$$G_i(\xi, \eta) = G_g(\xi, \eta)H_l(\xi, \eta) \quad (4.4.16)$$

式中：

$$G_i(\xi, \eta) = F\{I_i(x_i, y_i)\} \quad (4.4.17a)$$

$$G_g(\xi, \eta) = F\{I_g(x_i, y_i)\} \quad (4.4.17b)$$

$$H_l(\xi, \eta) = F\{h_l(x_i, y_i)\} \quad (4.4.17c)$$

式中， $G_i(\xi, \eta)$ 、 $G_g(\xi, \eta)$ 和 $H_l(\xi, \eta)$ 分别表示像强度、物强度和强度脉冲响应函数的频谱函数。由于 $I_i(x_i, y_i)$ 、 $I_g(x_i, y_i)$ 和 $h_l(x_i, y_i)$ 都是强度分布，都是非负实函数，故其傅里叶变换是厄米函数，即有：

$$G_i(\xi, \eta) = G_i^*(-\xi, -\eta) \quad (4.4.18)$$

令

$$G_i(\xi, \eta) = A(\xi, \eta)e^{i\phi(\xi, \eta)} \quad (4.4.19)$$

由式(4.4.18)有：

$$A(\xi, \eta)e^{i\phi(\xi, \eta)} = A(-\xi, -\eta)e^{-i\phi(-\xi, -\eta)} \quad (4.4.20)$$

由此得到：

$$A(\xi, \eta) = A(-\xi, -\eta) \quad (4.4.21a)$$

$$\phi(\xi, \eta) = -\phi(-\xi, -\eta) \quad (4.4.21b)$$

即 $G_i(\xi, \eta)$ 的模是偶函数，幅角是奇函数。将式(4.4.19)式取逆傅里叶变换，得：

$$I_i(x_i, \eta) = F^{-1}\{A(\xi, \eta)e^{i\phi(\xi, \eta)}\} = \int \int_{-\infty}^{\infty} A(\xi, \eta)e^{i\phi(\xi, \eta)} e^{i2\pi(\xi x_i + \eta y_i)} d\xi d\eta \quad (4.4.22)$$

并取其中对应的正频率项与负频率项相加，按照欧拉公式可得到一个频率的余弦分量，即，

$$\begin{aligned} & A(\xi, \eta)e^{i\phi(\xi, \eta)} e^{i2\pi(\xi x_i + \eta y_i)} + A(-\xi, -\eta)e^{-i\phi(-\xi, -\eta)} e^{i2\pi(-\xi x_i - \eta y_i)} \\ &= A(\xi, \eta)[e^{i\phi(\xi, \eta)} e^{i2\pi(\xi x_i + \eta y_i)} + e^{-i\phi(-\xi, -\eta)} e^{i2\pi(-\xi x_i - \eta y_i)}] \\ &= A(\xi, \eta) \cdot 2 \cos[2\pi(\xi x_i + \eta y_i) + \phi(\xi, \eta)] \end{aligned} \quad (4.4.23)$$

像面上整个光强分布可视为各空间频率的余弦函数分布的光强分量的叠加求和。各余弦分量的模和幅角可以是互不相同的。最后得到：

$$I_i(x_i, y_i) = \int \int_{-\infty}^{\infty} d\xi d\eta \times A(\xi, \eta) \times 2 \cos[2\pi(\xi x_i + \eta y_i) + \phi(\xi, \eta)] \quad (4.4.24)$$

由于光强度不可能是负值，余弦分量的负值必然截止在零频率分量 $A(0,0)$ 上，故总和仍然是正的值。将 $\xi = \eta = 0$ 代入式(4.4.20)，得零频分量：

$$A(0,0)e^{i\phi(0,0)} = A(0,0)e^{-i\phi(0,0)} \quad (4.4.25)$$

即:

$$\phi(0,0) = 0, \text{ 表示零频无相位因子}$$

$A(0,0)$, 是一个正值实数

现在返回看(4.4.16)式。由于光强度总是非负的实函数, 因而必有一个常数分量即零频分量, 而且它的幅值大于任何非零分量的幅值, 即:

$$G(0,0) \geq |G_i(\xi, \eta)| \quad (4.4.26a)$$

$$G_g(0,0) \geq |G_g(\xi, \eta)| \quad (4.4.26b)$$

$$H_l(0,0) \geq |H_l(\xi, \eta)| \quad (4.4.26c)$$

实践表明, 人眼或光探测器对图像的视觉效果, 在很大程度上取决于像所携带的信息与直流背景的相对比值, 即像的清晰与否, 主要的不是包含零频分量在内的总光强有多大, 而在于携带有信息那部分光强相对于零频分量的比值有多大。

这就启示我们用零频对 $G_i(\xi, \eta)$ 、 $G_g(\xi, \eta)$ 和 $H_l(\xi, \eta)$ 进行归一化, 得到归一化频谱, 即:

$$G'_i(\xi, \eta) = \frac{G_i(\xi, \eta)}{G_i(0,0)} = \frac{\iint_{-\infty}^{\infty} I_i(x_i, y_i) e^{-i2\pi(\xi x_i + \eta y_i)} dx_i dy_i}{\iint_{-\infty}^{\infty} I_i(x_i, y_i) dx_i dy_i} \quad (4.4.27)$$

$$G'_g(\xi, \eta) = \frac{G_g(\xi, \eta)}{G_g(0,0)} = \frac{\iint_{-\infty}^{\infty} I_g(x_i, y_i) e^{-i2\pi(\xi x_i + \eta y_i)} dx_i dy_i}{\iint_{-\infty}^{\infty} I_g(x_i, y_i) dx_i dy_i} \quad (4.4.28)$$

$$H'_l(\xi, \eta) = \frac{H_l(\xi, \eta)}{H_l(0,0)} = \frac{\iint_{-\infty}^{\infty} h_l(x_i, y_i) e^{-i2\pi(\xi x_i + \eta y_i)} dx_i dy_i}{\iint_{-\infty}^{\infty} h_l(x_i, y_i) dx_i dy_i} \quad (4.4.29)$$

由于 $G_i(\xi, \eta) = G_g(\xi, \eta)H_l(\xi, \eta)$, 并且 $G_i(0,0) = G_g(0,0)H_l(0,0)$, 所以, 得到的归一化频谱满足公式:

$$G'_i(\xi, \eta) = G'_g(\xi, \eta)H'_l(\xi, \eta) \quad (4.4.30)$$

这样, 我们称 $H'_l(\xi, \eta)$ 为非相干成像系统的光学传递函数(Optical transfer function, OTF), 它描述非相干成像系统在频域的效应。它的模 $|H'_l(\xi, \eta)|$ 称为调制传递函数(Modulation transfer function, MTF), 描述了系统对各频率分量对比度的传递特性。其幅角称为相位传递函数(Phase transfer function, PTF), 描述了系统对各频率分量施加的相移。

由于 $G'_i(\xi, \eta)$ 、 $G'_g(\xi, \eta)$ 和 $H'_l(\xi, \eta)$ 一般都是复函数, 都可以用它的模和幅角表示, 于是有:

$$G'_i(\xi, \eta) = |G'_i(\xi, \eta)| e^{i\varphi_i(\xi, \eta)} \quad (4.4.31)$$

$$G'_g(\xi, \eta) = |G'_g(\xi, \eta)| e^{i\varphi_g(\xi, \eta)} \quad (4.4.32)$$

$$H'_I(\xi, \eta) = |H'_I(\xi, \eta)| e^{i\varphi_I(\xi, \eta)} = M(\xi, \eta) e^{i\varphi(\xi, \eta)} \quad (4.4.33)$$

由于 $I_i(x_i, y_i)$, $h_g(x_i, y_i)$ 和 $h_l(x_i, y_i)$ 都是非负实函数, 它们的归一化频谱, 即 $G'_i(\xi, \eta)$ 、 $G'_g(\xi, \eta)$ 和 $H'_I(\xi, \eta)$ 都是厄米函数。余弦函数是这种系统的本征函数, 即强度余弦分量在通过系统后仍为同频率的余弦输出, 其对比度和相位的变化取决于系统传递函数的模和幅角。换句话说, 如果把输入物看做强度透射率呈余弦变化的不同频率的光栅的线性组合, 在成像过程中, OTF 改变的影响是改变这些基元的对比度和相对相位。

如: 一个余弦输入的光强为:

$$\tilde{I}_g(\tilde{x}_0, \tilde{y}_0) = a + b \cos[2\pi(\xi_0 \tilde{x}_0 + \eta_0 \tilde{y}_0) + \varphi_g(\xi, \eta)] \quad (4.4.34)$$

则其频谱 $G_g(\xi, \eta)$ 为:

$$\begin{aligned} G_g(\xi, \eta) &= F\{\tilde{I}_g(\tilde{x}_0, \tilde{y}_0)\} \\ &= a\delta(\xi, \eta) + \frac{b}{2} \left\{ \delta(\xi - \xi_0, \eta - \eta_0) e^{i\varphi_g(\xi_0, \eta_0)} + \delta(\xi + \xi_0, \eta + \eta_0) e^{-i\varphi_g(\xi_0, \eta_0)} \right\} \end{aligned} \quad (4.4.35)$$

由于:

$$I_i(x_i, y_i) = I_g(x_i, y_i) \otimes h_l(x_i, y_i)$$

所以,

$$F\{I_i(x_i, y_i)\} = F\{I_g(x_i, y_i)\} F\{h_l(x_i, y_i)\} \quad (4.4.36)$$

根据 $H'_I(\xi, \eta)$ 的定义,

$$F\{h_l(x_i, y_i)\} = H_I(\xi, \eta) = H_I(0, 0) H'_I(\xi, \eta) \quad (4.4.37)$$

于是:

$$\begin{aligned} F\{I_i(x_i, y_i)\} &= H_I(0, 0) a \delta(\xi, \eta) H'_I(\xi, \eta) + \frac{1}{2} H_I(0, 0) b H'_I(\xi, \eta) \\ &\quad \times \left\{ \delta(\xi - \xi_0, \eta - \eta_0) e^{i\varphi_g(\xi_0, \eta_0)} + \delta(\xi + \xi_0, \eta + \eta_0) e^{-i\varphi_g(\xi_0, \eta_0)} \right\} \end{aligned}$$

对于确定的系统 $H_I(0, 0)$ 是一确定的常数, 对像强度的相对分布没有影响, 所以在下面取其逆变换得到 I_i 时可将其略去不写, 即,

$$I_i(x_i, y_i) = a H'_I(0, 0) + \frac{b}{2} \left[H'_I(\xi_0, \eta_0) e^{i\varphi_g(\xi_0, \eta_0)} e^{i2\pi(\xi_0 x_i + \eta_0 y_i)} + H'_I(-\xi_0, -\eta_0) e^{-i\varphi_g(\xi_0, \eta_0)} e^{-i2\pi(\xi_0 x_i + \eta_0 y_i)} \right]$$

由于有:

$$H'_I(0, 0) = 1$$

$$H'_I(\xi_0, \eta_0) = M(\xi_0, \eta_0) e^{i\varphi(\xi_0, \eta_0)}$$

$$H'_I(-\xi_0, -\eta_0) = M(-\xi_0, -\eta_0) e^{i\varphi(-\xi_0, -\eta_0)} = M(\xi_0, \eta_0) e^{-i\varphi(\xi_0, \eta_0)}$$

最后一步利用了 $H'_i(\xi, \eta)$ 的厄米性。将这些结果代入上式，得到像强度分布为：

$$I_i(x_i, y_i) = a + bM(\xi_0, \eta_0) \times \cos[2\pi(\xi_0 x_i + \eta_0 y_i) + \varphi_g(\xi_0, \eta_0) + \varphi(\xi_0, \eta_0)] \quad (4.4.38)$$

由于 (ξ_0, η_0) 是任意的，故上式可以写成一般形成：

$$I_i(x_i, y_i) = a + bM(\xi, \eta) \times \cos[2\pi(\xi x_i + \eta y_i) + \varphi_g(\xi, \eta) + \varphi(\xi, \eta)] \quad (4.4.39)$$

由此可见，余弦条纹通过线性空不变成像系统后，像仍然是同频率的余弦条纹，只是振幅减小了，相位变化了。振幅的减小和相位的变化都取决于系统的光学传递函数在该频率处的取值。

对于呈余弦变化的强度分布，其对比度或调制度，其定义为：

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \quad (4.4.40)$$

式中， I_{\max} 和 I_{\min} 分别是光强度分布的极大值和极小值。物(或理想像)和像的调制度为：

$$V_g = \frac{I_{g \max} - I_{g \min}}{I_{g \max} + I_{g \min}} = \frac{(a+b) - (a-b)}{(a+b) + (a-b)} = \frac{b}{a} \quad (4.4.41)$$

$$V_i = \frac{I_{i \max} - I_{i \min}}{I_{i \max} + I_{i \min}} = \frac{a + bM(\xi, \eta) - a - bM(\xi, \eta)}{a + bM(\xi, \eta) + a - bM(\xi, \eta)} = \frac{b}{a} M(\xi, \eta) \quad (4.4.42)$$

合关上面上式，得到

$$V_i = M(\xi, \eta) V_g \quad (4.4.43)$$

而 $H'_i(\xi, \eta)$ 的幅角 $\varphi(\xi, \eta)$ 显然是余弦像和余弦物(或理想像)的相位差，即：

$$\varphi_i(\xi, \eta) = \varphi_g(\xi, \eta) + \varphi(\xi, \eta) \quad (4.4.44)$$

即像的对比度等于物的对比度与相应频率的 MTF 的乘积，PTF 给出了相应的相移。空间余弦分布的相位差 $\varphi(\xi, \eta)$ ，体现了余弦像分布 $I_i(x_i, y_i)$ 相对于其物分布的 $I_g(\tilde{x}_0, \tilde{y}_0)$ 移动了多少。当 $\varphi(\xi, \eta)$ 为 2π 时，表示错开的一个条纹，当 $\varphi(\xi, \eta) = \theta$ 弧度时，说明错开了 $\frac{\theta}{2\pi}$ 个条纹。

光学传递函数的模通过系统后，振幅将衰减，即 $M(\xi, \eta) \leq 1$ 。

§ 4.4.2 OTF 和 CTF 的关系

光学传递函数 $H'_i(\xi, \eta)$ 与相干传递函数 $H(\xi, \eta)$ 分别描述同一系统采用非相干和相干照射时的传递函数，它们都决定于系统本身的物理性质，由自相关定理有巴塞伐定理可以得到：

$$\begin{aligned}
 H'_1(\xi, \eta) &= \frac{H_1(\xi, \eta)}{H_1(0, 0)} = \frac{F\{h_1(x_i, y_i)\}}{\iint_{-\infty}^{\infty} h_1(x_i, y_i) dx_i dy_i} = \frac{F\{|\tilde{h}(x_i, y_i)|^2\}}{\iint_{-\infty}^{\infty} |\tilde{h}(x_i, y_i)|^2 dx_i dy_i} = \frac{H(\xi, \eta) \otimes H(\xi, \eta)}{\iint_{-\infty}^{\infty} |H(\zeta, \iota)|^2 d\zeta d\iota} \\
 &= \frac{\iint_{-\infty}^{\infty} H^*(\zeta, \iota) H(\xi + \zeta, \eta + \iota) d\zeta d\iota}{\iint_{-\infty}^{\infty} |H(\zeta, \iota)|^2 d\zeta d\iota}
 \end{aligned}$$

(4.4.45)

因此，对于同一系统来说，光学传递函数等于相干传递函数的自相关归一化函数。这一结论对于有像差与无像差的情况都成立。

§ 4.4.3 衍射受限的 OTF

对于相干照射的衍射受限系统，由(4.3.6)，我们已经知道： $H(\xi, \eta) = P(\lambda d_i \xi, \lambda d_i \eta)$ ，代入(4.4.45)式，可得：

$$H'_1(\xi, \eta) = \frac{\iint_{-\infty}^{\infty} P(\lambda d_i \zeta, \lambda d_i \iota) P[\lambda d_i (\xi + \zeta), \lambda d_i (\eta + \iota)] d\zeta d\iota}{\iint_{-\infty}^{\infty} P(\lambda d_i \zeta, \lambda d_i \iota) d\zeta d\iota} \quad (4.4.46)$$

令 $x = \lambda d_i \zeta$ ， $y = \lambda d_i \iota$ ，积分变量的替换不会影响积分结果，于是得到：

$$H'_1(\xi, \eta) = \frac{\iint_{-\infty}^{\infty} P(x, y) P(x + \lambda d_i \xi, y + \lambda d_i \eta) dx dy}{\iint_{-\infty}^{\infty} P(x, y) dy dy} \quad (4.4.47)$$

上式表明，衍射由于光瞳函数只有 1 和 0 两个值，分母中的 P^2 与 P 是等价的，因为写成了 P 。这个式子的几何意义解释如下。

式中分母是光瞳的总面积 S_0 ，分子则中心位于 $(-\lambda d_i \xi, -\lambda d_i \eta)$ 的经过平称的光瞳与原光瞳的重叠面积 $S(\xi, \eta)$ ，求衍射受限系统的 OTF 只不过是计算归一化的重叠面积，即有：

$$H'_1(\xi, \eta) = \frac{\text{出瞳重叠瞳重}}{\text{出瞳总瞳总}} = \frac{S(\xi, \eta)}{S_0} \quad (4.4.48)$$

如图 4.4.2 所示，重叠面积取决于两个错开的光瞳的相对位置，也就是和频率 (ξ, η) 有关。对于简单几何形状的光瞳不难求得归一化重叠面积的数学表达式。对于复杂的光瞳，可用计算机计算在一系列分立频率上的 OTF。

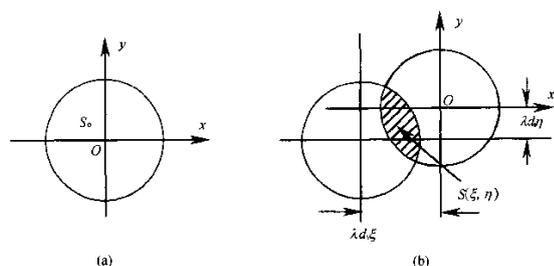


图 4.4.2 衍射受限系统 OTF 的几何解释

由此，我们可以总结衍射受限系统的 OTF 的一些性质：

(1) 由于 $H'_l(\xi, \eta)$ 是实的非负函数，因此衍射受限的非相干成像系统只改变各频率余弦分量的对比，而不改变它们的相位。即，只需考虑 MTF 而不必考虑 PTF。

(2) $H'_l(0,0) = 1$ 。当 $\xi = 0, \eta = 0$ 时，两个光瞳完全重叠，归一化重叠面积为 1，这正是 OTF 归一化的结果。这并不意味着物和像的背景光强相同。由于吸收、反射、散射及光阑挡光等原因，像面背光强总要弱于物面光强。但从对比度考虑，物像方零频分量的对比度都是零，无所谓衰减，所以，才有 $H'_l(0,0) = 1$ 。

(3) $H(\xi, \eta) \leq H'_l(0,0)$ 。这从两个光瞳错开后重叠的面积小于完全重叠面积，可以看出。

(4) 截止频率 ξ_0, η_0 。两出瞳从完全重合开始，分别朝相反方向平移，直至重叠面积刚好为零时，它们已经移开了 $2x_{\max}$ 和 $2y_{\max}$ ，于是有：

$$2x_{\max} = \lambda d_i \xi, \quad 2y_{\max} = \lambda d_i \eta$$

从而求得：

$$\xi_0 = \frac{2x_{\max}}{\lambda d_i} \quad (4.4.49a)$$

$$\eta_0 = \frac{2y_{\max}}{\lambda d_i} \quad (4.4.49b)$$

与相干成像系统相比，非相干成像系统的截止频率是相干成像系统的两倍。

由前面的讨论，尤其是由式(4.4.48)式，可以将计算 OTF 的步骤总结如下：

- (1) 确定系统出瞳的形状和大小，计算出瞳总面积 S_0 ；
- (2) 计算出瞳面至像平面之间的距离 d_i ；
- (3) 任意给定一组 (ξ, η) 值，算出 $(\lambda d_i \xi, \lambda d_i \eta)$ 值；将出瞳平移，使其中心落在 $(-\lambda d_i \xi, -\lambda d_i \eta)$ 处，计算移动前后两出瞳的重叠面积。
- (4) 相继再给定一组 (ξ, η) 值，再算出重叠面积。依次类推，就可算出 $S(\xi, \eta)$ 值。
- (5) 按公式(4.4.48)式计算得到 $H'_l(\xi, \eta)$ 。

例 4.4.1 衍射受限非相干成像系统的光瞳为边长 l 的正方形，求其光学传递函数。

解：此时的光瞳函数可表为：

$$P(x, y) = \text{rect}\left(\frac{x}{l}\right) \text{rect}\left(\frac{y}{l}\right)$$

显然光瞳总面积为： $S_0 = l^2$ ，当 $P(x, y)$ 在 x, y 方向分别位移 $-\lambda d_i \xi, -\lambda d_i \eta$ 以后，得 $P(x + \lambda d_i \xi, y + \lambda d_i \eta)$ ，从图 4.4.3 可以求出 $P(x, y)$ 和 $P(x + \lambda d_i \xi, y + \lambda d_i \eta)$ 和重叠面积 $S(\xi, \eta)$ 。由图可得：

$$S(\xi, \eta) = \begin{cases} (l - \lambda d_i \xi)(l - \lambda d_i \eta), & \xi, \eta > 0 \\ (l + \lambda d_i \xi)(l + \lambda d_i \eta), & \xi, \eta < 0 \\ 0 & \lambda d_i |\xi| > l, \lambda d_i |\eta| > l \end{cases}$$

即:

$$S(\xi, \eta) = \begin{cases} (l - \lambda d_i |\xi|)(l - \lambda d_i |\eta|), & \lambda d_i |\xi| < l, \\ & \lambda d_i |\eta| < l \\ 0, & \text{其他} \end{cases}$$

光学传递函数为:

$$H(\xi, \eta) = \frac{S(\xi, \eta)}{S_0} = \text{tri}\left(\frac{\xi}{2\rho_c}\right) \text{tri}\left(\frac{\eta}{2\rho_c}\right)$$

式中, $\rho_c = l/2\lambda d_i$, 是同一系统采用相干照明时的截止频率。非相干系统沿 ξ 和 η 轴方向上的截止频率是 $2\rho_c = l/\lambda d_i$ 。见图 4.4.3 所示。

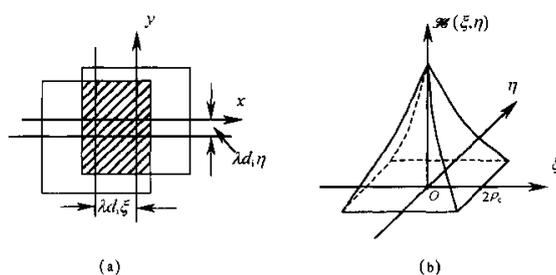


图 4.4.3 方形光瞳衍射受限 OTF 计算

例 4.4.2 衍射受限系统的出瞳直径为 D 的圆, 求此系统的光学传递函数。

解: 由于是圆形光瞳, OTF 应该是圆对称的。只要沿 ξ 轴计算 H 即可。参看图 4.4.4(a), 在 x 轴方向移动 $\lambda d_i \xi$ 后, 交叠积被 AB 分成两个面积相等的弓形。根据几何公式, 交叠面积为:

$$S(\xi, 0) = \frac{D^2}{2} (\theta - \sin \theta \cos \theta)$$

其中 $\cos \theta$ 由下式定义:

$$\cos \theta = \frac{\lambda d_i \xi / 2}{D/2} = \frac{\lambda d_i \xi}{D}$$

在截止频率范围内:

$$(H, 0) = \frac{S(\xi, 0)}{S_0} = \frac{S(\xi, 0)}{\pi D^2 / 4} = \frac{2}{\pi} (\theta - \sin \theta \cos \theta)$$

截止频率满足: $\lambda d_i \xi = D$, 也就是两个圆中心距离大于直径 D 时, 重叠面积为零。此种系统的相干传递函数的截止频率: $\rho_c = D/2\lambda d_i \xi$ 。显然光学传递函数的截止频率恰好又是 $2\rho_c$ 。图 4.4.4(b) 画出了光瞳函数为圆域函数时, $H(\xi, \eta)$ 示意图。 $H(\xi, \eta)$ 在极坐标中的表达式为:

$$H(\xi, 0) = \begin{cases} \frac{\pi}{2} (\theta - \sin \theta \cos \theta), & \rho \leq D / \lambda d_i \\ 0, & \text{其他} \end{cases}$$

式中:

$$\rho = \sqrt{\xi^2 + \eta^2} \quad \cos \theta = \frac{\lambda d_i \rho}{D}$$

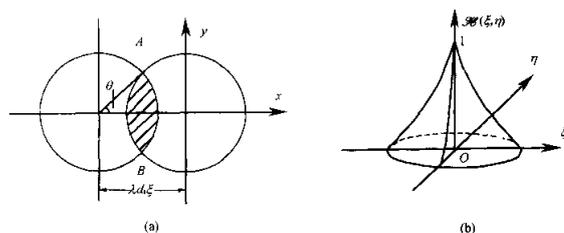


图 4.4.3 圆形光瞳衍射受限的 OTF 计算

§ 4.4.4 非相干线扩散函数和边缘扩散函数

在非相干照射下，平行于 y_0 轴的狭缝光源在像平面上产生的线响应函数称为线扩散函数。它与光学传递函数的关系是[参见(4.3.29)式]:

$$L(x_i) = F^{-1}\{H'_l(\xi, 0)\} \quad (4.4.50)$$

它是 OTF 沿 ξ 轴截面分布的一维傅里叶变换。虽然线扩散函数与传递函数之间的关系，在相干与非相干照射时都是相干照射都是相同的，但由于 OTF 和 CTF 的不同，其线扩散函数也是不同的。如：相干线扩散函数与孔径形状无关，总是 sinc 函数。而 OTF 是光瞳自相关的结果，这样，非相干线扩散函数自然就与孔径的形状有关了。

图 4.4.5(a) 画出了系统具有直径为 D 的圆形光瞳的线扩散函数。与相干线扩散函数的主要差别是，它没有零点。

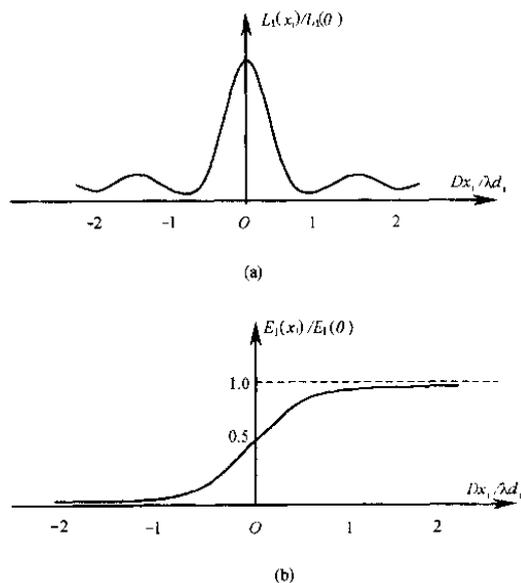


图 4.4.5 非相干线扩散函数与边缘扩散函数

非相干边缘扩散函数，由非相干线扩散函数的积分给出:

$$E_l(x_i) = \int_{-\infty}^{x_i} L_l(\zeta) d\zeta \quad (4.4.51)$$

图 4.4.5(b) 画出了非相干边缘扩散函数的曲线，可以看出它没有相干边缘扩散函数中的振荡现象。

§ 4.5 有像差系统的传递函数

上面, 我们只讨论与没有相干时, 相干照射与非相干照射下的光学传递函数, 这当然是理想的情况。任何一个实际系统, 总是有像差的。像差可能来自于构成系统的元件, 也可能来自成像平面的位置误差, 也可能来自理想球面透镜所固有的如球面像差等。对于有像差的光学系统, 不论造成像差的原因是什么, 总可以归结为波面对于理想球面波的偏离。像差会对传递函数产生影响, 在相干或非相干照射下, 往往都是复函数, 系统将对各频率成分的相位产生影响。

前面我们研究衍射受限系统时, 是通过点扩散函数 $\tilde{h}(x_i, y_i)$ 与光瞳函数傅里叶变化, 最后用光瞳函数来描述传递函数。对于有像差的系统, 我们仍然可以采用这种方法, 只需要对光瞳函数的概念加以推广, 然后用广义光瞳函数来描述有像差系统的传递函数。

如图 4.5.1 所示, 在衍射受限系统中, 单位脉冲 $\delta(\tilde{x}_0, \tilde{y}_0)$ 通过系统后投射到光瞳上的是以理想像点为中心的球面波, 由于系统像差的存在, 使与 O 点等相位的各点形成波面 Σ_1 , 若系统没有像差, 理想波面应该是 Σ_0 。 Σ_1 和 Σ_0 每一点的光程差用函数 $W(x, y)$ 表示, 它的具体形式由系统像差决定, 由它引起的相位变化是 $kW(x, y)$ 。定义:

$$P'(x, y) = P(x, y)e^{ikW(x, y)} = \begin{cases} e^{ikW(x, y)} & \text{Inside} \\ 0 & \text{Outside} \end{cases} \quad (4.5.1)$$

这样, $\tilde{h}(x_i, y_i)$ 可以看做是复振幅透射率为 $P'(x, y)$ 的光瞳被半径为 d_i 的球面波照射后所得的分布, 式中 $P(x, y)$ 即为没有像差时的光瞳函数, 这样, 我们可以把 $P'(x, y)$ 称为广义光瞳函数。 $\tilde{h}(x_i, y_i)$ 也就是广义光瞳函数的傅里叶变换。

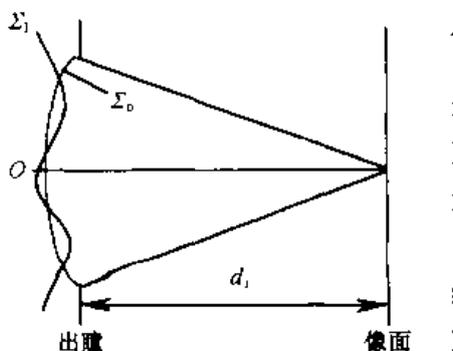


图 4.5.1 像差对于出瞳平面波的影响

我们用广义光瞳函数 $P'(x, y)$ 代替 $P(x, y)$ 就可以得到有像差系统的相干点扩散函数, 即:

$$\tilde{h}(x_i, y_i) = F\{P'(\lambda d_i \tilde{x}, \lambda d_i \tilde{y})\} = F\{P(\lambda d_i \tilde{x}, \lambda d_i \tilde{y})e^{ikW(\lambda d_i \tilde{x}, \lambda d_i \tilde{y})}\} \quad (4.5.2)$$

由上式可知, 相干脉冲响应不再单纯是孔径的夫琅和费衍射图样, 必须考虑波像差的影响。若像差是对称的, 如球差和离焦, 点物的像斑仍具有对称性。如果像差是非对称的, 如彗差、像散点, 点物的像斑也不具有圆对称性。

由前面可知, 相干传递函数是相干点扩散函数的傅里叶变换, 即有:

$$H(\xi, \eta) = P'(\lambda d_i \xi, \lambda d_i \eta) = P(\lambda d_i \xi, \lambda d_i \eta)e^{ikW(\lambda d_i \xi, \lambda d_i \eta)} \quad (4.5.3)$$

由此可见, 系统的通频带的范围仍由光瞳的大小决定, 截止频率和无像差的情况相同。像差的唯一影响是在通带引入了与频率有关的相位畸变, 从而使像质变坏。

在非相干照射下，强度点扩散函数仍然是相干点扩散函数模的平方，即 $h_I = |\tilde{h}|^2$ 。对于圆形光瞳， h_I 不再是爱里斑的强度分布。由于像差的影响，点扩散函数的峰值明显小于没有像差时系统点扩散函数的峰值。可以把这两个峰值之比作为像差大小的指标，称为斯特列尔(Strehl) 清晰度。

同样，由 H 和 H'_I 以及和孔径函数的关系可知，有像差系统的 OTF 应该是广义光瞳函数的归一化自相关函数为：

$$H'_I(\xi, \eta) = \frac{\iint_{-\infty}^{\infty} P'(x, y) P'(x + \lambda d_i \xi, y + \lambda d_i \eta) dx dy}{\iint_{-\infty}^{\infty} P(x, y) dy dy} \quad (4.5.4)$$

由于广义光瞳函数的相位因子不影响式中的分母的积分值，所以，分母部分依然是光瞳的总面积。上式中，分子的积分区域不变，还是 $P(x, y)$ 和 $P(x + \lambda d_i \xi, y + \lambda d_i \eta)$ 的重叠区，于是上式可以写为：

$$H'_I(\xi, \eta) = \frac{\iint_{S(\xi, \eta)} e^{-ikW(x, y)} e^{ikW(x + \lambda d_i \xi, y + \lambda d_i \eta)} dx dy}{S_0} \quad (4.5.5)$$

上式给出了像差引起的相位畸变与 OTF 的直接关系，当波像差为零时，便与衍射受限的 OTF。对于像差不为零的情况，OTF 是复函数。像差不为零不仅影响输入频率成分的对比度，而且也产生相移，利用许瓦兹 (Schwarz) 不等式，像差的存在不会增大 MTF 的值。许瓦兹不等式为：

$$\left| \iint_{-\infty}^{\infty} XY d\xi d\eta \right|^2 \leq \iint_{-\infty}^{\infty} |X|^2 d\xi d\eta \cdot \iint_{-\infty}^{\infty} |Y|^2 d\xi d\eta \quad (4.5.6)$$

$$\text{令： } X = e^{-ikW(\zeta, t)}, Y = e^{ikW(\zeta + \lambda d_i \xi, t + \lambda d_i \eta)}$$

将上式应该到(4.5.5)式，两端取模的平方，再应用许瓦兹不等式，同时注意到 $|X|^2 = |Y|^2 = 1$ ，便得到：

$$\begin{aligned} H'_I(\xi, \eta) \Big|_{\text{像差}} &= \frac{\left| \iint_{S(\xi, \eta)} e^{-ikW(\zeta, t)} e^{ikW(\zeta + \lambda d_i \xi, t + \lambda d_i \eta)} d\zeta dt \right|^2}{S_0} \\ &\leq \frac{\iint_{S(\xi, \eta)} \left| e^{-ikW(\zeta, t)} \right|^2 d\zeta dt \cdot \iint_{S(\xi, \eta)} \left| e^{ikW(\zeta + \lambda d_i \xi, t + \lambda d_i \eta)} \right|^2 d\zeta dt}{S_0} \\ &= \frac{\left| \iint_{S(\xi, \eta)} d\zeta dt \right|^2}{S_0} = H'_I(\xi, \eta) \Big|_{\text{像差}} \end{aligned} \quad (4.5.7)$$

由此可见，像差的存在会使光学系统的调制传递函数下降，像面光强度分布的各个空间频率分量的对比度降低，也就是像差会进一步降低成像质量。但可以证明，只要是同样大小和形状出射光瞳，则对于有像差的系统和无像差的系统，其截止空间频率是一样的。

由于 h_I 是实数，无论有无像差， H'_I 都是厄米型的，即有： $H'_I(\xi, \eta) = H'_I(-\xi, -\eta)$ 。它的模和幅角分别为偶函数和奇函数，即：

$$M(\xi, \eta) = M(-\xi, -\eta) \quad (4.5.8)$$

$$\varphi(\xi, \eta) = -\varphi(-\xi, -\eta) \quad (4.5.9)$$

这样，我们在考察 MTF 或 PTF 的截面时，只需画出正频部分。

-----* * * * *-----

§ 4.6 相干与非相干成像系统的比较

现在, 我们对相干成像和非相干成像两种系统作一些比较。这种比较, 虽然并不能简单地得出哪一种照射方式与更优的结果, 但有助于我们更好地理解两种系统之间的联系以及差异。从而, 在实际应用中, 根据具体情况判断选择哪一种方式会更优一些。

§ 4.6.1 截止频率

由上面讨论我们知道, 非相干衍射受限系统的 OTF, 其截止频率扩展到相干系统 CTF 的截止频率的两倍处。因此, 似乎可以得出结论: 对于同一个光学成像系统, 使用非相干照射一定会比采用相干照射得到更好的像。可是从下面的讨论看到, 这一结论一般是不正确的。主要原因是, 相干系统截止频率是确定像的复振幅的最高频率分量, 而非相干系统截止频率是对像的强度的最高频率分量而言。虽然这两种情况中, 最后的可观察量都是像的强度分布, 但由于两种截止频率所描述的物理量不同, 所以不能直接对它们进行比较, 从而简单地得到结论。即使比较的物理量一致, 而要判断绝对好坏也是困难的。

§ 4.6.2 像强度频谱

要对相干照射和非相干照射下像强度地行比较, 可以考察其频谱特性。由前可知, 在相干和非相干照射下, 像强度可分别表示为:

$$I_c(x_i, y_i) = \left| U_g(x_i, y_i) * \tilde{h}(x_i, y_i) \right|^2 \quad (4.6.1)$$

$$I_i(x_i, y_i) = I_g(x_i, y_i) * h_l(x_i, y_i) \quad (4.6.2)$$

式中, I_c 和 I_i 分别是相干和非相干照射下像面的强度分布, U_g 和 I_g 分别为物(或理想像)的复振幅分布和强度分布。对上述两式进行傅里叶变换, 并利用卷积定理和相关定理, 得到相干照射和非相干照射下的像强度分布如下:

$$G_c(\xi, \eta) = [G_{gc}(\xi, \eta)H(\xi, \eta)] \star [G_{gc}(\xi, \eta)H(\xi, \eta)] \quad (4.6.3)$$

$$G_i(\xi, \eta) = [G_{gc}(\xi, \eta) \star G_{gc}(\xi, \eta)] [H(\xi, \eta) \star H(\xi, \eta)] \quad (4.6.4)$$

式中, G_c 和 G_i 分别是相干和非相干照射下像强度的频谱, G_{gc} 是物的复振幅分布的频谱, H 是相干传递函数。

由上两式可以看到, 在两种情况下, 像强度的频谱可能很不相同, 但并不因此, 而简单地得出结论来说明, 一种照射方式比另一种照射方式更好。这是因为, 成像不仅与照射方式有关, 也与系统的结构和物的空间结构有关。这一点, 可从下面的例题中得到进一步的理解。

例 4.6.1 有一单透镜成像系统, 其圆形边框的直径为 7.2cm, 焦距为 10cm, 且物和像等大。设物的透射率函数为:

$$t(x) = |\sin(2\pi x/b)|$$

式中, $b = 0.5 \times 10^{-3}$ cm。今用 $\lambda = 600$ nm 的单色光垂直照射该物, 试解析说明在相干光和非相干光照射情况下, 像面上能否出现强度起伏?

解: 按题设条件, 有:

$$\text{物周期: } T_1 = b/2, \text{ 其频率为: } \rho_1 = \frac{1}{T_1} = \frac{2}{b} = 400 \text{ 线/mm}$$

$d_0 = d_i = 2\rho = 200 \text{ mm}$, 故

$$\rho_c = \frac{D}{2\lambda d_i} = 300 \text{ 线/mm},$$

$$\rho_0 = \frac{D}{\lambda d_i} = 600 \text{ 线/mm}$$

显然, 在相干照射条件下, $\rho_c < \rho_1$, 系统的截止频率小于物的基频, 此时, 系统只允许零频分量通过, 其他频谱分量均被挡住, 所以物不能成像, 像面呈均匀分布。在非相干照射下, $\rho_1 < \rho_c$, 系统的截止频率大于物的基频, 故零频和基频均能通过系统参与成像, 在像面上将有图像存在。基于这种分析, 非相干成像要比相干成像好。但在别的物结构下, 情况将发生变化, 如见下例。

例 4.6.2 在上例中, 如果物的透射率函数为:

$$t(x) = \sin(2\pi x / b)$$

结论又如何?

解: 这时, 物周期: $T_1 = b$, 其频率为: $\rho_1 = \frac{1}{b} = 200 \text{ 线/mm}$; 根据上例的数据, 显然:

$\rho_1 < \rho_c < \rho_0$ 。即在相干照射下, 这个呈正弦分布的物函数复振幅能够不受衰减地通过此系统成像。而对于非相干照射方式, 物函数的基频也小于其截止频率, 故此物函数也能通过系统成像, 但其幅度要随空间频率的增加受到逐渐增大的衰减, 即对比度降低。由此可见, 在这种物结构中, 相干照射方式比非相干照射方式要好。

此外, 相干照射具有严重的散斑效应, 且光学缺陷易在相干照射下观察到, 并容易产生一些木纹状的附加干涉花纹, 对成像的清晰度带来干扰。

另外, 相干照射方式与非相干照射方式对锐边的响应也迥然不同。其原因可作如下解释。我们知道, 相干成像系统的传递函数为:

$$H_c(\xi, \eta) = P(\lambda d_i \xi, \lambda d_i \eta) = \begin{cases} 1 & \text{在出瞳外} \\ 0 & \text{在出瞳内} \end{cases}$$

可知, 具有陡峭的不连续性, 且在截止频率确定的通频带内不衰减, 因而具有较小的误差。

而非相干成像系统传递函数为:

$$H_0(\xi, \eta) = [H_c(\xi, \eta) \star H_c(\xi, \eta)] / \int \int_{-\infty}^{\infty} H_c(\xi, \eta) d\xi d\eta$$

它的截止频率所确定的通频带内, 不像 $H_c(\xi, \eta)$ 那样恒等于 1, 而是随着空间频率的增大逐渐减小, 其结果是降低了像的对比度。

§ 4.6.3 两点分辨

分辨率是评价光学系统成像质量的一个重要指标。对于衍射受限的圆形光瞳情况, 在非相干照射方式下, 根据瑞利分辨判据, 对两个强度相等的非相干点源, 若一个点源产生的爱里斑中心恰好落在另一个点源所产生的的爱里斑的第一个极小上, 则称它们对于非相干衍射受限系统“刚刚能够分辨”的两个点源。由前面所讲的圆孔的夫琅和费衍射花样公式可知, 像斑的归一化强度可表示为:

$$I(r_0) = \left[\frac{2J_1(kdr_0/2z)}{kdr_0/2z} \right]^2 = \left[\frac{2J_1(\pi x)}{\pi x} \right]^2$$

式中, $x = dr_0 / \lambda z$ 。由前可知, 第一个暗环的角

半径为 $x = 1.22$, 所以, 如果把两个点源像的中心沿 x 轴方向分别放在 $x = \pm 0.61$ 处, 则它们正好满足瑞利分辨判据的条件, 且其光强分布可表示为:

$$I(x) = \left\{ \frac{2J_1[\pi(x-0.61)]}{\pi(x-0.61)} \right\}^2 + \left\{ \frac{2J_1[\pi(x+0.61)]}{\pi(x+0.61)} \right\}^2 \quad (4.6.5)$$

图 4.6.1 画出了此强度分布的剖面图，此时，两个点源的爱里斑图样在其中心处约下降 19%。

相对于相干照射方式，两个点源产生的爱里斑则必须按复振幅叠加后，再求其合强度。此强度记为：

$$I(x) = \left\{ \frac{2J_1[\pi(x-0.61)]}{\pi(x-0.61)} + \frac{2J_1[\pi(x+0.61)]}{\pi(x+0.61)} e^{i\phi} \right\}^2 \quad (4.6.6)$$

式中， ϕ 是两物点之间的相位差。显然， $I(x)$ 的值与 ϕ 有关。在图 4.6.2 中画出了 $\phi = 0, \pi/2, \pi$ 时的光强分布。根据图 4.6.1 中的曲线，可以对系统在相干照射和非相干照射条件下的分辨能力进行比较后，得出如下结论：

1) $\phi = 0$ 时，即两点源同相位时， $I(x)$ 不出现中心凹陷，因而两物点完全不能分辨，其系统的分辨能力不如非相干情形好。

2) $\phi = \pi/2$ 时，相干照射的强度 $I(x)$ 与非相干照射所得结果完全相同，从而在两种照射方式下，系统的分辨能力都一样。

3) $\phi = \pi$ 时，即两点源相位相反时，相干照射的强度分布 $I(x)$ 的中心凹陷取极小值，远完低于 19%，故这两点间的分辨要比非相干照射方式下更清楚。

由此可见，到底哪种照射方式对提高两点源间的分辨更为有利，不可能得出一个普遍适用的结论。故瑞利判据仅适用于非相干成像系统，而对于相干成像系统，能否分辨两个点源，则要考虑它们的相位关系。

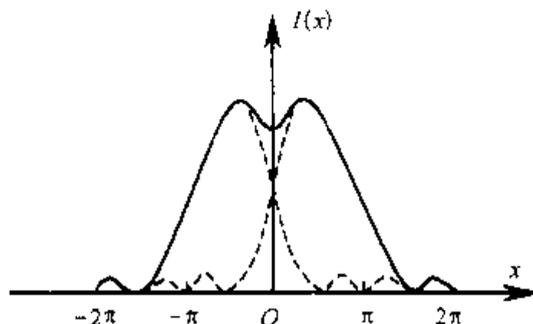


图 4.6.1 刚能分辨两个非相干点源的像强度分布

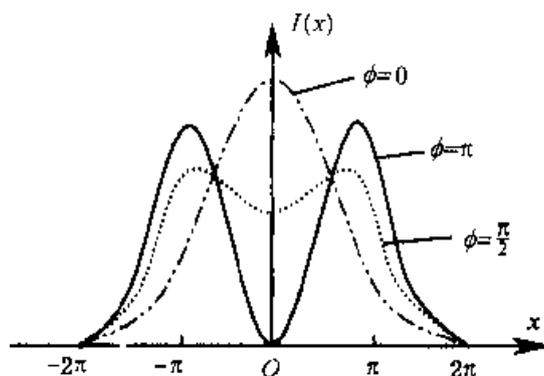


图 4.6.2 相距为瑞利间隔的两个相干点源的像强度分布

第五章 部分相干理论

光学相干性(Optical Coherence)包含时间相干和空间相干, 分别产生于光源的单色性程度和光源的有限尺寸。相干性的好坏使常用其干涉条纹的对比度相描述。在讨论过程中, 为了简便起见, 作了两种理想化的假设, 一种情况是把光源假设为一个理想的点光源, 且具有严格的单色性。这样的光振动具有完全的相干性, 其干涉条纹的对比度可以达到 1。另一种情况, 则假设光源是完全不相干的。用完全不相干的光源照明时得不到干涉条纹, 干涉图的对比度等于 0。

但是, 点光源和单色光的概念都只是一种在数学上的理想化的抽象。严格的单色光在时间上是无限延续的, 传播的波列也是无限长的, 这自然是实际不存在的。同样, 任何点光源也是不存在的, 也是一种理想模型。任何光源都包含有一定的波长范围和尺寸, 这些都会影响到光源的相干性。与此对应, 完全不相干的光源也是一种理想化模型, 即使采用通常认为完全全不相干的太阳光束照射, 在一定条件下也能产生干涉效应。例如, 在杨氏干涉装置中, 只要两个小孔靠的很近 (约 0.02mm), 用太阳光来照射双孔, 也能看到干涉条纹。由于严格相干场和严格的非相干场实际上都不可能得到。因此, 应该研究实际存在于完全相干与完全不相干之间的中间状态, 称为部分相干性 (Partial coherence), 这就我们这章所要讲述的内容。

部分相干性理论是现代光学中较为活跃的一个研究领域, 它既是处理光场统计性质的一种理论 (统计光学方法), 又涉及到光场的量子力学描述 (量子光学)。这里, 我们仅限于讲述部分相干光理论的基本概念和基本规律, 只采用对光场的统计描述, 而不涉及量子光学处理方法。

§ 5.1 多色光场解析信号表示

为了研究光场的相干性, 需要一个矢量函数全面地描述辐射场。但假设只考虑小角度发射的辐射, 不考虑辐射场的偏振效应, 则用一个标量函数来描述辐射场, 对于研究经典相干理论的大部分概念, 这是令人满意的。

在单色光波的讨论中, 我们知道, 对于线性系统, 把一个实函数表示成一个复函数, 常常是很方便。对于单色光场, 由已知的实函数构造一个相应的复函数比较容易, 但对于多色光场就不那么简单明了了。下面我们介绍多色光场的复值表示, 即解析信号表示法。

假设空间某点 \vec{r} 在时刻 t 的光场可以用一个实标量函数 $u^r(\vec{r}, t)$ 来描述。为了简化书写, 下面略去空间变量 \vec{r} , 记为 $u^r(t)$ 。则对于线性系统, 常常把 $u^r(t)$ 表示成与之相关联的一个复函数:

$$u(t) = u^r(t) + iu^i(t)$$

更便于运算, 式中 $u^r(t)$ 、 $u^i(t)$ 分别表示 $u(t)$ 的实部和虚部。 $u(t)$ 称为 $u^r(t)$ 的解析信号 (Analytic signal)。

一、单色信号的复表示

对一个单色信号, 其实函数的方式可表示如下:

$$u^r(t) = A \cos(2\pi\nu_0 t - \phi) \quad (5.1.1)$$

式中, A 是常数振幅, ν_0 为光的频率, ϕ 为初始相位。这个信号的复表示为:

$$u(t) = A e^{-i(2\pi\nu_0 t - \phi)} \quad (5.1.2)$$

其实部正好等于原来的信号 $u^r(t)$ 。这个信号的复振幅定义为:

$$\tilde{A} = A e^{-i\phi} \quad (5.1.3)$$

这个复振幅表示了单色信号的振幅与相位。

一个实函数表示成为一个复函数, 其虚部不是任意的, 它是与原来的实信号密切相关的。那么, 用什么样的方法, 才能得到如 (5.1.2) 式那样的复表示呢? 这个问题在频域可以看得更清楚。

对 (5.1.1) 应用欧拉公式, 得到:

$$u^r(t) = \frac{A}{2} \left[e^{i(2\pi\nu_0 t - \phi)} + e^{-i(2\pi\nu_0 t - \phi)} \right]$$

对上式两边作傅里叶变换得:

$$U^r(\nu) = \frac{A}{2} \left[e^{-i\phi} \delta(\nu + \nu_0) + e^{i\phi} \delta(\nu - \nu_0) \right]$$

式中, $U^r(\nu)$ 是单色实信号的傅里叶谱。然后, 对(5.1.2)式所表示的复信号 $u(t)$, 我们有:

$$F\{u(t)\} = Ae^{i\phi} \delta(\nu - \nu_0) \quad (5.1.4)$$

因此, 从原来的实信号 $u^r(t)$ 变到复信号 $u(t)$, 通过频域中的比较可以看出, 去掉实信号的负频成分, 加倍实信号的正频成分。由此可见, 单色复信号是只有正频分量的单边谱。

二、多色信号的复表示

如果多色信号表示为 $u^r(t)$, 其傅里叶变换存在如下:

$$u^r(t) = \int_{-\infty}^{+\infty} U^r(\nu) e^{j2\pi\nu t} d\nu \quad (5.1.5)$$

式中, $U^r(\nu)$ 是 $u^r(t)$ 的傅里叶谱。由于 $u^r(t)$ 是一个实函数, 故 $U^r(\nu)$ 应是一个厄米函数:

$$U^r(\nu) = [U^r(-\nu)]^* \quad (5.1.6)$$

上式表明 $U^r(\nu)$ 的负频率分量与正频率分量载有同样的信息, 亦即仅正频率(或负频率分量)就携带了实函数的全部信息。因此, 只用正频率分量并不会丢失光场的任何信息。令:

$$U^r(\nu) = A(\nu) e^{i\phi(\nu)} \quad (5.1.7)$$

则有

$$A(\nu) = A(-\nu), \quad \phi(\nu) = -\phi(-\nu) \quad (5.1.8)$$

即 $A(\nu)$ 是偶函数, $\phi(\nu)$ 是奇函数。将式(5.1.7)

式代入(5.1.5)式, 并取其中对应的正频率和负频率项相加, 利用欧拉公式可将(5.1.5)式最后表示成:

$$u^r(t) = 2 \int_0^{+\infty} A(\nu) \cos[2\pi\nu t + \phi(\nu)] d\nu \quad (5.1.9)$$

上式包含所有正频率分量的积分。若把这些频率分量都相移 $\frac{\pi}{2}$, 则可在定义函数 $u^i(t)$ 为:

$$u^i(t) = 2 \int_0^{+\infty} A(\nu) \sin[2\pi\nu t + \phi(\nu)] d\nu \quad (5.1.10)$$

因而与 $u^r(t)$ 相关联的解析信号可写成:

$$\begin{aligned} u(t) &= u^r(t) + u^i(t) = 2 \int_0^{+\infty} A(\nu) e^{i[2\pi\nu t + \phi]} d\nu \\ &= 2 \int_0^{+\infty} U^r(\nu) e^{i2\pi\nu t} d\nu \\ &= 2 \int_0^{+\infty} U(\nu) e^{i2\pi\nu t} d\nu \end{aligned} \quad (5.1.11)$$

这样, $U(\nu)$ 与 $U^r(\nu)$ 的关系可以表示成:

$$U(\nu) = \begin{cases} 2U^r(\nu) & \nu > 0 \\ 0 & \nu < 0 \end{cases} \quad (5.1.12)$$

式 (5.1.11) 式表明, 去掉实函数 $u^r(t)$ 的所有负频率分量, 并把正频率分量的幅值加倍后叠加起来, 就得

到了解析信号 $u(t)$ 。反之，实函数 $u^r(t)$ 可由它的解析信号 $u(t)$ 唯一确定：

$$u^r(t) = \operatorname{Re}[u(t)] \quad (5.1.13)$$

同样，实函数 $u^i(t)$ 可由它的解析信号 $u(t)$ 唯一确定，因为从 $u^r(t)$ 中把第一个傅里叶分量的相位变化 $\frac{\pi}{2}$ 后，就可得到 $u^i(t)$ 。因此，积分式 (5.1.9) 和 (5.1.10) 称为同源的傅里叶积分，亦称为相缩合的函数 (Associated function)。

在上面构造解析信号时，我们只考虑去掉 $u^r(t)$ 的负频分量而将其正频率分量加倍。那么 $u^r(t)$ 的零频分量将如何处理呢？当 $u^r(t)$ 包含有常数项时，就属于这种情况。这时频域里对应于在 $\nu = 0$ 处的一个 δ 函数，这在构造解析信号时应该保留，即要求：

$$U(\nu) = \begin{cases} 2U^r(\nu) & \nu > 0 \\ U^r(\nu) & \nu = 0 \\ 0 & \nu < 0 \end{cases} \quad (5.1.14)$$

若用符号函数表示：

$$\operatorname{sgn}(\nu) = \begin{cases} 1 & \nu > 0 \\ 0 & \nu = 0 \\ -1 & \nu < 0 \end{cases} \quad (5.1.15)$$

则 (5.1.14) 式可写为

$$U(\nu) = [1 + \operatorname{sgn}(\nu)]U^r(\nu) \quad (5.1.16)$$

因此，

$$u(t) = \int_{-\infty}^{\infty} [1 + \operatorname{sgn}(\nu)]U^r(\nu)e^{i2\pi\nu t} d\nu \quad (5.1.17)$$

由此可见，在构造解析信号时，我们应该去掉 $u^r(t)$ 的负频分量，保留零频分量，加倍正频分量。

现在我们从 (5.1.16) 式出发，讨论解析信号的另一个表示方法。将 (5.1.16) 式两边作傅里叶逆变换：

$$\begin{aligned} u(t) &= F^{-1}\{U(\nu)\} \\ &= F^{-1}\{U^r(\nu)\} + F^{-1}\{\operatorname{sgn}(\nu)U^r(\nu)\} \\ &= u^r(t) + F^{-1}\{\operatorname{sgn}(\nu)\} \otimes F^{-1}\{U^r(\nu)\} \quad (5.1.18) \\ &= u^r(t) + \frac{i}{\pi t} \otimes u^r(t) \\ &= u^r(t) - \frac{i}{\pi} \int_{-\infty}^{\infty} \frac{u^r(\alpha)}{\alpha - t} d\alpha \end{aligned}$$

上式中的被积函数在 $\alpha = t$ 时有一奇点，式中的积分要取柯西积分主值，即：

$$\int_{-\infty}^{\infty} \frac{u^r(\alpha)}{\alpha - t} d\alpha = \lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{t-\varepsilon} \frac{u^r(\alpha)}{\alpha - t} d\alpha + \lim_{\varepsilon \rightarrow 0} \int_{t+\varepsilon}^{\infty} \frac{u^r(\alpha)}{\alpha - t} d\alpha \quad (5.1.19)$$

上式所表示的积分称为希尔伯特 (Hilbert) 变换。函数 $u^r(t)$ 的希尔伯特变换用 $H\{u^r(t)\}$ 来表示。由 (5.1.18) 式和 (5.1.19) 可以看出，解析信号 $u(t)$ 的虚部 $u^i(t)$ 不是任意的，而是实信号 $u^r(t)$ 的希尔伯特变换，即：

$$u^i(t) = H\{u^r(t)\} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{u^r(\alpha)}{\alpha - t} d\alpha \quad (5.1.20)$$

根据解析信号的这一性质我们又得到了一个由实信号构造解析信号的方法：给定一个实信号 $u^r(t)$ ，对它实行希尔伯特变换而得到 $u^i(t)$ ，则所求的解析信号为：

$$u(t) = u^r(t) - iu^i(t)$$

由希尔伯特变换式 (5.1.20) 式可以看出，函数 $u^r(t)$ 的希尔伯特变换，可以持做函数 $u^r(t)$ 与 $-\frac{1}{\pi t}$ 的卷积，即：

$$u^i(t) = H\{u^r(t)\} = u^r(t) * \left(-\frac{1}{\pi t}\right) \quad (5.1.21)$$

换言之，希尔伯特变换可以看作是一个线性平移不变系统，该系统的脉冲响应为：

$$h(t) = -\frac{1}{\pi t} \quad (5.1.22)$$

于是 (5.1.18) 式可以写成

$$u(t) = \left[\delta(t) + \frac{i}{\pi t}\right] * u^r(t) \quad (5.1.23)$$

与脉冲响应 $h(t)$ 相应的传递函数为：

$$H(\nu) = F\left\{-\frac{1}{\pi t}\right\} = i \operatorname{sgn}(\nu) \quad (5.1.24)$$

若设解析信号的虚部 $u^i(t)$ 的频谱 $U^i(\nu)$ ，则

$$U^i(\nu) = i \operatorname{sgn}(\nu) U^r(\nu) \quad (5.1.25)$$

-----* * * * *-----

§ 5.2 互相干函数

从微观角度来说，一个发光光源包含有大量微观的辐射基元，光场是由这些辐射基元叠加而成的。因此，我们难以精确地描述这样的光场，而只能作为随机过程来讨论其统计性质。要对这种随机过程做完备的讨论，理论是可能的，但实际上却因过于复杂而难以实现。由于大多数情况下，我们无需知道光波的完整统计模型，只要知道某些阶的矩即可。如，只要用二阶矩就可以研究光场的相干性。

为了讨论一个具有有限带宽和有限大小的光源发出的光场的相干性问题，就是要确定光场中两上不同点，某一相对时间延迟时相关性。我们可以把光场中的两点看作次波源，考察它们发出的两束光波在空间另一点 Q 的干涉现象。时间变量可以通过光程差而得到。在时空坐标中，研究两个时空点的相关性，转化为在空间坐标系中研究三个点的相关性。

§ 5.2.1 互相干函数

光场的相干性可以用相干度 (Degree of coherence) 来度量。为此，首先定义互相干函数 (Mutual coherence function)。如图 5.2.1 所示，这时需要假定光源 S 具有一定大小并发出多色光。光照射到不透明屏上的两个小孔 P_1 和 P_2 到观察点 Q 的距离分别为 r_1 和 r_2 ， t 时刻 P_1 和 P_2 点的光振动解析信号为： $u(P_1, t)$ 和 $u(P_2, t)$ ，则两个小孔 P_1 和 P_2 实际上是把从光源 S 来的光衍射到观察屏上，它们形成了两个新光场。

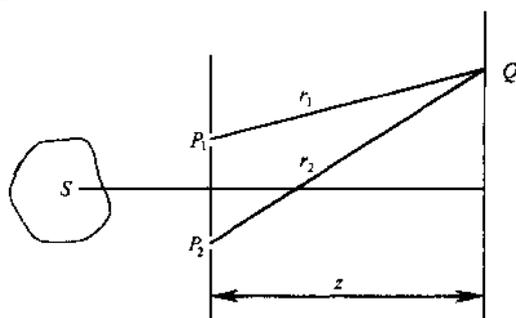


图 5.2.1 扩展光源的杨氏干涉实验

按照前面的讨论，这两个光场在观察屏上 Q 点叠加在一起后的合成光场和光强可表示为：

$$u(Q, t) = u(P_1, t) + u(P_2, t) = K_1 u(P_1, t - t_1) + u(P_2, t - t_2) \quad (5.2.1)$$

式中， $t_1 = \frac{r_1}{c}, t_2 = \frac{r_2}{c}$ ， c 是真空中光速， K_1 和 K_2 称为传播因子，它们分别与 r_1 和 r_2 成反比，与小孔的大小及光路的几何布局有关。对于窄带光，小孔足够小时，与根据惠更斯-菲涅耳原理的数学表达式，可知 K_1 和 K_2 是一个纯虚数。

由于探测器相对光频来说，是慢响应的，因而在 Q 探测到的光强是一个时间平均值：

$$I(Q) = \langle u(Q, t) u^*(Q, t) \rangle \quad (5.2.2)$$

式中，尖括号表示对时间求平均，即：

$$\langle f(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t) dt$$

将上式代入(5.2.2)式，得到：

$$\begin{aligned}
 I(Q) = & K_1^2 \langle u(P_1, t-t_1)u^*(P_1, t-t_1) \rangle + K_2^2 \langle u(P_2, t-t_2)u^*(P_2, t-t_2) \rangle \\
 & + K_1 K_2 \langle u(P_1, t-t_1)u^*(P_2, t-t_2) \rangle + K_1 K_2 \langle u^*(P_1, t-t_1)u(P_2, t-t_2) \rangle
 \end{aligned} \quad (5.2.3)$$

如果光场是平稳的，也就是说，其统计性质不随时间改变，互相关函数只与时间差： $\tau = t_2 - t_1 = \frac{r_2 - r_1}{c}$ 有关。若光场还是各态经历的，则时间互相关函数等于统计互相关函数。因此得出：

$$\langle u(P_1, t-t_1)u^*(P_2, t-t_2) \rangle = \langle u^*(P_1, t-t_1)u(P_2, t-t_2) \rangle = \Gamma_{12}(\tau) \quad (5.2.4)$$

式中 $\Gamma_{12}(\tau)$ 称为光场的互相干函数，显然：

$$\langle u^*(P_1, t-t_1)u(P_2, t-t_2) \rangle = \langle u(P_1, t-t_1)u(P_2, t-t_2) \rangle = \Gamma_{12}^*(\tau) \quad (5.2.5)$$

当 P_1 和 P_2 点重合时，则为自相干函数，定义为：

$$\langle u(P_1, t+\tau)u^*(P_1, t) \rangle = \Gamma_{11}(\tau) \quad (5.2.6)$$

$$\langle u(P_2, t+\tau)u^*(P_2, t) \rangle = \Gamma_{22}(\tau) \quad (5.2.7)$$

由于自相干函数只涉及一个空间点，仅仅是时间差 τ 的函数，在不致混淆的情况下，可以不再写下标，而记为 $\Gamma(\tau)$ ，当 $\tau = 0$ 时，便有：

$$\langle u(P_1, t)u^*(P_1, t) \rangle = \Gamma_{11}(0) = I_1 \quad (5.2.8)$$

$$\langle u(P_2, t)u^*(P_2, t) \rangle = \Gamma_{22}(0) = I_2 \quad (5.2.9)$$

I_1 和 I_2 分别是 P_1 和 P_2 点的光强，单孔 P_1 和 P_2 在 Q 点产生光强为：

$$I_1(Q) = K_1^2 \Gamma_{11}(0) = K_1^2 I_1 \quad (5.2.10)$$

$$I_2(Q) = K_2^2 \Gamma_{22}(0) = K_2^2 I_2 \quad (5.2.11)$$

这样，式(5.2.3)式可以写成：

$$\begin{aligned}
 I(Q) = & I_1(Q) + I_2(Q) + K_1 K_2 [\Gamma_{12}(\tau) + \Gamma_{12}^*(\tau)] \\
 = & I_1(Q) + I_2(Q) + K_1 K_2 \text{Re}[\Gamma_{12}(\tau)]
 \end{aligned} \quad (5.2.12)$$

为了讨论方便，通常将互相干函数写成归一化方便：

$$\gamma_{12}(\tau) = \frac{\Gamma_{12}(\tau)}{\sqrt{\Gamma_{11}(0)\Gamma_{22}(0)}} = \frac{\Gamma_{12}(\tau)}{\sqrt{I_1 I_2}} \quad (5.2.13)$$

$\gamma_{12}(\tau)$ 称为光场 $u(P_1, t)$ 和 $u(P_2, t)$ 的复相干度(Complex degree of coherence)或相关度(Correlativity)。这样，

(5.2.12)式最终表示为：

$$I(Q) = I_1(Q) + I_2(Q) + 2\sqrt{I_1(Q)I_2(Q)} \text{Re}[\gamma_{12}(\tau)] \quad (5.2.14)$$

上式正是平稳光场的普遍干涉定律。利用许瓦兹不等式易于证明：

$$|\Gamma_{12}(\tau)| \leq \sqrt{\Gamma_{11}(0)\Gamma_{22}(0)} \quad (5.2.15)$$

再结合(5.2.13)，便可得到

$$0 \leq |\gamma_{12}(\tau)| \leq 1 \quad (5.2.16)$$

从式(5.2.14)出发，可将 $\gamma_{12}(\tau)$ 与 Q 点的干涉条纹可见度联系起来。条纹可见度是针对正弦型条纹而言的，对于平均频率为 $\bar{\nu}$ 的窄带光，可将互相干函数和复相干度分别表示如下：

$$\Gamma_{12}(\tau) = |\Gamma_{12}(\tau)| e^{i[2\pi\bar{\nu}\tau + \alpha_{12}(\tau)]} \quad (5.2.17)$$

$$\gamma_{12}(\tau) = |\gamma_{12}(\tau)| e^{i[2\pi\bar{\nu}\tau + \alpha_{12}(\tau)]} \quad (5.2.18)$$

于是，式(5.2.14)可写为：

$$I(Q) = I_1(Q) + I_2(Q) + 2\sqrt{I_1(Q)I_2(Q)} |\gamma_{12}(\tau)| \cos[2\pi\bar{\nu}\tau + \alpha_{12}(\tau)] \quad (5.2.19)$$

记：

$$\delta = 2\pi\bar{\nu}\tau = \frac{2\pi}{\lambda}(r_2 - r_1)$$

式中， $\bar{\lambda}$ 为平均波长，上式为光波从小孔 P_1 和 P_2 到达 Q 点的相位差，与光源性质无关。而 $\alpha_{12}(\tau)$ 为光波在 P_1 和 P_2 点的相位差，也是与光源的性质无关的。当 $|\gamma_{12}(\tau)|$ 取最大值 1 时， Q 点的光强与频率为 $\bar{\nu}$ 的单色光波在该点叠加所产生的干涉结果相同， P_1 和 P_2 点的光振动是相干的。当 $|\gamma_{12}(\tau)|$ 取最小值 0 时， Q 点的光强为两光束光波在 Q 点产生的光强的简单相加，因此， P_1 和 P_2 点的光振动是不相干的。当 $0 < |\gamma_{12}(\tau)| < 1$ 时， P_1 和 P_2 点的光振动就是部分相干的。

光场的相干性质，可通过实验测定干涉条纹的清晰度或可见度来确定。对于正弦型主，迈克逊定义的干涉条纹可见度为：

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \quad (5.2.21)$$

式中， I_{\max} 和 I_{\min} 分别是 Q 点附近干涉条纹的极大和极小强度值。由公式(5.2.20)可得：

$$I_{\max} = I_1(Q) + I_2(Q) + 2\sqrt{I_1(Q)I_2(Q)}$$

$$I_{\min} = I_1(Q) + I_2(Q) - 2\sqrt{I_1(Q)I_2(Q)}$$

于是，

$$V = \frac{2\sqrt{I_1(Q)I_2(Q)}}{I_1(Q) + I_2(Q)} |\gamma_{12}(\tau)| \quad (5.2.22)$$

上式表明，只要测出两光束在 Q 点产生的光强及干涉条纹的可见度，就可得到复相干度的模 $|\gamma_{12}(\tau)|$ 。从基础光学，我们知道，由 P_1 和 P_2 发出的单色光波在 Q 点形成的干涉条纹的可见度的表达式为：

$$V = \frac{2\sqrt{I_1(Q)I_2(Q)}}{I_1(Q) + I_2(Q)} \quad (5.2.23)$$

所以, $|\gamma_{12}(\tau)|$ 的物理意义为: 在 Q 点附近的干涉条纹的可见度达到了当 P_1 和 P_2 点完全相干时的多大程度。

$\gamma_{12}(\tau)$ 幅角 $\alpha_{12}(\tau) + \delta$ 的物理意义是: 第二项是光波从小孔 P_1 和 P_2 到达 Q 点的相位延时。而第一项 $\alpha_{12}(\tau)$ 为光波在 P_1 和 P_2 点的相位差, 称为有效相位延迟。

根据杨氏干涉实验, 相位延迟 2π , 相当于干涉图在平行于 $P_1 P_2$ 方向上移动 $\frac{\bar{\lambda}z}{d}$ 。因此, 相对于当 P_1 和 P_2 处单色及同相位照射所形成的干涉条纹, 窄带光条纹在平行于双孔连线方向上有一位移量 Δ_1 , 它与 $\alpha_{12}(\tau)$ 有如下关系:

$$\frac{\alpha_{12}(\tau)}{2\pi} = \frac{\Delta_1}{\lambda z / d}$$

即有:

$$\alpha_{12}(\tau) = \frac{2\pi d}{\lambda z} \Delta_1$$

于是两束窄带光的复相干度的相位可通过测量干涉条纹的位置来确定。

当两束光波在 Q 点的强度相等时, 复相干度的模就等于干涉条纹的可见度:

$$|\gamma_{12}(\tau)| = V \quad (5.2.25)$$

§ 5.2.2 互相干函数的谱表示

下面导出互相干函数的傅里叶变换, 定义光谱密度, 得到互相干函数的光谱表达式。首先引入截断函数的定义:

$$u_T(P_1, t) = \begin{cases} u(P_1, t) & |t| \leq T \\ 0 & |t| > T \end{cases} \quad (5.2.26)$$

即先讨论限于 $-T < t < T$ 区间。 $u_T(P_1, t)$ 是与 $u_T^r(P_1, t)$ 相应的解析信号。由(5.1.8)得到其傅里叶变换为:

$$u_T(P_1, t) = \int_0^\infty U_T(P_1, \nu) e^{i2\pi\nu t} d\nu \quad (5.2.27)$$

式中, $U_T(P_1, \nu) = 2U_T^r(P_1, \nu)$ 。同样有:

$$u_T(P_2, t) = \int_0^\infty U_T(P_2, \nu) e^{i2\pi\nu t} d\nu \quad (5.2.28)$$

于是, 互相干函数可以写为:

$$\begin{aligned} \Gamma_{12}(\tau) &= \langle u(P_1, t + \tau) u^*(P_2, t) \rangle \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-\infty}^{\infty} u(P_1, t + \tau) u^*(P_2, t) dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-\infty}^{\infty} dt \int_0^\infty \int_0^\infty d\nu d\nu' \times U_T(P_1, \nu) U_T^*(P_2, \nu') e^{i2\pi(\nu - \nu')t} e^{i2\pi\nu\tau} \end{aligned}$$

由于:

$$\int_{-\infty}^{\infty} e^{i2\pi(\nu - \nu')t} dt = \delta(\nu - \nu')$$

于是有:

$$\Gamma_{12}(\tau) = \int_0^{\infty} G_{12}(\nu) e^{i2\pi\nu\tau} d\nu \quad (5.2.29)$$

式中,

$$G_{12}(\nu) = \lim_{T \rightarrow \infty} \left[\frac{U_T(P_1, \nu) U_T^*(P_2, \nu)}{2T} \right] \quad (5.2.30)$$

称为互光谱密度 (Mutual spectrum-density)。

对于自相干函数, 类似有:

$$\Gamma(\tau) = \int_0^{\infty} G(\nu) e^{i2\pi\nu\tau} d\nu \quad (5.2.31)$$

式中, $G(\nu)$ 称为光场的功率谱密度函数。如果把 P_1 点随时间变化的光振动看作是频率不同的许多单色光振动的线性组合, 则频率为 ν 的单色振动对强度的分布正比于 $G(\nu)$, 所以 $G(\nu)$ 也就是光源的光谱分布, 其定义为:

$$G_{12}(\nu) = \lim_{T \rightarrow \infty} \left[\frac{U_T(P, \nu) U_T^*(P, \nu)}{2T} \right] = \lim_{T \rightarrow \infty} \left[\frac{|U_T(P, \nu)|^2}{2T} \right] \quad (5.2.32)$$

对于复相干度也有类似的关系:

$$\gamma_{12}(\tau) = \int_0^{\infty} \hat{G}_{12}(\nu) e^{i2\pi\nu\tau} d\nu \quad (5.2.33)$$

式中, 归一化的互谱密度为:

$$g_{12}(\nu) = \frac{G_{12}(\nu)}{\sqrt{\Gamma_{11}(0)\Gamma_{22}(0)}} \quad (5.2.34)$$

相应地有:

$$\gamma(\tau) = \int_0^{\infty} \hat{g}(\nu) e^{i2\pi\nu\tau} d\nu \quad (5.2.35)$$

式中:

$$\hat{g}(\nu) = \frac{G(\nu)}{\Gamma(0)} \quad (5.2.36)$$

称为归一化的功率谱密度函数, 显然有:

$$\int_0^{\infty} \hat{g}(\nu) d\nu = 1 \quad (5.2.37)$$

-----** ** ** ** **-----

§ 5.3 时间相干性

在杨氏干涉的中，我们知道空间相干性和时间相干性都起作用。

如果在图 5.2.1 中的初级光源 S 为一个位于轴上的、具有有限带宽的点光源，那么时间相干性将是主要的。这时 P_1 和 P_2 点处的光振动将相同，两点间的互相干函数将变成自相干函数，即：

$$\Gamma_{12}(\tau) = \Gamma_{11}(\tau) = \Gamma(\tau)$$

$$\Gamma_{12}(\tau) = \Gamma_{11}(\tau) = \Gamma(\tau)$$

如果 S 是一个窄带扩展光源，那么空间相干性效应将是主要的，这时 P_1 和 P_2 点的光振动是不相同的，干涉条纹将取决于 Γ_{12} ，通过考察中心条纹附近的区域($r_2 - r_1 = 0, \tau = 0$)可以决定 $\Gamma_{12}(0)$ 和 $\gamma_{12}(0)$ 。实际上 $\gamma_{12}(0)$ 是 P_1 和 P_2 点两个点在同一时刻的复相干度，即空间相干度。这里，我们先讨论时间相干性。

5.3.1 时间相干性

在理想的单色光场中，空间任一点 P 的振动，振幅不变而相位随时间做线性变化。在实际光源所产生的光场中，情况就不是这样了，场中 P 点的振幅和相位都在随机地涨落，其涨落速度基本上取决于光源的有效频谱宽度 $\Delta\nu$ ，只有当时间间隔 τ 比 $\frac{1}{\Delta\nu}$ 小得多时，振幅才会大体保持不变。在这样一个时间间隔内，任何两个分量的相对相位变化都比 2π 小得多，并且这些分量的叠加所代表的振动在这个时间间隔内的表现，就像平均频率为 $\bar{\nu}$ 的单色光波一样，由： $\tau_c = \frac{1}{\Delta\nu}$ ，所决定的时间称为相干时间，而： $l_c = c\tau_c$ ，称为相干长度。相干时间大体上就是我们可以预言光波在某一给定点 P 的相位的那段时间间隔，这就是时间相干性的意思。也就是说，如果 τ_c 大，光波就有高的时间相干性，反之则时间相干性小。

同一特性可从不同的角度来解释。如：我们设想在一窄带点光源的半径上有两个分开的点 P_1 和 P_2 ，若相干长度 l_c 比两点间距 r_{12} 大得多，那么单个波列可以伸展在这个区间上，于是 P_1 点的振动和 P_2 点的振动是高度相关的。反之若 r_{12} 比 l_c 大得多，那么在距离 r_{12} 内会排下许多个波列，每个波列的相位都是不相关的。在这种情况下，空间两点的振动在任何时刻都是彼此无关的。相关的程度有时换个说法叫做纵向相关性的大小。不论我们用相干时间 τ_c 或用相干长度 l_c 思维，这个效应都是由光波的有效带宽引起的。

通过考察迈克耳孙干涉仪中光波的干涉可以更精确地描述和定义时间相干性。在图 5.3.1 所示的迈克耳孙干涉仪中，由点光源 S 发出的光束通过分束器 B 分成两支，一支反射到可动反射镜 M_2 ，经从反射后又折回而到达探测器 D 上。另一支透过 B 和补偿板 C 射到固定反射镜 M_1 ，再经从反射后折回也到达探测器 D 。当然，实际系统可加入准直透镜和会聚透镜等光学元件以改善性能。探测器位于干涉场内，这样一来，入射到探测器上的光强取决于干涉仪两支光路中的光的干涉，补偿板 C 的作用是保证光在干涉仪的两支光路中通过玻璃的光程相同。

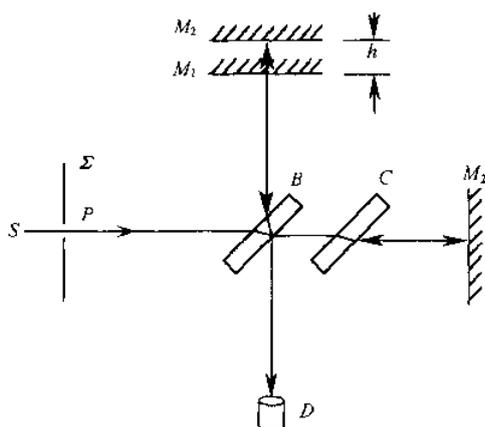


图 5.3.1 迈克耳逊干涉仪示意图

如果可动反射镜 M_1 从两支光路等光程的位置开始移动，其效果就是在两支光路个引入了一个相对时间延迟。反射镜每移动 $\frac{\lambda}{2}$ (光程为 λ)，入射到探测器上的光强变化一个周期。由于光源具有一定的光谱宽度，而不是理想单色光，如两支光路的光程差超过一定范围，干涉条纹的可见度下降到零。对于某些特种光源，当继续移动 M_1 加大程差，又能产生干涉条纹，只不过可见度已很小了。下面把光信号用解析信号构成的随机过程表示，对迈克耳孙干涉实验做出理论解释。

在迈克耳孙干涉仪中，光源 S 发出波场，我们研究空间 P 点场的时间相干性。为此，用一屏 Σ 挡住光场只露出 P 点。从 P 点发出的光波分别经干涉仪的两臂而到达 D 点相互干涉，用探测器 D 读出光强。设两臂之差为 h ，则两光路光程之差为 $2h$ 。这就相当于把 t 时刻 P 点的场与 $(t + 2h/c)$ 时刻 P 点的场二者进行了叠加。即同一点不同时刻的场发生干涉。移动 M_1 就相当于改变 $\tau = 2h/c$ ，于是可对任何 τ 进行测量。

我们把光源发出的光信号用解析信号构成的随机过程来表示。我们用 $u(t)$ 表示由 P 点发出的解析信号，由于 P 点位置固定，故可将 P 点的场写成只含时间变量的函数。 $u(t)$ 经分束器后通过两支光路而到达探测器 D 。在 D 处的两束光的解析信号为 $K_1u(t)$ 和 $K_2u(t + \tau)$ 。这里的 K_1 和 K_2 是两支光路的透过率所决定的实数， $\tau = 2h/c$ 是时间延迟。这样，我们知道，探测器上的合成解析信号为：

$$u_D(t) = K_1u(t) + K_2u(t + \tau) \quad (5.3.1)$$

探测器 D 只对照射到它上面的光强产生响应，是慢响应时间的，这样，对时间求平均，则入射到探测器上的光强可表示为：

$$\begin{aligned} I_D &= \langle u_D(t)u_D^*(t) \rangle \\ &= \langle |K_1u(t) + K_2u(t + \tau)|^2 \rangle \\ &= K_1^2 \langle |u(t)|^2 \rangle + K_2^2 \langle |u(t + \tau)|^2 \rangle \\ &\quad + K_1K_2 \langle u(t + \tau)u^*(t) \rangle + K_1K_2 \langle u^*(t + \tau)u(t) \rangle \end{aligned} \quad (5.3.2)$$

假设光场是平稳的和各态历经的，统计量的平均与时间原点无关，时间平均与统计平均相

同。则由 P 点发出的光强可表示为：

$$I_0 = \langle |u(t)|^2 \rangle = \langle |u(t+\tau)|^2 \rangle \quad (5.3.3)$$

解析信号 $u(t)$ 的自相关函数 $\Gamma(\tau)$ 称为光扰动的自相干函数，在各态历经条件假设下可用时间平均代替，即：

$$\Gamma(\tau) = \langle u(t+\tau)u^*(t) \rangle \quad (5.3.4)$$

将(5.3.4)式和(5.3.3)式代人(5.3.2)式得：

$$I_D = (K_1^2 + K_2^2)I_0 + 2K_1K_2 \operatorname{Re}\{\Gamma(\tau)\} \quad (5.3.5)$$

由于：

$$I_0 = \Gamma(0)$$

我们用它来归一化，于是得复相干度为：

$$\gamma(\tau) = \frac{\Gamma(\tau)}{\Gamma(0)} \quad (5.3.6)$$

显然：

$$\gamma(0) = 1, 0 \leq |\gamma(\tau)| \leq 1 \quad (5.3.7)$$

这样，(5.3.5)式可写成：

$$I_D = (K_1^2 + K_2^2)I_0 + 2K_1K_2I_0 \operatorname{Re}\{\gamma(\tau)\} \quad (5.3.8)$$

若认为两路光的透过率相等，即 $K_1 = K_2 = K$ ，更进一步若不考虑吸收，则 $K = 1$ ，于是(5.3.5)式和(5.3.8)式分别简化成：

$$I_D(\tau) = 2I_0 + 2 \operatorname{Re}\{\Gamma(\tau)\} \quad (5.3.9)$$

$$I_D(\tau) = 2I_0\{1 + \operatorname{Re}\{\gamma(\tau)\}\} \quad (5.3.10)$$

由于 $u(t)$ 是解析信号，于是它的自相关函数 $\Gamma(\tau)$ 也是具有单边频谱的解析信号。当然，复相干度 $\gamma(\tau)$ 也是一个解析信号，也具有单边频谱。我们曾用 $\hat{G}(\nu)$ 表示与其对应的归一化功率谱密度，它与 $\gamma(\tau)$ 构成一对傅里叶变换。即：

$$\gamma(\tau) = \int_0^\infty \hat{G}(\nu)e^{i2\pi\nu\tau} d\nu \quad (5.3.11)$$

对于窄带光， $\gamma(\tau)$ 可以写成：

$$\gamma(\tau) = \tilde{\gamma}(0)e^{i[2\pi\bar{\nu}\tau + \alpha(\tau)]} \quad (5.3.12)$$

式中， $\bar{\nu}$ 是光波的中心频率。于是(5.3.10)式写成：

$$I_D(\tau) = 2I_0\{1 + \tilde{\gamma}(\tau)\cos[2\pi\bar{\nu}\tau + \alpha(\tau)]\} \quad (5.3.13)$$

由(5.3.5)式所表示的干涉图的条纹可见度为：

$$V = \frac{2K_1K_2}{K_1^2 + K_2^2} \tilde{\gamma}(\tau) \quad (5.3.14)$$

当两支光路透射系数相等时，可见度为：

$$V = \tilde{\gamma}(\tau) \quad (5.3.15)$$

三种常见的谱线形状为：高斯型、洛伦兹型和矩形功率变形，其表达式分别为：

$$\hat{G}(\nu) = \frac{2\sqrt{\ln 2}}{\sqrt{\pi}\Delta\nu} e^{-\left(\frac{2\ln 2(\nu-\bar{\nu})}{\Delta\nu}\right)^2} \quad (5.3.16)$$

$$\hat{G}(\nu) = \frac{2\pi\Delta\nu}{[2\pi(\nu-\bar{\nu})]^2 + (\pi\Delta\nu)^2} \quad (5.3.17)$$

$$\hat{G}(\nu) = \frac{1}{\Delta\nu} \text{rect}\left(\frac{\nu-\bar{\nu}}{\Delta\nu}\right) \quad (5.3.18)$$

其对应的复相干度分别为：

$$\gamma(\tau) = e^{-\left(\frac{\pi\Delta\nu\tau}{2\ln 2}\right)^2} e^{j2\pi\bar{\nu}\tau} \quad (5.3.19)$$

与(5.3.12)相比，可得：

$$\tilde{\gamma}(\tau) = e^{-\pi|\tau|}, \alpha(\tau) = 0 \quad (5.3.20)$$

$$\gamma(\tau) = e^{-\pi\Delta\nu|\tau|} e^{j2\pi\bar{\nu}\tau} \quad (5.3.21)$$

$$\tilde{\gamma}(\tau) = e^{-\pi\Delta\nu|\tau|}, \alpha(\tau) = 0 \quad (5.3.22)$$

$$\gamma(\tau) = \text{sinc}(\Delta\nu\tau) e^{j2\pi\bar{\nu}\tau} \quad (5.3.23)$$

$$\tilde{\gamma}(\tau) = |\text{sinc}(\Delta\nu\tau)|,$$

$$\alpha(\tau) = \begin{cases} 0 & 2n < \Delta\nu\tau < 2n+1 \\ \pi & 2n+1 < \Delta\nu\tau < 2n+2 \end{cases} \quad (5.3.24)$$

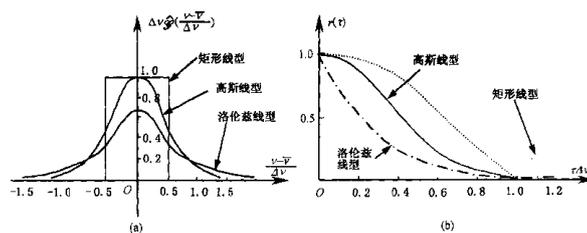


图 5.3.2 三种典型光源的功率谱和复相干度

§ 5.3.2 相干时间

在许多应用中，希望对相干时间有一个精确定义。可以借用复相干度来定义相干时间，但是借助复相干度定义相干时间的方法很多。按照曼德尔 (L. Mandel) 的意见，将相干时间定义为：

$$\tau_c = \int_{-\infty}^{\infty} |\gamma(\tau)| d\tau \quad (5.3.25)$$

如果这个定义有意义，则要求 τ_c 具有与 $\frac{1}{\Delta\nu}$ 有相同的数量级。对于前面给出的三种线性，分别有：

高斯型：

$$\tau_c = \sqrt{\frac{2\ln 2}{\pi}} \frac{1}{\Delta\nu} \approx 0.664 \frac{1}{\Delta\nu} \quad (5.3.25)$$

洛伦兹线型：

$$\tau_c = \frac{1}{\pi\Delta\nu} \approx 0.318 \frac{1}{\Delta\nu} \quad (5.3.26)$$

矩形线型：

$$\tau_c = \frac{1}{\Delta\nu} \quad (5.3.27)$$

但是在一般的应用中常用(5.3.27)式的结果来估计相干时间。

§ 5.3.3 傅里叶变换光谱技术

如果光波的功率谱密度已知，由迈克耳孙干涉仪观察到的干涉图的特征就可完全确定。利用干涉图和功率谱密度之间的这一关系，通过测量干涉图来确定未知的入射光的功率谱密度。这个原理就是傅里叶变换光谱术的基础。

由傅里叶变换光谱术得到光谱，首先必须测量干涉图。通常是在干涉仪的控制下，可动反射镜从零程差的位置移到大程差的范围内，将光强作为这个过程的时间函数进行测量。同时，把所得到的干涉图数字化，这可利用快速傅里叶变换技术，由数字傅里叶变换得到光谱。

在迈克耳孙干涉实验中， D 点光强随反射镜从移动而发生变化，为简单起见，在(5.3.5)式中令：

$K_1 = K_2 = K = 1$ 。于是有：

$$I_D = 2I_0 + \Gamma(\tau) + \Gamma^*(\tau) \quad (5.3.28)$$

即，

$$\begin{aligned} I(\tau) &= I_D(\tau) - 2I_0 = \Gamma(\tau) + \Gamma^*(\tau) \\ &= \int_0^\infty G(\nu)e^{i2\pi\nu\tau} d\nu + \int_0^\infty G(\nu)e^{-i2\pi\nu\tau} d\nu \\ &= 2\int_0^\infty G(\nu)\cos(2\pi\nu\tau) d\nu \end{aligned} \quad (5.3.29)$$

其中利用了功率若为实函数的性质。由于在迈克耳孙干涉仪中将 τ 换为 $-\tau$ 时。干涉强度不变，所以 $I_D(\tau)$ 是 τ 的偶函数，因而 $I(\tau)$ 也是的 τ 偶函数。实偶函数的傅里叶变换也是实偶函数，由实偶函数的余弦变换性质，由公式(5.3.29)得出：

$$G(\nu) = 2\int_0^\infty I(\tau)\cos(2\pi\nu\tau) d\tau \quad (5.3.30)$$

用迈克耳孙干涉仪记录下 $I(\tau)$ ，再借助于傅里叶余弦变换就可获得光源的光谱分布，这种方法称为傅里叶变换光谱术。当然，傅里叶光谱术不是一种直接的方法。根据色散原理，用一个普通的棱镜光谱仪或光栅光谱仪也可直接得到光源的光谱。但这种方法有一个缺点，即在每一波长位置上只能接收光源总能量的极

§ 5.4 空间相干性

我们知道，在杨氏干涉实验的中，时间相干性和空间相干性都起作用。如果在图 5.2.1 中， S 是一个窄带扩展光源，那么空间相干效应将是主要的。这时 P_1 和 P_2 点的光振动是不相同的，干涉条纹将取决于 $\Gamma_{12}(\tau)$ ，通过考察中心条纹附近区域($r_2 - r_1 = 0, \tau = 0$)可以决定 $\Gamma_{12}(0)$ 和 $\gamma_{12}(0)$ 。实际上 $\gamma_{12}(0)$ 是 P_1 和 P_2 两个点在同一时刻的复相干度。这样，我们知道，在零程差位置形成干涉条纹的能力反映了空间相干效应。

当图 5.2.1 中的 Q 点移向 O 点时，在公式(5.2.4)和(5.2.13)式中 $\tau = 0$ ，可得：

$$\Gamma_{12}(0) = \langle u(P_1, t)u^*(P_2, t) \rangle \quad (5.4.1)$$

$$\gamma_{12}(0) = \frac{\Gamma_{12}(0)}{\sqrt{\Gamma_{11}(0)\Gamma_{22}(0)}} = \frac{\Gamma_{12}(0)}{\sqrt{I_1 I_2}} \quad (5.4.2)$$

我们就把 $\Gamma_{12}(0)$ 称为互相干函数， $\gamma_{12}(0)$ 称为复空间相干度，它们描述在同一时刻 t ，光场中两点的空间相干性，它一般是复数，可写成：

$$\Gamma_{12}(0) = \tilde{\Gamma}_{12}(0)e^{j\alpha_{12}(0)} \quad (5.4.3)$$

$$\gamma_{12}(0) = \tilde{\gamma}_{12}(0)e^{j\alpha_{12}(0)} \quad (5.4.4)$$

这时， Q 点的光强为：

$$I(Q) = I_1(Q) + I_2(Q) + 2\sqrt{I_1(Q)I_2(Q)} \cos[\alpha_{12}(0)] \quad (5.4.5)$$

-----*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-----

§ 5.5 在准单色条件下的干涉

为了在杨氏干涉实验中将入射到观察点 Q 的场用针孔上的场(适当延迟)的加权和来简单表示, 必须假定光是窄带的。其次假定光的相干长度远大于所涉及范围内的最大光程差。用数学表示窄带和小程差的条件就是:

$$\Delta\nu \ll \bar{\nu}, \quad \frac{r_2 - r_1}{c} \ll \tau_c \quad (5.5.1)$$

满足条件(5.5.1)式的光称为准单色光。其中第二个条件的实质, 是认为在感兴趣的观察范围内, 条纹的可见度是一个常数, 这样, 我们就可简化互相干函数和复相干度的形式。

我们知道, 两解析信号的互相关函数仍为解析信号, 具有单边频谱, 这就是(5.2.28)式所表示的关系, 即

$$\Gamma_{12}(\tau) = \int_0^{\infty} G_{12}(\nu) e^{i2\pi\nu\tau} d\nu = e^{i2\pi\bar{\nu}\tau} \int_0^{\infty} G_{12}(\nu) e^{i2\pi(\nu-\bar{\nu})\tau} d\nu \quad (5.5.2)$$

考虑到第一个条件, 在满足 $|\nu - \bar{\nu}| \ll \bar{\nu}$ 的频率范围内, $G_{12}(\nu)$ 才有明显不为零的值, 或者说上式中对积分的主要贡献来自很窄的频率范围 $\Delta\nu$ 内。这个很窄的范围 $\Delta\nu$ 决定了相干时间, 即 $\tau_c = 1/\Delta\nu$ 。

根据准单色的第二个条件, 必有: $\tau \ll 1/\Delta\nu, \Delta\nu\tau \ll 1$ 。因此在这两个条件下, 公式(5.5.2)积分中的指数函数近似等于 1, 因而有:

$$\Gamma_{12}(\tau) = e^{i2\pi\bar{\nu}\tau} \int_0^{\infty} G_{12}(\nu) d\nu = e^{i2\pi\bar{\nu}\tau} \Gamma_{12}(0) \quad (5.5.3)$$

由(5.4.3), 我们有:

$$\Gamma_{12}(0) = \tilde{\Gamma}_{12}(0) e^{i\alpha_{12}(0)} \quad (5.5.4)$$

$$\text{令: } J_{12}(0) = J_{12}(0), \beta_{12} = \alpha_{12}(0) \quad (5.5.5)$$

则有:

$$J_{12}(0) = \tilde{J}_{12}(0) e^{i\beta_{12}} \quad (5.5.6)$$

们称 J_{12} 为 P_1 和 P_2 点的互强度, 表示 P_1 和 P_2 两点在相对时间延迟 $\tau = 0$ 的情况下的互相关。于是(5.5.3)式所表示 $\Gamma_{12}(\tau)$ 的可写成:

$$\Gamma_{12}(\tau) = J_{12} e^{i2\pi\bar{\nu}\tau} = \tilde{J}_{12} e^{i(2\pi\bar{\nu}\tau + \beta_{12})} \quad (5.5.7)$$

$$\gamma_{12}(\tau) = \frac{\Gamma_{12}(\tau)}{\sqrt{\Gamma_{11}(0)\Gamma_{12}(0)}} = \frac{\Gamma_{12}(\tau)}{\sqrt{I_1 I_2}} = \frac{\Gamma_{12}(0) e^{i2\pi\bar{\nu}\tau}}{\sqrt{I_1 I_2}} = \frac{J_{12}}{\sqrt{I_1 I_2}} e^{i2\pi\bar{\nu}\tau} = \gamma_{12}(0) e^{i2\pi\bar{\nu}\tau} = \mu_{12} e^{i2\pi\bar{\nu}\tau} \quad (5.5.8)$$

式中:

$$\mu_{12} = \gamma_{12}(0) = \tilde{\mu}_{12} e^{i\beta_{12}} \quad (5.5.9)$$

称为复相干系数。于是(5.5.8)式最后写成:

$$\gamma_{12}(\tau) = \tilde{\mu}_{12} e^{i(2\pi\bar{\nu}\tau + \beta_{12})} \quad (5.5.10)$$

显然 μ_{12} 满足:

$$0 \leq |\mu_{12}| \leq 1 \quad (5.5.11)$$

这就是说, 在准单色条件下互相干函数 $\Gamma_{12}(\tau)$ 和复相干度 $\gamma_{12}(\tau)$ 可以用(5.5.7)式和(5.5.10)式来表示。

辐射场的干涉定律(5.2.14)式变成:

$$I(Q) = I_1(Q) + I_2(Q) + 2\sqrt{I_1(Q)I_2(Q)}\tilde{\mu}_{12} \cos[\beta_{12} + 2\pi\bar{\nu}\tau] \quad (5.5.12)$$

式中 β_{12} 是与 τ 无关的量。如果 $I_1(Q)$ 和 $I_2(Q)$ 在观察区内近似不变, 在该区域干涉图样具有几乎恒定的可见度和相位。这时, 条纹可见度为:

$$V = \frac{2\sqrt{I_1(Q)I_2(Q)}}{I_1(Q) + I_2(Q)} \mu_{12} \quad (5.5.13)$$

若两支光路的光强相等, 则有:

$$V = \mu_{12} \quad (5.5.14)$$

由公式(5.5.12)可以看出, 准单色光场的特点似乎类似于频率为 $\bar{\nu}$ 的严格单色光场。区别在于准单色光场的干涉条纹的可见度和位置分别决定于复相干度的模和相位。

-----*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-*-----

§ 5.6 互相干的传播

当光波在空间传播时，其详细结构会发生变化。互相干函数的详细结构也以同样的方式在变化。在这个意义上说互相干函数在传播。在这两种情况下，传播的最基本物理原因是基于光波本身服从波动方程。我们这里只给出在惠更斯-菲涅耳原理基础上的解。在讨论互相干函数传播之前，引述一下非单色光场的衍射积分公式。

一个单色光波入射到一个无限大表面 Σ 上，如图 5.6.1 所示，我们希望借助于 Σ 上的光场表示出右边场中一点 Q 的复振幅。根据惠更斯-菲涅耳原理，可写出 Q 点的复振幅表达式：

$$u(Q) = \frac{1}{i\lambda} \iint_{\Sigma} u(P) K(\theta) \frac{e^{i2\pi r/\lambda}}{r} dS \quad (5.6.1)$$

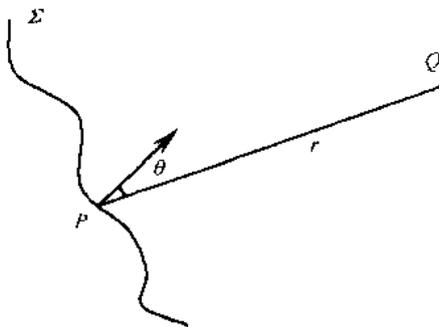


图 5.6.1 传播的空间几何关系

现在考虑入射到面 Σ 上的场是一个非单色光波，在 Σ 面上的光场分布为 $\tilde{u}(P, t)$ ，与其对应的解析信号为 $u(P, t)$ 。 Σ 面上的光场分布在其右边场中 Q 点处产生的光场用 $\tilde{u}(Q, t)$ 表示，其对应的解析信号为 $u(Q, t)$ ，可以证明 $u(Q, t)$ 可由 Σ 面上的场分布 $u(P, t)$ 表示出来，即：

$$u(Q, t) = \iint_{\Sigma} \frac{du(P, t - r/c)}{2\pi cr} K(\theta) dS \quad (5.6.2)$$

对于窄带光，这个表达式可以简化为：

$$u(Q, t) = \iint_{\Sigma} \frac{u(P, t - r/c)}{j\bar{\lambda}r} K(\theta) dS \quad (5.6.3)$$

式中 $\bar{\lambda}$ 是中心波长。

现在讨论互相干函数的传播问题。如图 5.6.2 所示，若有任意相干性的光波从左向右传播，已知在 Σ_1 面上的互相干函数为 $\Gamma(P_1, P_2, \tau)$ 。我们希望找到 Σ_2 面上的互相干函数 $\Gamma(Q_1, Q_2, \tau)$ ，也就是说，我们的目的是已知针孔 P_1 和 P_2 的杨氏干涉结果时，预测针孔 Q_1 和 Q_2 的杨氏干涉实验结果。

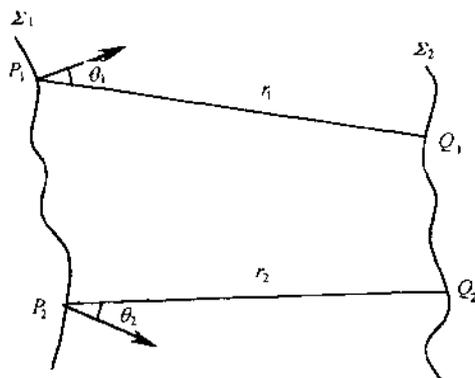


图 5.6.2 互相干传播的几何关系

我们这里限于讨论窄带情形， Σ_2 面上的互相干函数定义为：

$$\Gamma(Q_1, Q_2, \tau) = \langle u(Q_1, t + \tau) u^*(Q_2, t) \rangle \quad (5.6.4)$$

通过窄带光传播规律的(5.6.3)式，便可以把 Σ_2 面上的光场同 Σ_1 面上的光场联系起来，于是有：

$$u(Q_1, t) = \iint_{\Sigma} \frac{u(P_1, t + \tau - r_1/c)}{j\bar{\lambda}r_1} K(\theta_1) dS_1 \quad (5.6.5)$$

$$u^*(Q_2, t) = \iint_{\Sigma} -\frac{u(P_2, t + \tau - r_2/c)}{j\bar{\lambda}r_2} K(\theta_2) dS_2 \quad (5.6.6)$$

式中， θ_1, θ_2 分别是 r_1, r_2 与该点处波面法线的夹角。将(5.6.5)和(5.6.6)代入(5.6.4)式，并交换积分和求平均的次序得：

$$\Gamma(Q_1, Q_2; \tau) = \iint_{\Sigma_1} \iint_{\Sigma_2} dS_1 dS_2 K(\theta_1) K(\theta_2) \times \frac{\langle u(P_1, t + \tau - \frac{r_1}{c}) u^*(P_2, t - \frac{r_2}{c}) \rangle}{\bar{\lambda}^2 r_1 r_2} \quad (5.6.7)$$

被积函数中的时间平均可借助于 Σ_1 面上的互相干函数来表示，这就得出了在窄带假设下互相干传播的基本定律：

$$\Gamma(Q_1, Q_2; \tau) = \iint_{\Sigma_1} \iint_{\Sigma_2} \Gamma\left(P_1, P_2; \tau + \frac{r_2 - r_1}{c}\right) \frac{K(\theta_1)}{\bar{\lambda}r_1} \frac{K(\theta_2)}{\bar{\lambda}r_2} dS_1 dS_2 \quad (5.6.8)$$

(5.6.8)式适用于窄带条件，现在对它作进一步的限制以适用于准单色条件。为此我们要求最大光程差远小于相干长度。在这个假设下，我们会找到相应的互强度的传播规律。当准单色条件被满足时，按定义 Σ_2 面上的互强度为：

§ 5.7 范西特-泽尼克定理

如图 5.7.1 所示, 当光场由 Σ_1 面传播到 Σ_2 面时, 面上任何一点 Q_1 或 Q_2 的光振动都是由 Σ_1 面上各点贡献叠加而成的。因此, 即使 Σ_1 面上的光场是非相干的, 在 Σ_2 面上的各点对 (Q_1, Q_2) 的光振动之间都存在一定的联系, 也就是有一定的相干性。作为近代光学中最重要的定理之一的范西特-泽尼克(Van Cittert-Zernike)定理, 就是讨论一种由准单色(空间)非相干光源照明面产生的光场的互强度。

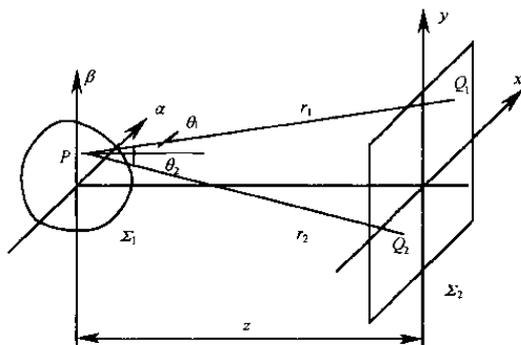


图 5.7.1 范西特-泽尼克定理的几何关系

§ 5.7.1 范西特-泽尼克定理

如图 5.7.1 所示, Σ_1 和 Σ_2 平面相互平行, 相距为 z 。 Σ_1 是一个准单色扩展光源, 由它发出的非相干光照射到 Σ_2 面上, 我们现在想求出 Σ_2 面上任意两点 Q_1 和 Q_2 的互强度和复相干系数。

扩展光源 Σ_1 上的互强度 $J(P_1, P_2)$ 和观察面 Σ_2 上的互强度 $J(Q_1, Q_2)$ 由下式联系:

$$J(Q_1, Q_2) = \iint_{\Sigma_1} \iint_{\Sigma_1} J(P_1, P_2) e^{\frac{j2\pi}{\lambda}(r_2 - r_1)} \frac{K(\theta_1)}{\lambda r_1} \frac{K(\theta_2)}{\lambda r_2} dS_1 dS_2 \quad (5.7.1)$$

对于空间非相干光源这种特殊情况, 两个不同点的光振动是统计无关的, 因而有:

$$J(P_1, P_2) = I(P_1) \delta(P_1 - P_2) \quad (5.7.2)$$

把它代入(5.7.1)中, 并利用 δ 函数的筛选性质, 使可得到观察屏幕上的互强度为:

$$J(Q_1, Q_2) = \frac{1}{\lambda^2} \iint_{\Sigma_1} I(P_1) e^{\frac{j2\pi}{\lambda}(r_2 - r_1)} \frac{K(\theta_1)}{\lambda r_1} \frac{K(\theta_2)}{\lambda r_2} dS \quad (5.7.3)$$

为了将(5.7.3)式进一步简化, 我们如下假设和近似:

- (1) 光源和观察区的线度与两者之间的距离 z 相比很小, 因此有: $r_1 r_2 = z^2$;
- (2) 只涉及小角度, 以致: $K(\theta_1) \approx K(\theta_2) \approx 1$ 。

于是在观察区的互强度取如下形式:

$$J(Q_1, Q_2) = \frac{1}{(\lambda z)^2} \iint_{\Sigma_1} I(P_1) e^{\frac{j2\pi}{\lambda}(r_2 - r_1)} dS \quad (5.7.4)$$

现在再对指数函数中的 r_1 和 r_2 引入傍轴近似:

$$r_2 = \sqrt{z^2 + (x_2 - \alpha)^2 + (y_2 - \beta)^2} \approx z + \frac{(x_2 - \alpha)^2 + (y_2 - \beta)^2}{2z}$$

$$r_1 = \sqrt{z^2 + (x_1 - \alpha)^2 + (y_1 - \beta)^2} \approx z + \frac{(x_1 - \alpha)^2 + (y_1 - \beta)^2}{2z}$$

由于 α, β 在有限的光源范围 Σ_1 之外时, $I(\alpha, \beta) = 0$, 于是范西特—泽尼克定理的最后形式为:

$$J(x_1, y_1; x_2, y_2) = \frac{e^{i\psi}}{(\lambda z)^2} \iint_{-\infty}^{\infty} I(\alpha, \beta) e^{-i\frac{2\pi}{\lambda z}(\Delta x \alpha + \Delta y \beta)} d\alpha d\beta \quad (5.7.5)$$

式中: $\Delta x = x_2 - x_1, \Delta y = y_2 - y_1$, 相位因子为:

$$\psi = \frac{\pi}{\lambda z} [(x_2^2 + y_2^2) - (x_1^2 + y_1^2)] = \frac{\pi}{\lambda z} (\rho_2^2 - \rho_1^2) \quad (5.7.6)$$

其中: ρ_1, ρ_2 分别是点 (x_1, y_1) 和 (x_2, y_2) 离光轴的距离。

把这一定理表示成归一化形式往往更方便, 为此先计算 Q_1 和 Q_2 点的强度。在公式(5.7.5)中令: $x_1 = x_2$,

$y_1 = y_2$ 便可得出 Q_1 点或 Q_2 点的强度表达式, 显然它们相等, 即:

$$I(x_1, y_1) = I(x_2, y_2) = \frac{1}{(\lambda z)^2} \iint_{-\infty}^{\infty} I(\alpha, \beta) d\alpha d\beta \quad (5.7.7)$$

于是:

$$\mu(x_1, y_1; x_2, y_2) = \frac{J(x_1, y_1; x_2, y_2)}{\sqrt{I(x_1, y_1)I(x_2, y_2)}} = \frac{e^{i\psi} \iint_{-\infty}^{\infty} I(\alpha, \beta) e^{-i\frac{2\pi}{\lambda z}(\Delta x \alpha + \Delta y \beta)} d\alpha d\beta}{\iint_{-\infty}^{\infty} I(\alpha, \beta) d\alpha d\beta} \quad (5.7.8)$$

(5.7.8)式给出了一个十分重要的结论: 即当光源本身的线度以及观察区域的线度都比二者间的距离小得多时, 观察区域上的复相干系数正比于光源强度分布的归一化傅里叶变换。

相位因子 $e^{i\psi}$ 并不影响复相干系数的模 $\tilde{\mu}(Q_1, Q_2)$ 。也就是说不影响我们判断 Q_1 和 Q_2 两点在杨氏干涉实验中产生干涉条纹的可见度。 $\mu(Q_1, Q_2)$ 只和观察平面上选定的 Q_1 和 Q_2 两点的坐标差 $(\Delta x, \Delta y)$ 有关。

此外, 当两点到光轴的距离相等时, $\psi = 0$; 或者当 $z \gg \frac{\pi(\rho_2^2 - \rho_1^2)}{\lambda}$ 时, $\psi \ll \frac{\pi}{2}$ 。这两种情况下, (5.7.8)

式中的相位因子 $e^{i\psi}$ 便可略去。

$\mu(Q_1, Q_2)$ 和 $I(\alpha, \beta)$ 之间存在着傅里叶变换关系。这种运算关系类似于夫琅禾费衍射。但是, 范西特—泽尼克定理在更宽的空间范围内成立, 因为我们只涉及了傍轴近似, 在衍射问题中对菲涅耳衍射和夫琅禾费衍射都适用。

§ 5.7.2 相干面积

从范西特—泽尼克定理出发可以导出准单色扩展光源相干性的量度, 即空间相干性。由于复相干系数的模只与 x, y 平面两点的坐标差 $\Delta x, \Delta y$ 有关, 由于我们可以定义一个相干面积 A_c :

$$A_c = \iint_{-\infty}^{\infty} |\mu(\Delta x, \Delta y)|^2 d\Delta x d\Delta y \quad (5.7.9)$$

它在性质上完全类似于(5.3.27)式所定义的相干时间 τ_c 。

我们可以证明，对于形状任意，面积为 A_s 的均匀非相干被单色光源，在离光源 z 处的相干面积是：

$$A_c = \frac{(\bar{\lambda}z)^2}{A_s} \approx \frac{(\bar{\lambda})^2}{\Omega_s} \quad (5.7.10)$$

式中 Ω_s 是光源对观察区原点所张的立体角。上式可以由范西特—泽尼克定理得以证明。

§ 5.7.3 均匀圆形光源

下面我们计算一个亮度均匀、非相干准单色、半径为 a 的圆盘形光源所产生的光场，作为应用范西特—泽尼克定理的一个例子。设光源的强度分布：

$$I(\alpha, \beta) = I_0 \text{circ} \left(\frac{\sqrt{\alpha^2 + \beta^2}}{a} \right) = I_0 \text{circ} \left(\frac{r}{a} \right) \quad (5.7.11)$$

为了求出 $J(\Delta x, \Delta y)$ ，我们采用极坐标，对强度分布作傅里叶—贝塞尔变换：

$$B \left\{ I_0 \text{circ} \left(\frac{r}{a} \right) \right\} = 2\pi I_0 \int_0^a r J_0(2\pi r \rho) dr$$

式中， ρ 为领域中的极坐标半径。为完成上式的积分，令 $r' = 2\pi r \rho$ ，并利用第一类零阶和一阶贝塞尔函数的积分关系：

$$\int_0^x \alpha J_0(\alpha) d\alpha = x J_1(x)$$

于是：

$$B \left\{ J_0 \text{circ} \left(\frac{r}{a} \right) \right\} = \frac{I_0}{2\pi \rho^2} \int_0^{2\pi a \rho} r' J_0(r') dr' = \pi a^2 \frac{2J_1(2\pi a \rho)}{2\pi a \rho} \quad (5.7.12)$$

由(5.7.5)式我们得：

$$J(\Delta x, \Delta y) = \frac{e^{i\psi}}{(\bar{\lambda}z)^2} \iint_{-\infty}^{\infty} d\alpha d\beta \cdot I(\alpha, \beta) e^{-i2\pi \left(\frac{\Delta x}{\lambda z} \alpha + \frac{\Delta y}{\lambda z} \beta \right)} = \frac{\pi a^2 I_0}{(\lambda z)^2} e^{i\psi} \left[\frac{2J_1(2\pi a \rho)}{2\pi a \rho} \right] \quad (5.7.13)$$

式中：

$$\rho = \sqrt{\left(\frac{\Delta x}{\lambda z} \right)^2 + \left(\frac{\Delta y}{\lambda z} \right)^2} = \frac{1}{\lambda z} \sqrt{(\Delta x)^2 + (\Delta y)^2} = \frac{s}{\lambda z} \quad (5.7.14)$$

式中： $s = \sqrt{(\Delta x)^2 + (\Delta y)^2}$ 是两点间的距离。相应的复相干系数为：

$$\mu(\Delta x, \Delta y) = e^{i\psi} \left[\frac{2J_1(2\pi a \rho)}{2\pi a \rho} \right] \quad (5.7.15)$$

上式的第一个因子： $e^{i\psi}$ 取决于 (x_1, y_1) 和 (x_2, y_2) ；而第二个因子仅仅取决于两点之间的距离 s 。这样复相干系数的模 $\tilde{\mu}$ 仅取决于 Δx 和 Δy ，如图 5.7.2 所示。一阶贝塞尔函数的第一个零点为 3.83，因此 $\tilde{\mu}$ 的第一个

零点发生在间距:

$$s_0 = 0.61 \frac{\bar{\lambda}z}{a} = 1.22 \frac{\bar{\lambda}}{\theta} \quad (5.7.16)$$

的地方。式中 $\theta = \frac{2a}{z}$ 是小角度近似下光源对观察面坐标原点的张角。

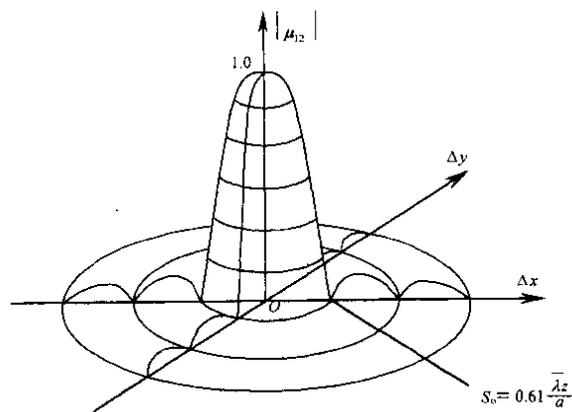


图 5.7.2 复相干系数与坐标差的关系

假定点 (x_1, y_1) 和 (x_2, y_2) 相应于一个不透明屏上的针孔，且在屏后一定距离上观察干涉条纹。根据 μ 的特征可以预示在每种一个可能的针孔间距下得到的条纹特征。

第六章 光电器件的物理基础

本章着重介绍两个主要内容：一个是辐射量和光度量的定义及它们之间的换算关系；另一个是半导体光电器件的物理基础，如能带理论、PN 结理论、半导体光电导效应和光电发射过程等。这些是以后各章所述具体光电器件的理论基础，对于正确理解和掌握各种光电器件的原理、性能和用法是十分重要的。

6.1 光谱与光子能量

光谱与光子能量

光具有波粒二象性，既是电磁波，又是光子流。1860 年麦克斯韦 (C. Maxwell) 提出光是电磁波的理论。光在传播时表现出波动性，如光的干涉、衍射、偏振、反射、折射。1900 年，普朗克 (Max. Planck) 提出了辐射的量子论，1905 年，爱因斯坦 (Albert. Einstein) 将量子论用于光电效应之中，提出光子理论。光与物质作用时表现出粒子性，如光的发射、吸收、色散、散射。

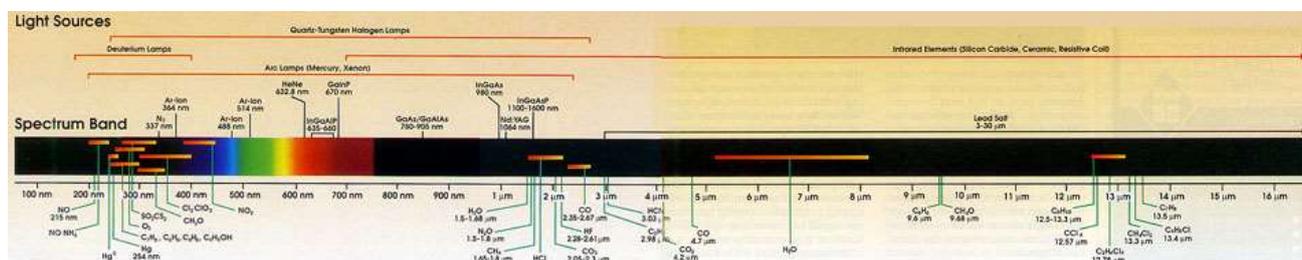


麦克斯韦 (1831-1879)

普朗克 (1858-1947)

爱因斯坦 (1879-1955)

电磁波谱分为长波区、光学区、射线区。光电技术只涉及光学谱区。在光学谱区内，具有相同的辐射与吸收机理，许多辐射源的光谱分布和接收器的灵敏阈都同时覆盖此区域。使用光学透镜来接收辐射或聚焦成象。



电磁波谱及光谱图

光子能量公式： $\epsilon = h \nu$

光子动量公式： $p = h \nu / c = h / \lambda$

h : 普朗克常数

上面两公式等号左边表示光为微粒性质 (光子能量与动量)，等号右边表示光为波动性质 (电磁波频率和波长)。

光电转换一般使用固体材料，利用其量子效应。从固体能级来说，具有从 0.1eV 到几个 eV 能量的转换比较容易，即比较容易在十几微米的红外到 0.2 微米左右的紫外范围内进行高效率的能量转换。

6.2 光度学(Photometry)与辐射度学(Radiometry)

光度学(Photometry)与辐射度学(Radiometry)

光电系统可以看作是光能的传递和接收系统。辐射能从目标(辐射源)发出后经过中间介质、光学系统,最后被光电器件接收。光能的强弱是否能使接收器感受,这是光电系统一个很重要的指标。

光度学研究对可见光的能量的计算,它使用的参量称为光度量。以人的视觉习惯为基础建立。

辐射度学适用于整个电磁波谱的能量计算。主要用于 X 光、紫外光、红外光以及其他非可见的电磁辐射。

光度学是辐射度学的一部分或特例。这两套参量的名称、符号、定义式彼此对应,基本都相同,只是单位不同。为了区别这两种量,规定用下标 e 和 v 表示。

常用辐射量和光度量一览表

名称	符号	定义	单位
辐[射]能	Q_e	以辐射形式发射、传播或接收的能量。	焦[耳](J)
光量	Q_v	光通量对时间的积分。	流[明]秒($lm \cdot s$)
辐[射能]通量	Φ_e	以辐射形式发射、传播或接收的功率。	瓦[特](W)
光通量	Φ_v	发光强度为 I_v 的光源,在立体角元 $d\Omega$ 内的辐通量, $d\Phi_v = I_v \cdot d\Omega$ 。	流[明](lm)
辐[射]出[射]度	M_e	离开表面一点处面元的辐通量除以该面元面积。	瓦每平方米 ($W \cdot m^{-2}$)
光出射度	M_v	离开表面一点处面元的光通量除以该面元面积。	流[明]每平方米 ($lm \cdot m^{-2}$)
辐[射]照度	E_e	照射到表面一点处面元上的辐通量除以该面元的面积。	瓦每平方米 ($W \cdot m^{-2}$)
[光]照度	E_v	照射到表面一点处面元上的光通量除以该面元的面积。	勒[克斯](lx)
辐[射]强度	I_e	在给定方向上的立体角元内,离开点辐射源或辐射源面元的辐射功率除以该立体角元。	瓦每球面度 ($W \cdot sr^{-1}$)
发光强度	I_v	光度量中的基本量,单位为坎德拉 cd。cd 的意义为: 频率为 $540 \times 10^{10} Hz$ 的单色辐射在给定方向上的辐射强度 $I_e = 1/683 W \cdot sr^{-1}$ 时,规定为 1cd。	坎[德拉](cd)
辐射亮度	L_e	表面一点处的面元在给定方向上的辐射强度除以该面元在垂直于给定方向平面上的正投影面积。	瓦每球面度平 方米($W \cdot sr^{-1} \cdot m^{-2}$)
[光]亮度	L_v	表面一点处的面元在给定方向上的发光强度除以该面元在垂直于给定方向平面上的正投影面积。	坎[德拉]每平方 米($cd \cdot m^{-2}$)
曝光量	H	光照度对时间的积分。	勒[克斯]秒 ($lx \cdot s$)

6.3 辐通量的光谱分布与接收器的光谱响应

辐射一般由各种波长组成，每种波长的辐通量各不相同。总的辐通量为各个组成波长的辐通量的总和。下图为某辐通量的连续分布曲线。

如图，给定波长 λ_0 处极小波长间隔 $d\lambda$ 内的辐通量 $d\Phi_e$ 称为单色辐通量。

$\Phi_{e\lambda} = d\Phi_e / d\lambda$ ， $\Phi_{e\lambda}$ 称为光谱辐通量。

单色辐通量的积分为

$$\Phi_e = \int_{\lambda_1}^{\lambda_2} \Phi_{e\lambda} d\lambda$$

此式中 Φ_e 称为多色辐通量。

$$\Phi_e = \int_0^{\infty} \Phi_{e\lambda} d\lambda$$

此式中 Φ_e 称为全色辐通量。



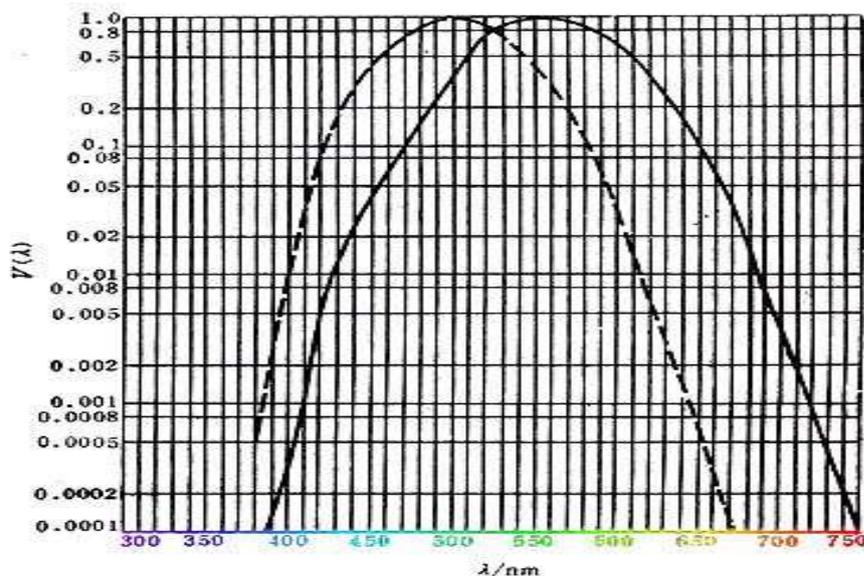
许多接收器所能感受的波长是有选择性的，接收器对不同波长电磁辐射的响应程度（反应灵敏度）称为光谱响应度或光谱灵敏度。对人眼来说采用光谱光视效能 $K(\lambda)$ 来表征不同波长辐射下的响应能力，光谱光视效能 $K(\lambda)$ 为同一波长下光谱光通量与光谱辐通量之比，即

$$K(\lambda) = \Phi_v \lambda / \Phi_e \lambda$$

由于人眼在频率为 $540 \times 10^{12} \text{ Hz}$ ($\lambda_m = 555 \text{ nm}$ ，该波长称为峰值波长) 的辐射下， $K(\lambda)$ 最大，记以 K_m ， $K_m = 683 \text{ lm} \cdot \text{W}^{-1}$ 对于某给定波长下的 K_m ，定义光谱光视效率 $V(\lambda)$ 为

$$V(\lambda) = K(\lambda) / K_m$$

$V(\lambda)$ 又称为视见函数。根据对许多正常人眼的研究，可统计出各种波长的平均相对灵敏度。列于下表：



光谱光视效率 $V(\lambda)$ 曲线

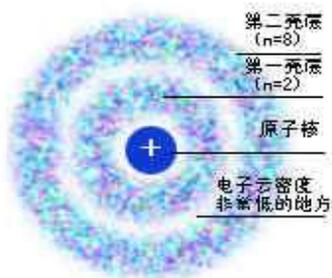
图中实线为在视场较亮时测得的，称为明视觉 $V(\lambda)$ 曲线；虚线为在视场较暗时测得的，称为暗视觉 $V(\lambda)$ 曲线。对于暗视觉， $\lambda_{m'} = 507\text{nm}$ ， $K_{m'} = 683\text{lm} \cdot \text{W}^{-1}$ 。所有光度计量均以明视觉的 $K(\lambda)$ 为基础。

6.4 半导体基础知识

1. 能带理论

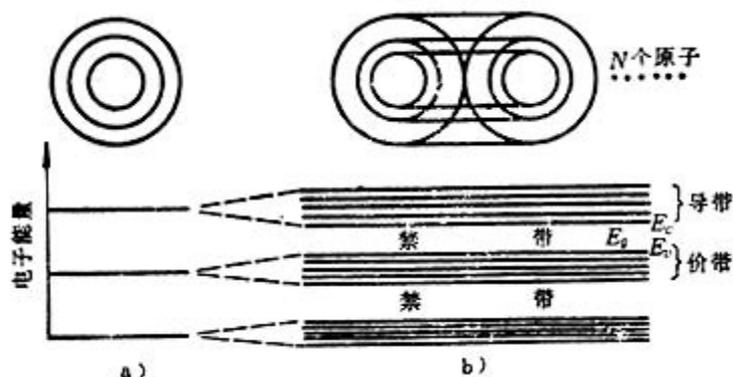
(1) 原子能级与晶体能带

能级 (Energy Level): 在孤立原子中，原子核外的电子按照一定的壳层排列，每一壳层容纳一定数量的电子。每个壳层上的电子具有分立的能量值，也就是电子按能级分布。为简明起见，在表示能量高低的图上，用一条条高低不同的水平线表示电子的能级，此图称为电子能级图。



能带 (Energy Band): 晶体中大量的原子集合在一起，而且原子之间距离很近，以硅为例，每立方厘米的体积内有 5×10^{22} 个原子，原子之间的最短距离为 0.235nm 。致使离原子核较远的壳层发生交叠，壳层交叠使电子不再局限于某个原子上，有可能转移到相邻原子的相似壳层上去，也可能从相邻原子运动到更远的原子壳层上去，这种现象称为电子的共有化。从而使本来处于同一能量状态的电子产生微小的能量差异，

与此相对应的能级扩展为能带。



电子共有化，能级扩展为能带示意图

a) 单个原子 b) N个原子

禁带 (Forbidden Band): 允许被电子占据的能带称为允许带，允许带之间的范围是不允许电子占据的，此范围称为禁带。原子壳层中的内层允许带总是被电子先占满，然后再占据能量更高的外面一层的允许带。被电子占满的允许带称为满带，每一个能级上都没有电子的能带称为空带。

价带 (Valence Band): 原子中最外层的电子称为价电子，与价电子能级相对应的能带称为价带。

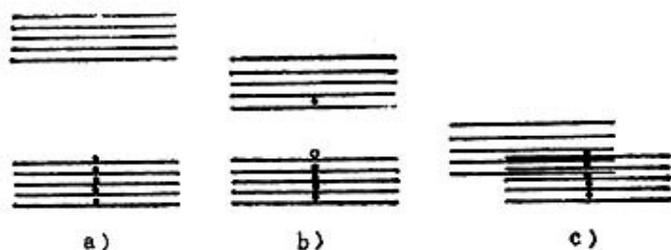
导带 (Conduction Band): 价带以上能量最低的允许带称为导带。

导带的底能级表示为 E_c ，价带的顶能级表示为 E_v ， E_c 与 E_v 之间的能量间隔称为禁带 E_g 。

导体或半导体的导电作用是通过带电粒子的运动（形成电流）来实现的，这种电流的载体称为载流子 (carrier)。导体中的载流子是自由电子，半导体中的载流子则是带负电的电子和带正电的空穴。对于不同的材料，禁带宽度不同，导带中电子的数目也不同，从而有不同的导电性。例如，绝缘材料 SiO_2 的 E_g 约为 5.2eV，导带中电子极少，所以导电性不好，电阻率大于 $10^{12} \Omega \cdot \text{cm}$ 。

半导体 Si 的 E_g 约为 1.1eV，导带中有一定数目的电子，从而有一定的导电性，电阻率为 $10^{-3} - 10^{12} \Omega \cdot \text{cm}$ 。

金属的导带与价带有一定程度的重合， $E_g=0$ ，价电子可在金属中自由运动，导电性好，电阻率 $10^{-6} - 10^{-3} \Omega \cdot \text{cm}$ 。



绝缘体、半导体、金属的能带图

a) 绝缘体 b) 半导体 c) 金属

(2) 本征半导体与杂质半导体

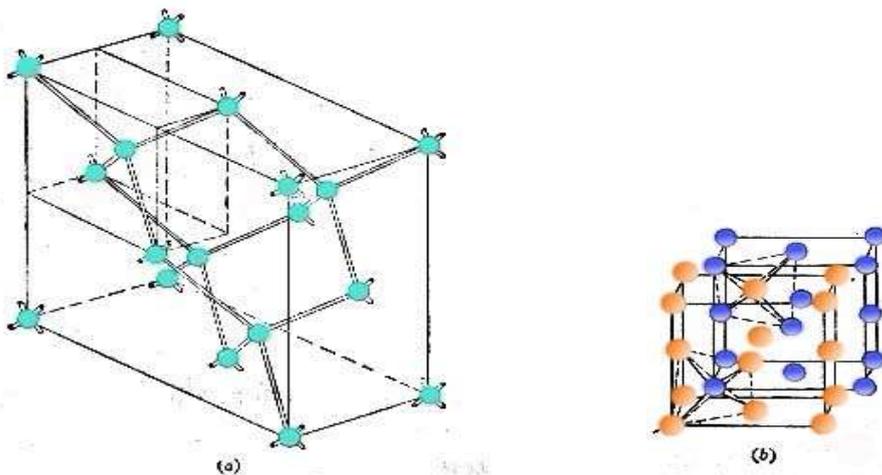
现代固体电子与光电子器件大多由半导体材料制备，半导体材料大多为晶体（晶体中原子有序排列，非晶体中原子无序排列。）晶体分为单晶与多晶：

单晶——在一块材料中，原子全部作有规则的周期排列。

多晶——只在很小范围内原子作有规则的排列，形成小晶粒，而晶粒之间有无规则排列的晶粒界隔开。



晶体实物图



晶体构造示意图

(a) 金刚石结构 (Ge、Si 晶体) (b) 闪锌矿结构 (GaAs 晶体)

本征半导体:

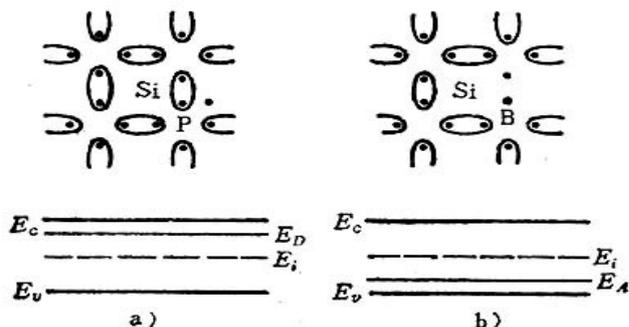
结构完整、纯净的半导体称为本征半导体。例如纯净的硅称为本征硅。本征硅中，自由电子和空穴都是由共价键断裂而产生的，所以电子浓度 n 等于空穴浓度 p ，并称之为本征载流子浓度 n_i ， n_i 随温度升高而增加，随禁带宽度的增加而减小，室温下硅的 n_i 约为 $10^{10}/\text{cm}^3$ 。

杂质半导体:

半导体中人为地掺入少量杂质形成掺杂半导体，杂质对半导体导电性能影响很大。在技术上通常用控制杂质含量（即掺杂）来控制半导体导电特性。

N 型半导体:

在四价原子硅 (Si) 晶体中掺入五价原子，例如磷 (P) 或砷 (As)，形成 N 型半导体。在晶格中某个硅原子被磷原子所替代，五价原子用四个价电子与周围的四价原子形成共价键，而多余一个电子，此多余电子受原子束缚力要比共价键上电子所受束缚力小得多，容易被五价原子释放，游离跃迁到导带上形成自由电子。易释放电子的原子称为施主，施主束缚电子的能量状态称为施主能级 E_D 。 E_D 位于禁带中，较靠近材料的导带底。 E_D 与 E_c 间的能量差称为施主电离能。N 型半导体由施主控制材料导电性。



杂质硅的原子图象和能带图

a) N 型半导体 b) P 型半导体

P 型半导体:

在四价原子硅 (Si) 晶体中掺入三价原子, 例如硼 (B), 形成 P 型半导体。晶体中某个硅原子被硼原子所替代, 硼原子的三个价电子和周围的硅原子中四个价电子要组成共价键, 形成八个电子的稳定结构, 尚缺一个电子。于是很容易从硅晶体中获取一个电子形成稳定结构, 使硼原子外层多了一个电子变成负离子, 而在硅晶体中出现空穴。容易获取电子的原子称为受主。受主获取电子的能量状态称为受主能级 E_A , 也位于禁带中。在价带顶 E_v 附近, E_A 与 E_v 间能量差称为受主电离能。P 型半导体由受主控制材料导电性。

N 型半导体与 P 型半导体的比较

半导体	所掺杂质	多数载流子 (多子)	少数载流子 (少子)	特性
N 型	施主杂质	电子	空穴	电子浓度 $n_n \gg$ 空穴浓度 p_n
P 型	受主杂质	空穴	电子	电子浓度 $n_p \ll$ 空穴浓度 p_p

掺杂对半导体导电性能的影响:

半导体中不同的掺杂或缺陷都能在禁带中产生附加的能级, 价带中的电子若先跃迁到这些能级上然后再跃迁到导带中去, 要比电子直接从价带跃迁到导带容易得多。因此虽然只有少量杂质, 却会明显地改变导带中的电子和价带中的空穴数目, 从而显著地影响半导体的电导率。

2. 热平衡态下的载流子

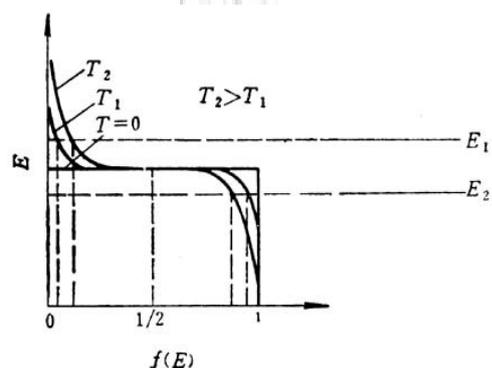
在一定温度下, 若没有其他的外界作用, 半导体中的自由电子和空穴是由热激发产生的。电子从不断热振动的晶体中获得一定的能量, 从价带跃迁到导带, 形成自由电子, 同时在价带中出现自由空穴。在热激发同时, 电子也从高能级的量子态跃迁到低能级的量子状态, 向晶格放出能量, 这就是载流子的复合。在一定温度下, 激发和复合两种过程形成平衡, 称为热平衡状态, 此时载流子浓度即为某一稳定值。

热平衡时半导体中自由载流子浓度与两个参数有关: 一是在能带中能态 (或能级) 的分布, 二是这些能态中每一个能态可能被电子占据的概率。

根据量子理论和泡利不相容原理, 能态分布服从费米统计分布规律。

在某温度下热平衡态, 能量为 E 的能态被电子占据的概率由费米-狄拉克函数给出, 即

$$f(E) = \frac{1}{1 + e^{(E - E_f)/kT}}$$



费米-狄拉克函数曲线

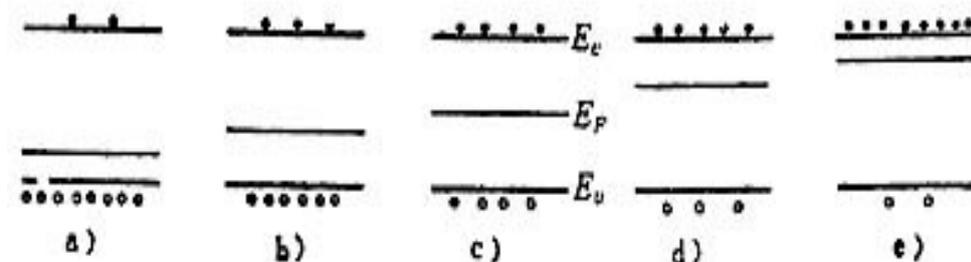
$f(E)$: 费米分布函数, 能量 E 的概率函数

k : 玻耳兹曼常数, $1.38 \times 10^{-23} \text{J/K}$

T : 绝对温度

E_F : 费米能级

E_F 为表征电子占据某能级 E 的概率的“标尺”, 它定性表示导带中电子或价带中空穴的多少。当 $E=E_F$ 时, $f(E)=1/2$, 它并不代表可为电子占据的真实能级, 只是个参考能量。在量子统计中 E_F 应视为固体中电子的化学势。常温下 E_F 随材料掺杂程度而变化。



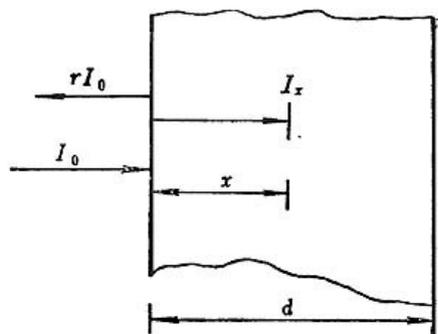
以 E_F 来定性表示两能带中载流子的浓度

- a) 重掺杂 P 型 b) 轻掺杂 P 型 c) 本征型
- d) 轻掺杂 N 型 e) 重掺杂 N 型

3. 半导体对光的吸收

半导体材料吸收光子能量转换成电能是光电器件的工作基础。光垂直入射到半导体表面时, 进入到半导体内的光强遵照吸收定律:

$$I_x = I_0(1-r)e^{-\alpha x}$$



光垂直入射于半导体表面时发生反射与吸收

I_x : 距离表面 x 远处的光强

I_0 : 入射光强

r : 材料表面的反射率

α : 材料吸收系数, 与材料、入射光波长等因素有关

本征吸收与非本征吸收:

本征吸收:

半导体吸收光子的能量使价带中的电子激发到导带, 在价带中留下空穴, 产生等量的电子与空穴, 这种吸收过程叫本征吸收。

产生本征吸收的条件: 入射光子的能量 ($h\nu$) 至少要等于材料的禁带宽度 E_g 。即

$$h\nu \geq E_g$$

从而有

$$\nu_0 \geq E_g/h$$

$$\lambda_0 \leq hc/E_g = 1.24 \mu\text{m} \cdot \text{eV}/E_g$$

h: 普朗克常数

c: 光速

ν_0 : 材料的频率阈值

λ_0 : 材料的波长阈值

几种重要半导体材料的波长阈值

材料	温度/K	E_g /eV	λ / μm	材料	温度/K	E_g /eV	λ / μm
Se	300	1.8	0.69	InSb	300	0.18	6.9
Ge	300	0.81	1.5	GaAs	300	1.35	0.92
Si	290	1.09	1.1	Gap	300	2.24	0.55
PbS	295	0.43	2.9				

非本征吸收:

[各种非本征吸收的示意图, 表示出不同能量光子的不同吸收]

非本征吸收包括杂质吸收、自由载流子吸收、激子吸收和晶格吸收等。

杂质吸收: 杂质能级上的电子(或空穴)吸收光子能量从杂质能级跃迁到导带(空穴跃迁到价带), 这种吸收称为杂质吸收。杂质吸收的波长阈值多在红外区或远红外区。

自由载流子吸收: 导带内的电子或价带内的空穴也能吸收光子能量, 使它在本能带内由低能级迁移到高能级, 这种吸收称为自由载流子吸收, 表现为红外吸收。

激子吸收: 价带中的电子吸收小于禁带宽度的光子能量也能离开价带, 但因能量不够还不能跃迁到导带成为自由电子。这时, 电子实际还与空穴保持着库仑力的相互作用, 形成一个电中性系统, 称为激子。能产生激子的光吸收称为激子吸收。这种吸收的光谱多密集与本征吸收波长阈值的红外一侧。

晶格吸收: 半导体原子能吸收能量较低的光子, 并将其能量直接变为晶格的振动能, 从而在远红外区形成一个连续的吸收带, 这种吸收称为晶格吸收。

半导体对光的吸收主要是本征吸收。对于硅材料, 本征吸收的吸收系数比非本征吸收的吸收系数要大几十倍到几万倍, 一般照明下只考虑本征吸收, 可认为硅对波长大于 $1.15 \mu\text{m}$ 的可见光透明。

4. 非平衡态下的载流子

半导体在外界条件有变化(如受光照、外电场作用、温度变化)时, 载流子浓度要随之发生变化, 此时系统的状态称为非热平衡态。载流子浓度对于热平衡状态时浓度的增量称为非平衡载流子。

电注入: 通过半导体界面把载流子注入半导体, 使热平衡受到破坏。

光注入: 光注入下产生非平衡载流子表现为价带中的电子吸收了光子能量从价带跃迁到导带, 同时在价带中留下等量的空穴。

1. 产生与复合

使非平衡载流子浓度增加的运动称为产生, 单位时间、单位体积内增加的电子空穴对数目称为产生率 G 。

使非平衡载流子浓度减少的运动称为复合, 单位时间、单位体积内减少的电子空穴对数目称为复合率

R。

以 N 型半导体为例，在非平衡状态下载流子浓度为：

$$n_n = n_{n0} + \Delta n_n$$

$$p_n = p_{n0} + \Delta p_n$$

$$\Delta n_n \approx \Delta p_n$$

n_n ：N 型半导体中多数载流子电子的浓度

p_n ：N 型半导体中少数载流子空穴的浓度

n_{n0} ：光照前一定温度下热平衡时电子的浓度

p_{n0} ：光照前一定温度下热平衡时空穴的浓度

Δn_n ：非平衡载流子电子的浓度

Δp_n ：非平衡载流子空穴的浓度

光注入分为强光注入与弱光注入：

满足

$$n_n p_n \gg n_{n0} p_{n0} = n_i^2$$

$$n_{n0} < \Delta n_n = \Delta p_n$$

条件的注入称为强光注入；

满足

$$n_n p_n > n_{n0} p_{n0} = n_i^2$$

$$n_{n0} > \Delta n_n = \Delta p_n$$

条件的注入称为弱光注入。

对于弱光注入

$$n_n = n_{n0} + \Delta n_n \approx n_{n0}$$

$$p_n = p_{n0} + \Delta p_n \approx \Delta p_n$$

此时受影响最大的是少子浓度，可认为一切半导体光电器件对光的响应都是少子行为。例如：一 N 型硅片，室温下， $n_{n0} = 5.5 \times 10^{15} \text{cm}^{-3}$ ， $p_{n0} = 3.5 \times 10^4 \text{cm}^{-3}$ ；弱光注入下， $\Delta n = \Delta p = 10^{10} \text{cm}^{-3}$ ，此时非平衡载流子浓度

$$n_n = n_{n0} + \Delta n_n = 10^{15} + 10^{10} \approx 10^{15} \text{cm}^{-3}$$

$$p_n = p_{n0} + \Delta p_n = 10^4 + 10^{10} \approx 10^{10} \text{cm}^{-3}$$

在光照过程中，产生与复合同时存在，在恒定持续光照下产生率保持在高水平，同时复合率也随非平衡载流子的增加而增加，直至二者相等，系统达到新的平衡。当光照停止，光致产生率为零，系统稳定态遭到破坏，复合率大于产生率，使非平衡载流子浓度逐渐减少，复合率随之下降，直至复合率等于热致的产生率时，非平衡载流子浓度将为零，系统恢复热平衡状态。

2. 复合与非平衡载流子寿命 τ

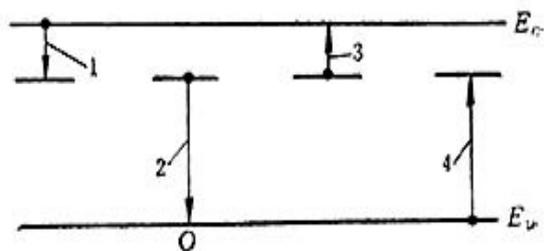
复合是指电子与空穴相遇时，成对消失，以热或发光方式释放出多余的能量。

非平衡载流子寿命 τ ：非平衡载流子从产生到复合之前的平均存在时间。它表征复合的强弱， τ 小表示复合快， τ 大表示复合慢。它决定了光电器件的时间特性，采用光激发方式的光生载流子寿命与光电转换的效果有直接关系。 τ 的大小与材料的微观复合结构、掺杂、缺陷有关。

三种复合机制：

直接复合：导带中电子直接跳回到价带，与价带中的空穴复合。

通过复合中心复合：复合中心指禁带中杂质及缺陷。通过复合中心间接复合包括四种情况：电子从导带落入到复合中心称电子俘获；电子从复合中心落入价带称空穴俘获；电子从复合中心被激发到导带称电子发射；电子从价带被激发到复合中心称空穴发射。



通过复合中心进行的复合-产生过程

1—电子俘获 2—空穴俘获 3—电子发射 4—空穴发射

表面复合：材料表面在研磨、抛光时会出现许多缺陷与损伤，从而产生大量复合中心。发生于半导体表面的复合过程称为表面复合。

5. 载流子的输运——扩散与漂移

电子在晶体中的运动与气体分子的热运动类似。当没有外加电场时，电子作无规则运动，其平均定向速度为零。一定温度下半导体中电子和空穴的热运动是不能引起载流子净位移，从而也就没有电流。但漂移和扩散可使载流子产生净位移，从而形成电流。

漂移：

载流子在外电场作用下，电子向正电极方向运动，空穴向负电极方向运动称为漂移。

在强电场作用下，由于饱和或雪崩击穿半体会偏离欧姆定律。在弱电场作用下，半导体中载流子漂移运动服从欧姆定律。

讨论漂移运动的重要参量：

迁移率 μ (电子迁移率 μ_n ，空穴迁移率 μ_p)， μ 的大小主要决定于晶格振动及杂质对载流子的散射作用。

从欧姆定律的微分形式：

$$j = \sigma E$$

j ：电流密度

σ ：材料导电率

E ：电场强度

从电流密度的定义：

$$j = nq v$$

n ：电子浓度

q ：电子电量

v ：电子漂移平均速度

故 $nq v = \sigma E$ (上面两式恒等)

$$v = (\sigma / nq) E = \mu_n \cdot E$$

表明电子漂移的平均速度与场强成正比。

在电场中电子所获得的加速度

$$a = qE / m^*$$

qE ：表征电场力

m^* ：电子有效质量，考虑了晶格对电子运动的影响并对电子静止质量进行修正后得到的值。

在漂移运动中，因电子与晶格碰撞发生散射，故每次碰撞后漂移速度降到零。如两次碰撞之间的平均时间为 t_c ，则经 t_c 后载流子的

$$v = a \cdot t_c = (qE/m^* \cdot t_c) = (qt_c/m^*) \cdot E$$

有

$$\mu = qt_c/m^*$$

表明 μ 与 t_c 、 m^* 有关。

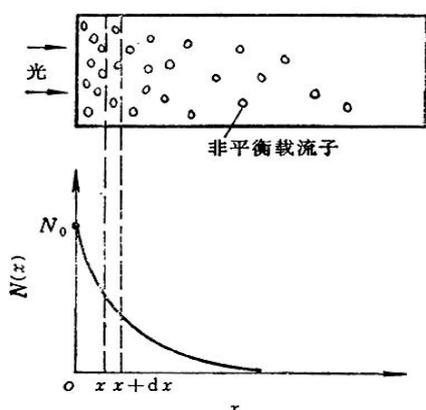
在同一种半导体中，因电子与空穴运动状态不同， m^* 各不相同，故 μ_p 、 μ_n 不同。

同一种载流子在导电类型不同的半导体中，因浓度不同，平均自由程不同， t_c 也不同，故 μ 也不同。半导体中杂质浓度增加时，载流子碰撞机会增多， t_c 减小， μ 将随之减小。

扩散：

载流子因浓度不均匀而发生的从浓度高的点向浓度低的点运动。

下图为光注入，非平衡载流子扩散示意图。光在受照表面很薄一层内即被吸收掉。受光部分将产生非平衡载流子，其浓度随离开表面距离 x 的增大而减小，因此非平衡载流子就要沿 x 方向从表面向体内扩散，使自己在晶格中重新达到均匀分布。



光注入，非平衡载流子扩散示意图

扩散流面密度 j 与浓度梯度 $dN(x)/dx$ 成正比：

$$j = -D \cdot dN(x)/dx$$

D 为扩散系数，表征非平衡载流子扩散能力。式中负号表示扩散流方向与浓度梯度方向相反。

下列关系式成立：

$$(-D \cdot dN(x)/dx)_x - (-D \cdot dN(x)/dx)_{x+\Delta x} = N(x) \Delta x / \tau$$

τ ：非平衡载流子平均寿命

非平衡载流子沿 x 轴分布是在边扩散边复合中形成的，定态下， $N(x)$ 分布稳定，单位时间内复合的非平衡载流子数必然要靠净扩散流补偿。

上式两边同除以 Δx ，并对等号左边取 $\Delta x \rightarrow 0$ 极限得扩散方程：

$$d^2N(x)/dx^2 = N(x) / (\tau D)$$

利用边界条件 $x=0, N(x)=N_0$ ； $x=\infty, N(x)=0$ ，得

$$N(x) = N_0 e^{-x/L}$$

$L = (\tau D)^{1/2}$ 称为扩散长度，表示 $N(x)$ 减小到的 N_0 的 $1/e$ 时所对应的距离 x 。光生的非平衡载流子复合有先后，在复合前扩散的距离有远近之分，从而形成 $N(x)$ 分布曲线。 L 表示非平衡载流子复合前在半导体中扩散的平均深度。

在扩散与漂移同时存在（半导体既受光照，又外加电场时）的情况下，扩散系数 D (D 表示扩散的难易) 与迁移率 μ (μ 表示迁移的快慢) 之间有爱因斯坦关系式：

$$D = (kT/q) \mu$$

kT/q 为比例系数，室温下为 $0.026V$ 。

D 与 μ 成正比。

电子与空穴沿 x 轴扩散，但 $D_n \neq D_p$ ，故它们引起的扩散流不能抵消。在电场中多子、少子均作漂移运动，因多子数目远比少子多，所以漂移流主要是多子的贡献；在扩散情况下，如光照产生非平衡载流子，此时非平衡少子的浓度梯度最大，所以对扩散流的贡献主要是少子。

6.5 光电效应

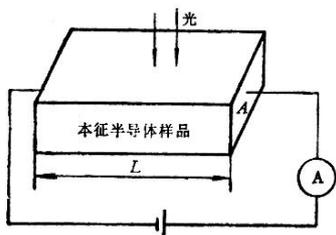
因光照而引起物体电学特性的改变统称为光电效应。

本世纪最伟大的科学家之一[爱因斯坦](#)以他在 1905 年发表的相对论而闻名于世，而他在 1921 年获得诺贝尔奖是由于他对发现光电效应的贡献。

6.5.1 光电导效应

光照变化引起半导体材料电导变化的现象称光电导效应。当光照射到半导体材料时，材料吸收光子的能量，使非传导态电子变为传导态电子，引起载流子浓度增大，因而导致材料电导率增大。

稳态光电导



本征半导体光电导效应图

半导体无光照时为暗态，此时材料具有暗电导；有光照时为亮态，此时具有亮电导。如果给半导体材料外加电压，通过的电流有暗电流与亮电流之分。亮电导与暗电导之差称为光电导，亮电流与暗电流之差称为光电流。

暗态下

$$G_d = \sigma_d \cdot A/L, \quad I_d = G_d U = \sigma_d \cdot AU/L$$

亮态下

$$G_l = \sigma_l \cdot A/L, \quad I_l = G_l U = \sigma_l \cdot AU/L$$

亮态与暗态之差

$$G_p = G_l - G_d = (\sigma_l - \sigma_d) \cdot A/L = \Delta \sigma \cdot A/L$$

$$I_p = I_l - I_d = (G_l - G_d) \cdot U = \Delta \sigma \cdot AU/L$$

A: 半导体材料横截面面积

L: 半导体材料长度

I: 电流

U: 外加电压

G: 电导

σ : 电导率

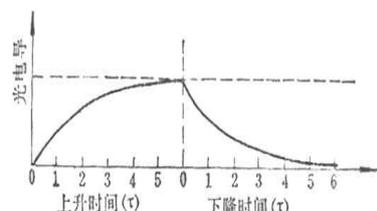
$\Delta \sigma$: 光致电导率的变化量

下标 d 代表暗，l 代表亮，p 代表光。

光电导弛豫过程

光电导材料从光照开始到获得稳定的光电流是要经过一定时间的。同样光照停止后光电流也是逐渐消失的。这些现象称为弛豫过程或惰性。

对光电导体受矩形脉冲光照时，常有上升时间常数 τ_r 和下降时间常数 τ_f 来描述弛豫过程的长短。 τ_r 表示光生载流子浓度从零增长到稳态值 63% 时所需的时间， τ_f 表示从停光前稳态值衰减到 37% 时所需的时间。



矩形脉冲光照弛豫过程图

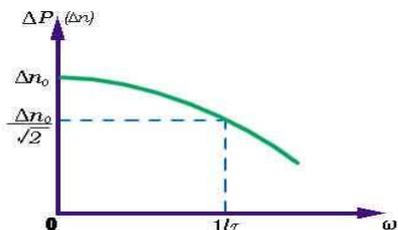
当输入光功率暗正弦规律变化时，光生载流子浓度（对应于输出光电流）与光功率频率变化的关系，是一个低通特性，说明光电导的弛豫特性限制了器件对调制频率高的光功率的响应：

$$\Delta n = \frac{\Delta n_0}{\sqrt{1 + (\omega\tau)^2}}$$

Δn_0 ：中频时非平衡载流子浓度。

ω ：圆频率， $\omega = 2\pi f$ 。

τ ：非平衡载流子平均寿命，在这里称时间常数。



正弦光照弛豫过程图

可见 Δn 随 ω 增加而减小，当 $\omega = 1/\tau$ 时， $\Delta n = \Delta n_0 / \sqrt{2}$ ，称此时 $f = 1/2\pi\tau$ 为上限截止频率或带宽。

光电增益与带宽之积为一常数，

$$\begin{aligned} Mf &= (\tau_n/t_n + \tau_p/t_p) \cdot (1/2\pi\tau) \\ &= (1/t_n + 1/t_p) \cdot (1/2\pi) \\ &= \text{常数} \end{aligned}$$

表明材料的光电灵敏度与带宽是矛盾的：材料光电灵敏度高，则带宽窄；材料带宽宽，则光电灵敏度低。此结论对光电效应现象有普遍性。

光电效应：

1 外光电效应

光电发射效应：当光照射到物体上使物体向真空中发射电子

2 内光电效应

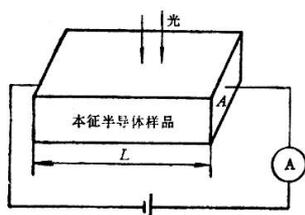
光电导效应：电导率发生变化

光伏效应：产生光电势

6.4.1 光电导效应

光照变化引起半导体材料电导变化的现象称光电导效应。当光照射到半导体材料时，材料吸收光子的能量，使非传导态电子变为传导态电子，引起载流子浓度增大，因而导致材料电导率增大。

稳态光电导



本征半导体光电导效应图

半导体无光照时为暗态，此时材料具有暗电导；有光照时为亮态，此时具有亮电导。如果给半导体材料外加电压，通过的电流有暗电流与亮电流之分。亮电导与暗电导之差称为光电导，亮电流与暗电流之差称为光电流。

暗态下

$$G_d = \sigma_d \cdot A/L, \quad I_d = G_d U = \sigma_d \cdot AU/L$$

亮态下

$$G_l = \sigma_l \cdot A/L, \quad I_l = G_l U = \sigma_l \cdot AU/L$$

亮态与暗态之差

$$G_p = G_l - G_d = (\sigma_l - \sigma_d) \cdot A/L = \Delta \sigma \cdot A/L$$

$$I_p = I_l - I_d = (G_l - G_d) \cdot U = \Delta \sigma \cdot AU/L$$

A: 半导体材料横截面面积

L: 半导体材料长度

I: 电流

U: 外加电压

G: 电导

σ : 电导率

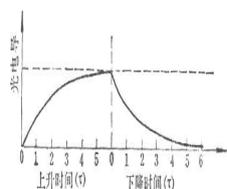
$\Delta \sigma$: 光致电导率的变化量

下标 d 代表暗, l 代表亮, p 代表光。

光电导弛豫过程

光电导材料从光照开始到获得稳定的光电流是要经过一定时间的。同样光照停止后光电流也是逐渐消失的。这些现象称为弛豫过程或惰性。

对光电导体受矩形脉冲光照时，常有上升时间常数 τ_r 和下降时间常数 τ_f 来描述弛豫过程的长短。 τ_r 表示光生载流子浓度从零增长到稳态值 63% 时所需的时间， τ_f 表示从停光前稳态值衰减到 37% 时所需的时间。



矩形脉冲光照弛豫过程图

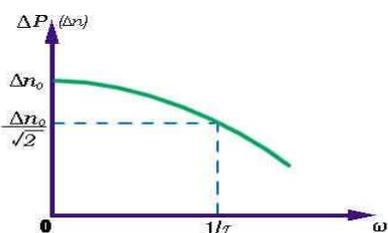
当输入光功率暗正弦规律变化时，光生载流子浓度（对应于输出光电流）与光功率频率变化的关系，是一个低通特性，说明光电导的弛豫特性限制了器件对调制频率高的光功率的响应：

$$\Delta n = \frac{\Delta n_0}{\sqrt{1 + (\omega\tau)^2}}$$

Δn_0 ：中频时非平衡载流子浓度。

ω ：圆频率， $\omega = 2\pi f$ 。

τ ：非平衡载流子平均寿命，在这里称时间常数。



正弦光照弛豫过程图

可见 Δn 随 ω 增加而减小，当 $\omega = 1/\tau$ 时， $\Delta n = \Delta n_0 / \sqrt{2}$ ，称此时 $f = 1/2\pi\tau$ 为上限截止频率或带宽。

光电增益与带宽之积为一常数，

$$\begin{aligned} Mf &= (\tau_n/t_n + \tau_p/t_p) \cdot (1/2\pi\tau) \\ &= (1/t_n + 1/t_p) \cdot (1/2\pi) \\ &= \text{常数。} \end{aligned}$$

表明材料的光电灵敏度与带宽是矛盾的：材料光电灵敏度高，则带宽窄；材料带宽宽，则光电灵敏度低。此结论对光电效应现象有普遍性。

6.4.2 光 伏 效 应

光生伏特效应简称为光伏效应，指光照使不均匀半导体或半导体与金属组合的不同部位之间产生电位差的现象。

产生这种电位差的机理有好几种，主要的一种是由于阻挡层的存在。以下以 P-N 结为例说明。

制作 P-N 结的材料，可以是同一种半导体（同质结），也可以是由两种不同的半导体材料或金属与半导体的结合（异质结）。“结合”指一个单晶体内部根据杂质的种类和含量的不同而形成的接触区域，严格来说是指其中的过渡区。

结有多种：P-N 结、P-I 结、N-I 结、P⁺-P 结、N⁺-N 结等。I 型指本征型，P⁺、N⁺分别指相对于 p、n 型半导体受主、施主浓度更大些。

热平衡态下的 P-N 结

P-N 结的形成：

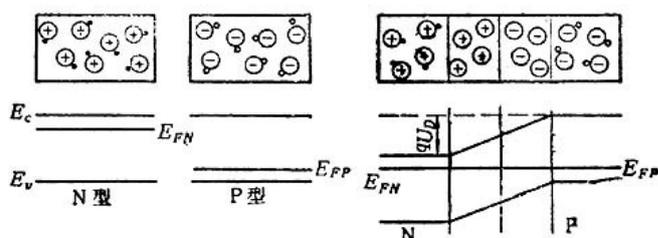
同质结可用一块半导体经掺杂形成 P 区和 N 区。由于杂质的激活能量 ΔE 很小，在室温下杂质差

不多都电离成受主离子 N_A^- 和施主离子 N_D^+ 。在 PN 区交界面处因存在载流子的浓度差，故彼此要向对方扩散。

设想在结形成的一瞬间，在 N 区的电子为多子，在 P 区的电子为少子，使电子由 N 区流入 P 区，电子与空穴相遇又要发生复合，这样在原来是 N 区的结面附近电子变得很少，剩下未经中和的施主离子 N_D^+ 形成正的空间电荷。

同样，空穴由 P 区扩散到 N 区后，由不能运动的受主离子 N_A^- 形成负的空间电荷。

在 P 区与 N 区界面两侧产生不能移动的离子区（也称耗尽区、空间电荷区、阻挡层），于是出现空间电偶层，形成内电场（称内建电场）此电场对两区多子的扩散有抵制作用，而对少子的漂移有帮助作用，直到扩散流等于漂移流时达到平衡，在界面两侧建立起稳定的内建电场。



热平衡下 P-N 结模型及能带图

P-N 结能带与接触电势差：

在热平衡条件下，结区有统一的 E_F ；在远离结区的部位， E_c 、 E_F 、 E_v 之间的关系与结形成前状态相同。

从能带图看，N 型、P 型半导体单独存在时， E_{FN} 与 E_{FP} 有一定差值。

当 N 型与 P 型两者紧密接触时，电子要从费米能级高的一方向费米能级低的一方流动，空穴流动的方向相反。同时产生内建电场，内建电场方向为从 N 区指向 P 区。

在内建电场作用下， E_{FN} 将连同整个 N 区能带一起下移， E_{FP} 将连同整个 P 区能带一起上移，直至将费米能级拉平为 $E_{FN}=E_{FP}$ ，载流子停止流动为止。

在结区这时导带与价带则发生相应的弯曲，形成势垒。势垒高度等于 N 型、P 型半导体单独存在时费米能级之差：

$$qU_D = E_{FN} - E_{FP}$$

得

$$U_D = (E_{FN} - E_{FP}) / q$$

q：电子电量

U_D ：接触电势差或内建电势

对于在耗尽区以外的状态：

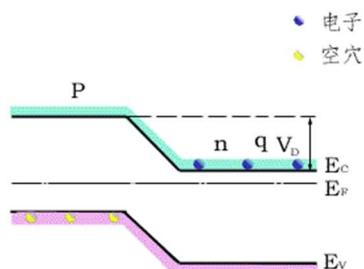
$$U_D = (KT/q) \ln(N_A N_D / n_i^2)$$

N_A 、 N_D 、 n_i ：受主、施主、本征载流子浓度。

可见 U_D 与掺杂浓度有关。在一定温度下，P-N 结两边掺杂浓度越高， U_D 越大。禁带宽的材料， n_i 较小，故 U_D 也大。

光照下的 P-N 结

P-N 结光电效应：



光照下光电流的产生动态图

当 P-N 结受光照时，样品对光子的本征吸收和非本征吸收都将产生光生载流子。但能引起光伏效应的只能是本征吸收所激发的少数载流子。

因 P 区产生的光生空穴，N 区产生的光生电子属多子，都被势垒阻挡而不能过结。只有 P 区的光生电子和 N 区的光生空穴和结区的电子空穴对（少子）扩散到结电场附近时能在内建电场作用下漂移过结。

光生电子被拉向 N 区，光生空穴被拉向 P 区，即电子空穴对被内建电场分离。这导致在 N 区边界附近有光生电子积累，在 P 区边界附近有光生空穴积累。它们产生一个与热平衡 P-N 结的内建电场方向相反的光生电场，其方向由 P 区指向 N 区。此电场使势垒降低，其减小量即光生电势差，P 端正，N 端负。于是有结电流由 P 区流向 N 区，其方向与光电流相反。

实际上，并非所产生的全部光生载流子都对光生电流有贡献。设 N 区中空穴在寿命 τ_p 的时间内扩散距离为 L_p ，P 区中电子在寿命 τ_n 的时间内扩散距离为 L_n 。 $L_n+L_p=L$ 远大于 P-N 结本身的宽度。故可以认为在结附近平均扩散距离 L 内所产生的光生载流子都对光电流有贡献。而产生的位置距离结区超过 L 的电子空穴对，在扩散过程中将全部复合掉，对 P-N 结光电效应无贡献。

光照下的 P-N 结电流方程：

与热平衡时比较，有光照时，P-N 结内将产生一个附加电流（光电流） I_p ，其方向与 P-N 结反向饱和电流 I_0 相同，一般 $I_p \geq I_0$ 。此时

$$I = I_0 e^{qU/KT} - (I_0 + I_p)$$

令 $I_p = SE$ ，则

$$I = I_0 e^{qU/KT} - (I_0 + SE)$$

开路电压 U_{oc} ：

光照下的 P-N 结外电路开路时 P 端对 N 端的电压，即上述电流方程中 $I=0$ 时的 U 值：

$$0 = I_0 e^{qU_{oc}/KT} - (I_0 + SE)$$

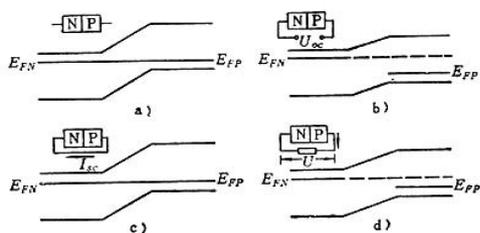
$$U_{oc} = (KT/q) \ln (SE + I_0) / I_0$$

$$\approx (KT/q) \ln (SE / I_0)$$

短路电流 I_{sc} ：

光照下的 P-N 结，外电路短路时，从 P 端流出，经过外电路，从 N 端流入的电流称为短路电流 I_{sc} 。即上述电流方程中 $U=0$ 时的 I 值，得 $I_{sc} = SE$ 。

U_{oc} 与 I_{sc} 是光照下 P-N 结的两个重要参数，在一定温度下， U_{oc} 与光照度 E 成对数关系，但最大值不超过接触电势差 U_D 。弱光照下， I_{sc} 与 E 有线性关系。



不同状态下 P-N 结的能带图

a) 无光照时热平衡态，NP 型半导体有统一的费米能级，势垒高度为 $qU_D = E_{FN} - E_{FP}$ 。

b) 稳定光照下 P-N 结外电路开路，由于光生载流子积累而出现光生电压 U_{oc} ，不再有统一费米能级，势垒高度为 $q(U_D - U_{oc})$ 。

c) 稳定光照下 P-N 结外电路短路，P-N 结两端无光生电压，势垒高度为 qU_D ，光生电子空穴对被内建电场分离后流入外电路形成短路电流。

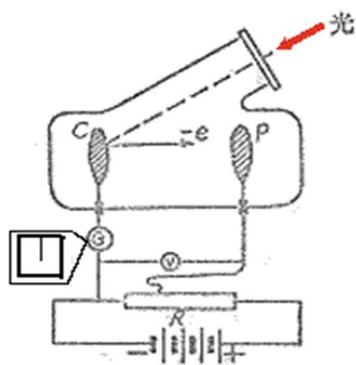
d) 有光照有负载，一部分光电流在负载上建立起电压 U_f ，另一部分光电流被 P-N 结因正向偏压引起的

正向电流抵消，势垒高度为 $q(U_b - U_f)$ 。

6.4.3 光电发射效应

金属或半导体受光照时，如果入射的光子能量 $h\nu$ 足够大，它和物质中的电子相互作用，使电子从材料表面逸出现象，也称为外光电效应。它是真空光电器件光电阴极的物理基础。

外光电效应的两个基本定律



光电发射定律解释图

1. 光电发射第一定律——斯托列托夫定律：

当照射到光阴极上的入射光频率或频谱成分不变时，饱和光电流（即单位时间内发射的光电子数目）与入射光强度成正比：

$$I_k = S_k F_0$$

I_k ：光电流

S_k ：光强

F_0 ：该阴极对入射光线的灵敏度

2. 光电发射第二定律——爱因斯坦定律

光电子的最大动能与入射光的频率成正比，而与入射光强度无关：

$$E_{\max} = (1/2) m v_{\max}^2 = h\nu - h\nu_0 = h\nu - A$$

E_{\max} ：光电子的最大初动能。

h ：普朗克常数。

ν_0 ：产生光电发射的极限频率，频率阈值。

A ：金属电子的逸出功（从材料表面逸出时所需的最低能量），单位 eV，与材料有关的常数，也称功函数。

入射光子的能量至少要等于逸出功时，才能发生光电发射。

波长阈值：

$$\begin{aligned} \lambda_0 = c / \nu_0 &\leq hc / A = 1.24 (\mu\text{m} \cdot \text{eV}) / A \\ &= 1240 / A (\text{nm}) \end{aligned}$$

当入射光波长大于 λ_0 时，不论光强如何，以及照射时间多长，都不会有光电子产生。要用红外光（ $\lambda > 0.76 \mu\text{m}$ ）发射电子，必须寻求低于 1.8eV 的低能阈值材料。

光电发射的基本过程

光电发射大致可分三个过程：

- 1) 光射入物体后，物体中的电子吸收光子能量，从基态跃迁到能量高于真空能级的激发态。
- 2) 受激电子从受激地点出发，在向表面运动过程中免不了要同其它电子或晶格发生碰撞，而失去一部分能量。
- 3) 达到表面的电子，如果仍有足够的能量足以克服表面势垒对电子的束缚（即逸出功）时，即可从表面逸出。



光电发射过程图

- 好的光电发射材料应该是：
- (1) 对光的吸收系数大，以便体内有较多的电子受到激发；
 - (2) 受激电子最好是发生在表面附近，这样向表面运动过程中损失的能量少；
 - (3) 材料的逸出功要小，使到达真空界面的电子能够比较容易地逸出；
 - (4) 另外，作为光阴极，其材料还要有一定的电导率，以便能够通过外电源来补充因光电发射所失去的电子。

金属的光电发射

金属反射掉大部分入射的可见光（反射系数达 90%以上），吸收效率很低。光电子与金属中大量的自由电子碰撞，在运动中丧失很多能量。只有很靠近表面的光电子，才有可能到达表面并克服势垒逸出，即金属中光电子逸出深度很浅，只有几 nm，而且金属逸出功大多为大于 3eV（ $\lambda > 410\text{nm}$ ）的可见光来说，很难产生光电发射，只有铯（2eV 逸出功）对可见光最灵敏，故可用于光阴极。但纯金属铯量子效率很低，小于 0.1%，因在光电发射前两个阶段能量损耗太大。

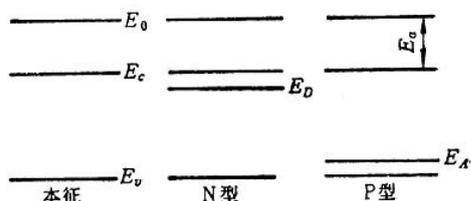
半导体的光电发射

半导体光电逸出参量：

电子亲和势：(Electron Affinity)

导带底上的电子向真空逸出时所需的最低能量，数值上等于真空能级（真空中静止电子能量）与导带底能级 E_c 之差。

它有表面电子亲和势 E_a 与体内电子亲和势 E_{ae} 之分。 E_a 是材料的参量，与掺杂、表面能带弯曲等因素无关。而 E_{ae} 不是材料参量，可随表面能带弯曲变化。



电子亲和势

电子逸出功：电子逸出功是描述材料表面对电子束缚强弱的物理量，在数量上等于电子逸出表面所需的最低能量，也可以说是光电发射的能量阈值。

金属有大量的自由电子，没有禁带，费米能级以下基本上为电子所填满，费米能级以上基本上是空的，表面能带受内外电场影响很小， E_f 只决定于材料。所以金属的电子逸出功定义为 $T=0\text{K}$ 时真空能级与 E_f 之差，

它是材料的参量，可以用来作为光电发射的能量阈值。

半导体自由电子较少，且有禁带，费米能级 E_F 一般都在禁带当中，且随掺杂和内外电场变化，所以真空能级与费米能级之差不是材料参量。

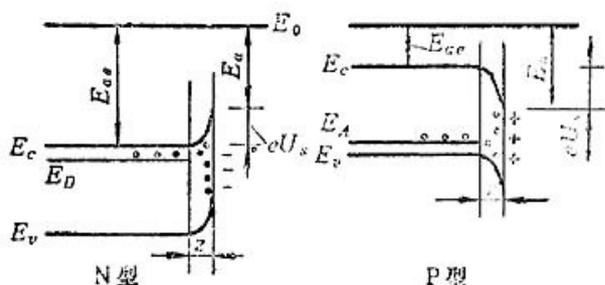
半导体电子逸出功定义为 $T=0K$ 时真空能级与电子发射中心的能级之差，而电子发射中心的能级有的是价带顶，有的是杂质能级，有的是导带底，情况复杂，因此对于半导体很少用电子逸出功的概念。

由于电子逸出功不管从哪里算起，其中都包含有亲和势（真空能级与导带底之差），因此为了表示光电发射的难易，使用亲和势的概念比使用逸出功的概念更有实际意义。所以，对于半导体一般不用逸出功的概念，而用电子亲和势的概念。为了表示光电发射的能量阈值，许多资料都是按真空能级与价带顶之差（亲和势加上禁带宽度）来计算。

表面能带弯曲：

半导体无界时，能带结构是平直的，有界时表面处破坏了晶格排列周期性（势场）而且表面易氧化及被杂质污染，因而在禁带中引入附加能级（表面能级），由于表面能级的存在，在表面处引起能带弯曲。

表面能带弯曲，对于体内的光电子发射是有影响的。因为表面电子亲和势 E_a 是材料的参量，它不随表面能带弯曲而变化。而体内电子亲和势则要随着表面能带弯曲而增减。



表面能带弯曲

对于 N 型半导体，施主能级上的电子跃迁到表面能级时，半导体表面将产生一个负的空间电荷区。而距离表面稍远一点的体内则分布有等量正的体电荷，因此表面能带向上弯。向上弯的程度，可用表面势垒 eU_s 表示， e 为电子电荷， U_s 为表面势，在数值上等于体内与表面的电势差。对于 N 型半导体来说，因表面能带向上弯，体内的电子亲和势 E_{ae} 要比表面能带不发生弯曲时增加一个势垒高度 eU_s ，使得体内光电子发射变得更困难。

对于 P 型半导体，情况正好相反。表面能级中能量高于受主能级的电子有的要跃迁到受主能级上，于是半导体表面即产生一个正的空间电荷区，距离表面稍远一点的体内则分布有等量的负电荷，因此表面能带向下弯。特别是 P 型半导体表面吸附有带正电性的原子（例如铯原子）或 N 型材料的时候，表面上偶电层正电性在外，能带弯曲就更厉害。能带弯曲的程度也用表面势垒 eU_s 表示。表面能带向下弯，使得体内电子亲和势比能带不发生弯曲时减少一个势垒高度 eU_s 。这样，这种表面能带弯曲对于体内的光电子发射十分有利。

因此，现在各种实用光电阴极几乎全是用 P 型半导体材料为衬底，然后在它的表面上再涂上带正电性的金属或 N 型材料而制成的。这样，就能得到向下弯曲的表面能带，减小逸出功。如果再能使带弯曲足够小（带曲区宽度为 z ），以至比材料吸收系数的倒数还小得多（ $z \leq 1/a$ ）时，则可以使光电子发射的主要部位来自于体内。这时，量子效率要比单纯能带弯曲大得多。另外，强 P 型半导体的费米能级十分靠近于价带，这可使热电子发射（暗电流）较小。

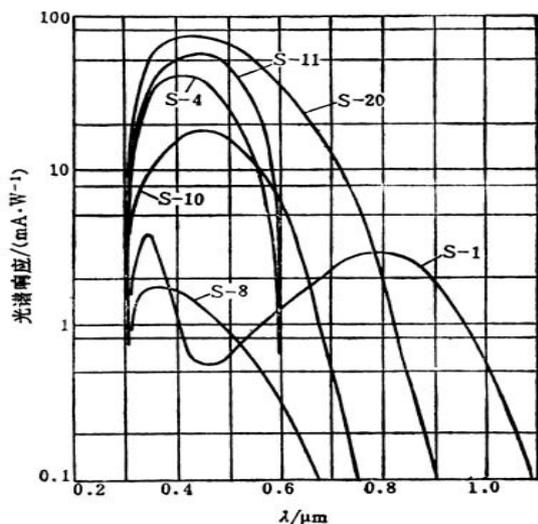
实用光电阴极：分为常规光电阴极、负电子亲和势阴极两大类。

常规光电阴极：用于常用真空光电器件中。根据国际电子工业协会规定，按其出现的先后顺序，以 S 为字头排成序号。

常规光电阴极编号及特性表

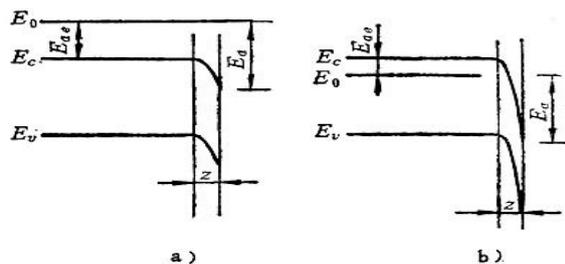
光谱响应 编号	光电发射材料	窗材料	工作方式 半透为T 反射为R	峰值波长 $\lambda_{max}/\mu m$	积分灵敏度 $\mu A \cdot lm^{-1}$	λ_{max} 处辐射 响应度 $/(mA \cdot W^{-1})$	λ_{max} 处量子 效率 (%)	25℃下暗电流 $/(fA \cdot mm^{-2})$
S-1	Ag-O-Cs	石灰玻璃	T, R	0.800	30	2.8	0.43	900
S-3	Ag-O-Rb	石灰玻璃	R	0.42	6.5	1.8	0.53	—
S-4	Cs-Sb	石灰玻璃	R	0.400	40	40	12.4	0.2
S-5	Cs-Sb	9741玻璃	R	0.340	40	60	18.2	0.3
S-8	Cs-Sb	石灰玻璃	R	0.365	3	2.3	0.78	0.13
S-9	Cs-Sb	7052玻璃	T	0.480	30	20.5	5.3	0.3
S-10	Ag-Bi-O-Cs	石灰玻璃	T	0.450	40	20	5.5	70
S-11	Cs-Sb	石灰玻璃	T	0.440	70	66	15.7	3
S-13	Cs-Sb	熔凝玻璃	T	0.440	60	48	13.5	4
S-14	Ge	石灰玻璃	—	1.500	12400 ^①	520 ^①	43 ^①	—
S-16	CdSe	石灰玻璃	—	0.730	—	—	—	—
S-17	Cs-Sb	石灰玻璃	R	0.490	125	83	21	1.2
S-19	Cs-Sb	石灰玻璃	R	0.330	40	65	24.4	0.3
S-20	Na-K-Cs-Sb	熔凝玻璃	R	0.420	150	64	18.8	0.3
S-21	Cs-Sb	9741玻璃	T	0.440	30	23.4	6.6	4
S-23	Rb-Te	熔凝石英	T	0.240	—	4	2	0.001
S-24	K-Na-Sb	7056玻璃	T	0.380	45	67	21.8	0.003
S-25	Na-K-Cs-Sb	石灰玻璃	T	0.420	200	43	12.7	1

① 带45V的起偏电压。



常用光电阴极光谱特性曲线

负电子亲和势 (NEA, Negative Electron Affinity) 阴极:



a) 常规光电发射体物理模型

b) 负电子亲和势光电发射体物理模型

采用特殊工艺，例如在重掺杂 P 型硅表面涂一薄层 CsO₂，可形成 NEA 材料。负电子亲和势是指体内的有效电子亲和势，而不是指表面电子亲和势。

NEA 发射体和常规光电发射体的表面，电子状态是类似的，导带底上的电子能量都低于真空能级，其差值为 E_a。但是，两者体内电子能量则不同。NEA 发射体导带底的电子能量高于真空能级，而常规发射体电子亲和势仍是正的。

NEA 阴极的量子效率高于正电子亲和势阴极，可从其光电发射过程进行分析。价带中的电子吸收光子能量，跃迁到导带底以上，成为热电子（受激电子能量超过导带底的电子）。在向表面运动的过程中，由于碰撞散射而发生能量损失，故很快就落到导带底而变成冷电子（能量恰好等于导带底的电子）。热电子的平均寿命非常短，约 10⁻¹⁴~10⁻¹²s。如果在这么短的时间内能够运动到真空界面，自然能逸出。但是热电子的逸出深度只有几十纳米，绝大部分电子来不及到达真空界面，就已经落到导带底变成冷电子了。冷电子的平均寿命比较长，约 10⁻⁹~10⁻⁸s，其逸出深度可达 1000 纳米。因为体内冷电子能量仍高于真空能级，所以它们运动到真空界面时，可以很容易地逸出。因此 NEA 量子效率比常规发射体高得多。

6.5.2 光 伏 效 应

光生伏特效应简称为光伏效应，指光照使不均匀半导体或半导体与金属组合的不同部位之间产生电位差的现象。

产生这种电位差的机理有好几种，主要的一种是由于阻挡层的存在。以下以 P-N 结为例说明。

制作 P-N 结的材料，可以是同一种半导体（同质结），也可以是由两种不同的半导体材料或金属与半导体的结合（异质结）。“结合”指一个单晶体内部根据杂质的种类和含量的不同而形成的接触区域，严格来说是指其中的过渡区。

结有多种：P-N 结、P-I 结、N-I 结、P⁺-P 结、N⁺-N 结等。I 型指本征型，P⁺、N⁺分别指相对于 p、n 型半导体受主、施主浓度更大些。

热平衡态下的 P-N 结

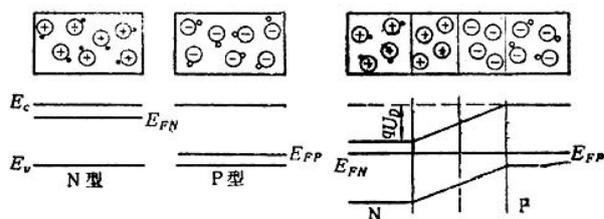
P-N 结的形成：

同质结可用一块半导体经掺杂形成 P 区和 N 区。由于杂质的激活能量 ΔE 很小，在室温下杂质差不多都电离成受主离子 N_a⁻和施主离子 N_b⁺。在 PN 区交界处因存在载流子的浓度差，故彼此要向对方扩散。

设想在结形成的一瞬间，在 N 区的电子为多子，在 P 区的电子为少子，使电子由 N 区流入 P 区，电子与空穴相遇又要发生复合，这样在原来是 N 区的结面附近电子变得很少，剩下未经中和的施主离子 N_b⁺形成正的空间电荷。

同样，空穴由 P 区扩散到 N 区后，由不能运动的受主离子 N_a⁻形成负的空间电荷。

在 P 区与 N 区界面两侧产生不能移动的离子区（也称耗尽区、空间电荷区、阻挡层），于是出现空间电偶层，形成内电场（称内建电场）此电场对两区多子的扩散有抵制作用，而对少子的漂移有帮助作用，直到扩散流等于漂移流时达到平衡，在界面两侧建立起稳定的内建电场。



热平衡下 P-N 结模型及能带图

P-N 结能带与接触电势差：

在热平衡条件下，结区有统一的 E_F；在远离结区的部位，E_c、E_F、E_v 之间的关系与结形成前状态相同。

从能带图看，N型、P型半导体单独存在时， E_{FN} 与 E_{FP} 有一定差值。

当N型与P型两者紧密接触时，电子要从费米能级高的一方向费米能级低的一方流动，空穴流动的方向相反。同时产生内建电场，内建电场方向为从N区指向P区。

在内建电场作用下， E_{FN} 将连同整个N区能带一起下移， E_{FP} 将连同整个P区能带一起上移，直至将费米能级拉平为 $E_{FN}=E_{FP}$ ，载流子停止流动为止。

在结区这时导带与价带则发生相应的弯曲，形成势垒。势垒高度等于N型、P型半导体单独存在时费米能级之差：

$$qU_D = E_{FN} - E_{FP}$$

得

$$U_D = (E_{FN} - E_{FP}) / q$$

q: 电子电量

U_D : 接触电势差或内建电势

对于在耗尽区以外的状态：

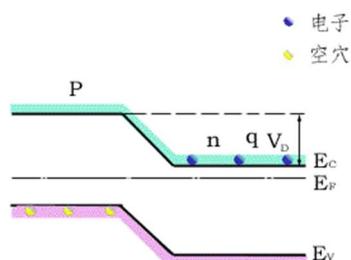
$$U_D = (KT/q) \ln(N_A N_D / n_i^2)$$

N_A 、 N_D 、 n_i : 受主、施主、本征载流子浓度。

可见 U_D 与掺杂浓度有关。在一定温度下，P-N结两边掺杂浓度越高， U_D 越大。禁带宽的材料， n_i 较小，故 U_D 也大。

光照下的P-N结

P-N结光电效应：



光照下光电流的产生动态图

当P-N结受光照时，样品对光子的本征吸收和非本征吸收都将产生光生载流子。但能引起光伏效应的只能是本征吸收所激发的少数载流子。

因P区产生的光生空穴，N区产生的光生电子属多子，都被势垒阻挡而不能过结。只有P区的光生电子和N区的光生空穴和结区的电子空穴对（少子）扩散到结电场附近时能在内建电场作用下漂移过结。

光生电子被拉向N区，光生空穴被拉向P区，即电子空穴对被内建电场分离。这导致在N区边界附近有光生电子积累，在P区边界附近有光生空穴积累。它们产生一个与热平衡P-N结的内建电场方向相反的光生电场，其方向由P区指向N区。此电场使势垒降低，其减小量即光生电势差，P端正，N端负。于是有结电流由P区流向N区，其方向与光电流相反。

实际上，并非所产生的全部光生载流子都对光生电流有贡献。设N区中空穴在寿命 τ_p 的时间内扩散距离为 L_p ，P区中电子在寿命 τ_n 的时间内扩散距离为 L_n 。 $L_n + L_p = L$ 远大于P-N结本身的宽度。故可以认为在结附近平均扩散距离L内所产生的光生载流子都对光电流有贡献。而产生的位置距离结区超过L的电子空穴对，在扩散过程中将全部复合掉，对P-N结光电效应无贡献。

光照下的P-N结电流方程：

与热平衡时比较，有光照时，P-N结内将产生一个附加电流（光电流） I_p ，其方向与P-N结反向饱和电

流 I_0 相同, 一般 $I_p \geq I_0$ 。此时

$$I = I_0 e^{qU/KT} - (I_0 + I_p)$$

令 $I_p = SE$, 则

$$I = I_0 e^{qU/KT} - (I_0 + SE)$$

开路电压 U_{oc} :

光照下的 P-N 结外电路开路时 P 端对 N 端的电压, 即上述电流方程中 $I=0$ 时的 U 值:

$$0 = I_0 e^{qU_{oc}/KT} - (I_0 + SE)$$

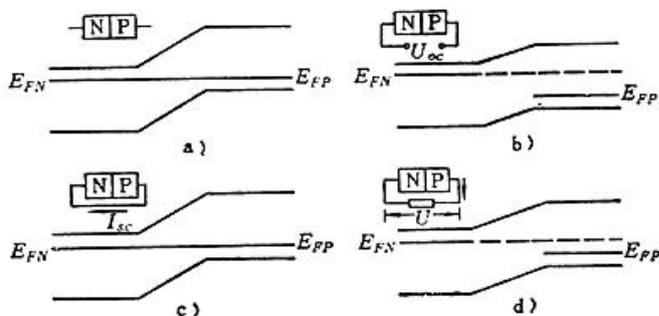
$$U_{oc} = (KT/q) \ln (SE + I_0) / I_0$$

$$\approx (KT/q) \ln (SE / I_0)$$

短路电流 I_{sc} :

光照下的 P-N 结, 外电路短路时, 从 P 端流出, 经过外电路, 从 N 端流入的电流称为短路电流 I_{sc} 。即上述电流方程中 $U=0$ 时的 I 值, 得 $I_{sc} = SE$ 。

U_{oc} 与 I_{sc} 是光照下 P-N 结的两个重要参数, 在一定温度下, U_{oc} 与光照度 E 成对数关系, 但最大值不超过接触电势差 U_0 。弱光照下, I_{sc} 与 E 有线性关系。



不同状态下 P-N 结的能带图

a) 无光照时热平衡态, NP 型半导体有统一的费米能级, 势垒高度为 $qU_0 = E_{FN} - E_{FP}$ 。

b) 稳定光照下 P-N 结外电路开路, 由于光生载流子积累而出现光生电压 U_{oc} , 不再有统一费米能级, 势垒高度为 $q(U_0 - U_{oc})$ 。

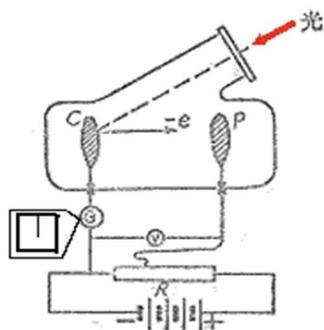
c) 稳定光照下 P-N 结外电路短路, P-N 结两端无光生电压, 势垒高度为 qU_0 , 光生电子空穴对被内建电场分离后流入外电路形成短路电流。

d) 有光照有负载, 一部分光电流在负载上建立起电压 U_f , 另一部分光电流被 P-N 结因正向偏压引起的正向电流抵消, 势垒高度为 $q(U_0 - U_f)$ 。

6.5.3 光电发射效应

金属或半导体受光照时, 如果入射的光子能量 $h\nu$ 足够大, 它和物质中的电子相互作用, 使电子从材料表面逸出现象, 也称为外光电效应。它是真空光电器件光电阴极的物理基础。

外光电效应的两个基本定律



光电发射定律解释图

1. 光电发射第一定律——斯托列托夫定律：

当照射到光阴极上的入射光频率或频谱成分不变时，饱和光电流（即单位时间内发射的光电子数目）与入射光强度成正比：

$$I_k = S_k F_0$$

I_k ：光电流

S_k ：光强

F_0 ：该阴极对入射光线的灵敏度

2. 光电发射第二定律——爱因斯坦定律

光电子的最大动能与入射光的频率成正比，而与入射光强度无关：

$$E_{\max} = (1/2) m v_{\max}^2 = h \nu - h \nu_0 = h \nu - A$$

E_{\max} ：光电子的最大初动能。

h ：普朗克常数。

ν_0 ：产生光电发射的极限频率，频率阈值。

A ：金属电子的逸出功（从材料表面逸出时所需的最低能量），单位 eV，与材料有关的常数，也称功函数。

入射光子的能量至少要等于逸出功时，才能发生光电发射。

波长阈值：

$$\begin{aligned} \lambda_0 = c / \nu_0 &\leq hc / A = 1.24 (\mu\text{m} \cdot \text{eV}) / A \\ &= 1240 / A (\text{nm}) \end{aligned}$$

当入射光波长大于 λ_0 时，不论光强如何，以及照射时间多长，都不会有光电子产生。要用红外光（ $\lambda > 0.76 \mu\text{m}$ ）发射电子，必须寻求低于 1.8eV 的低能阈值材料。

光电发射的基本过程

光电发射大致可分三个过程：

- 1) 光射入物体后，物体中的电子吸收光子能量，从基态跃迁到能量高于真空能级的激发态。
- 2) 受激电子从受激地点出发，在向表面运动过程中免不了要同其它电子或晶格发生碰撞，而失去一部分能量。
- 3) 达到表面的电子，如果仍有足够的能量足以克服表面势垒对电子的束缚（即逸出功）时，即可从表面逸出。



光电发射过程图

好的光电发射材料应该是：（1）对光的吸收系数大，以便体内有较多的电子受到激发；

（2）受激电子最好是发生在表面附近，这样向表面运动过程中损失的能量少；

（3）材料的逸出功要小，使到达真空界面的电子能够比较容易地逸出；（4）另外，作为光电阴极，其材料还要有一定的电导率，以便能够通过外电源来补充因光电发射所失去的电子。

金属的光电发射

金属反射掉大部分入射的可见光（反射系数达 90%以上），吸收效率很低。光电子与金属中大量的自由电子碰撞，在运动中丧失很多能量。只有很靠近表面的光电子，才有可能到达表面并克服势垒逸出，即金属中光电子逸出深度很浅，只有几 nm，而且金属逸出功大多为大于 3eV（ $\lambda > 410\text{nm}$ ）的可见光来说，很难产生光电发射，只有铯（2eV 逸出功）对可见光最灵敏，故可用于光阴极。但纯金属铯量子效率很低，小于 0.1%，因在光电发射前两个阶段能量损耗太大。

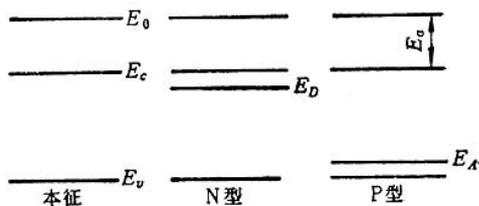
半导体的光电发射

半导体光电逸出参量：

电子亲和势：(Electron Affinity)

导带底上的电子向真空逸出时所需的最低能量，数值上等于真空能级（真空中静止电子能量）与导带底能级 E_c 之差。

它有表面电子亲和势 E_a 与体内电子亲和势 E_{ae} 之分。 E_a 是材料的参量，与掺杂、表面能带弯曲等因素无关。而 E_{ae} 不是材料参量，可随表面能带弯曲变化。



电子亲和势

电子逸出功：电子逸出功是描述材料表面对电子束缚强弱的物理量，在数量上等于电子逸出表面所需的最低能量，也可以说是光电发射的能量阈值。

金属有大量的自由电子，没有禁带，费米能级以下基本上为电子所填满，费米能级以上基本上是空的，表面能带受内外电场影响很小， E_F 只决定于材料。所以金属的电子逸出功定义为 $T=0\text{K}$ 时真空能级与 E_F 之差，它是材料的参量，可以用来作为光电发射的能量阈值。

半导体自由电子较少，且有禁带，费米能级 E_F 一般都在禁带当中，且随掺杂和内外电场变化，所以真空能级与费米能级之差不是材料参量。

半导体电子逸出功定义为 $T=0\text{K}$ 时真空能级与电子发射中心的能级之差，而电子发射中心的能级有的是价带顶，有的是杂质能级，有的是导带底，情况复杂，因此对于半导体很少用电子逸出功的概念。

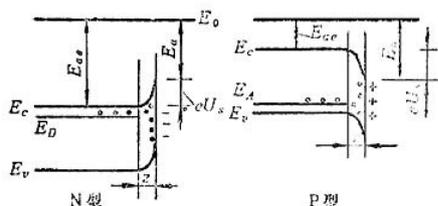
由于电子逸出功不管从哪里算起，其中都包含有亲和势（真空能级与导带底之差），因此为了表示光电

发射的难易，使用亲和势的概念比使用逸出功的概念更有实际意义。所以，对于半导体一般不用逸出功的概念，而用电子亲和势的概念。为了表示光电发射的能量阈值，许多资料都是按真空能级与价带顶之差（亲和势加上禁带宽度）来计算。

表面能带弯曲：

半导体无界时，能带结构是平直的，有界时表面处破坏了晶格排列周期性（势场）而且表面易氧化及被杂质污染，因而在禁带中引入附加能级（表面能级），由于表面能级的存在，在表面处引起能带弯曲。

表面能带弯曲，对于体内的光电子发射是有影响的。因为表面电子亲和势 E_a 是材料的参量，它不随表面能带弯曲而变化。而体内电子亲和势则要随着表面能带弯曲而增减。



表面能带弯曲

对于 N 型半导体，施主能级上的电子跃迁到表面能级时，半导体表面将产生一个负的空间电荷区。而距离表面稍远一点的体内则分布有等量正的体电荷，因此表面能带向上弯。向上弯的程度，可用表面势垒 eU_s 表示， e 为电子电荷， U_s 为表面势，在数值上等于体内与表面的电势差。对于 N 型半导体来说，因表面能带向上弯，体内的电子亲和势 E_{ae} 要比表面能带不发生弯曲时增加一个势垒高度 eU_s ，使得体内光电子发射变得更困难。

对于 P 型半导体，情况正好相反。表面能级中能量高于受主能级的电子有的要跃迁到受主能级上，于是半导体表面即产生一个正的空间电荷区，距离表面稍远一点的体内则分布有等量的负电荷，因此表面能带向下弯。特别是 P 型半导体表面吸附有带正电性的原子（例如铯原子）或 N 型材料的时候，表面上偶电层正电性在外，能带弯曲就更厉害。能带弯曲的程度也用表面势垒 eU_s 表示。表面能带向下弯，使得体内电子亲和势比能带不发生弯曲时减少一个势垒高度 eU_s 。这样，这种表面能带弯曲对于体内的光电子发射十分有利。

因此，现在各种实用光电阴极几乎全是用 P 型半导体材料为衬底，然后在它的表面上再涂上带正电性的金属或 N 型材料而制成的。这样，就能得到向下弯曲的表面能带，减小逸出功。如果再能使带弯曲足够小（带曲区宽度为 z ），以至比材料吸收系数的倒数还小得多（ $z \leq 1/a$ ）时，则可以使光电子发射的主要部位来自于体内。这时，量子效率要比单纯能带弯曲大得多。另外，强 P 型半导体的费米能级十分靠近于价带，这可使热电子发射（暗电流）较小。

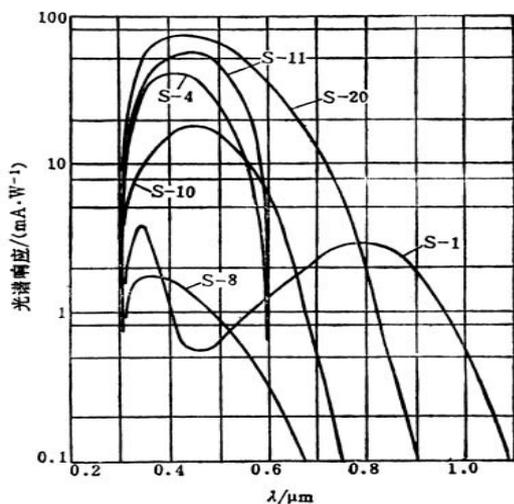
实用光电阴极：分为常规光电阴极、负电子亲和势阴极两大类。

常规光电阴极：用于常用真空光电器件中。根据国际电子工业协会规定，按其出现的先后顺序，以 S 为字头排成序号。

常规光电阴极编号及特性表

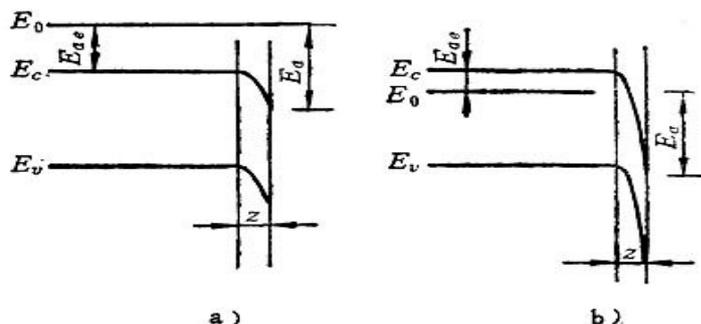
光谱响应 编号	光电发射材料	窗材料	工作方式 半透为T 反射为R	峰值波长 $\lambda_{max}/\mu m$	积分灵敏度 $\mu A \cdot lm^{-1}$	λ_{max} 处辐射 响应度 $/(mA \cdot W^{-1})$	λ_{max} 处量子 效率 (%)	25℃下暗电流 $\mu A \cdot mm^{-2}$
S-1	Ag-O-Cs	石灰玻璃	T, R	0.800	30	2.8	0.43	900
S-3	Ag-O-Rb	石灰玻璃	R	0.42	6.5	1.8	0.53	—
S-4	Cs-Sb	石灰玻璃	R	0.400	40	40	12.4	0.2
S-5	Cs-Sb	9741玻璃	R	0.340	40	60	18.2	0.3
S-8	Cs-Sb	石灰玻璃	R	0.365	3	2.3	0.78	0.13
S-9	Cs-Sb	7052玻璃	T	0.480	30	20.5	5.3	0.3
S-10	Ag-Bi-O-Cs	石灰玻璃	T	0.450	40	20	5.5	70
S-11	Cs-Sb	石灰玻璃	T	0.440	70	66	15.7	3
S-13	Cs-Sb	熔凝玻璃	T	0.440	60	48	13.5	4
S-14	Ge	石灰玻璃	—	1.500	12400 ^①	520 ^①	43 ^①	—
S-16	CdSe	石灰玻璃	—	0.730	—	—	—	—
S-17	Cs-Sb	石灰玻璃	R	0.490	125	83	21	1.2
S-19	Cs-Sb	石灰玻璃	R	0.380	40	65	24.4	0.3
S-20	Na-K-Cs-Sb	熔凝玻璃	R	0.420	150	64	18.8	0.3
S-21	Cs-Sb	9741玻璃	T	0.440	30	23.4	6.6	4
S-23	Rb-Tc	熔凝石英	T	0.240	—	4	2	0.001
S-24	K-Na-Sb	7056玻璃	T	0.380	45	67	21.8	0.003
S-25	Na-K-Cs-Sb	石灰玻璃	T	0.420	200	43	12.7	1

① 带45V的起偏电压。



常用光电阴极光谱特性曲线

负电子亲和势 (NEA, Negative Electron Affinity) 阴极:



a) 常规光电发射体物理模型

b) 负电子亲和势光电发射体物理模型

采用特殊工艺，例如在重掺杂 P 型硅表面涂一薄层 CsO_2 ，可形成 NEA 材料。负电子亲和势是指体内的有效电子亲和势，而不是指表面电子亲和势。

NEA 发射体和常规光电发射体的表面，电子状态是类似的，导带底上的电子能量都低于真空能级，其差值为 E_a 。但是，两者体内电子能量则不同。NEA 发射体导带底的电子能量高于真空能级，而常规发射体电子亲和势仍是正的。

NEA 阴极的量子效率高于正电子亲和势阴极，可从其光电发射过程进行分析。价带中的电子吸收光子能量，跃迁到导带底以上，成为热电子（受激电子能量超过导带底的电子）。在向表面运动的过程中，由于碰撞散射而发生能量损失，故很快就落到导带底而变成冷电子（能量恰好等于导带底的电子）。热电子的平均寿命非常短，约 $10^{-14} \sim 10^{-12} \text{s}$ 。如果在这么短的时间内能够运动到真空界面，自然能逸出。但是热电子的逸出深度只有几十纳米，绝大部分电子来不及到达真空界面，就已经落到导带底变成冷电子了。冷电子的平均寿命比较长，约 $10^{-9} \sim 10^{-8} \text{s}$ ，其逸出深度可达 1000 纳米。因为体内冷电子能量仍高于真空能级，所以它们运动到真空界面时，可以很容易地逸出。因此 NEA 量子效率比常规发射体高得多。

第七章 光电探测器

光电探测器件是光辐射探测器件中的一大类，它利用各种光电效应，使入射辐射强度转换成电学信息或电能。

7.1 真空光电管

真空光电管由玻壳、光电阴极和阳极三部分组成。为了防止氧化，将管内抽成真空。光电阴极即半导体光电发射材料，涂于玻壳内壁，受光照时，可向外发射光电子。阳极是金属环或金属网，置于光电阴极的对面，加正的高电压，用来收集从阴极发射出来的电子。



真空光电管构造示意图

特点：光电阴极面积大，灵敏度较高，一般积分灵敏度可达 $20\sim 200\ \mu\text{A}/\text{lm}$ ；暗电流小，最低可达 10^{-14}A ；光电发射弛豫过程极短。

缺点：真空光电管一般体积都比较大、工作电压高达百伏到数百伏、玻壳容易破碎等。

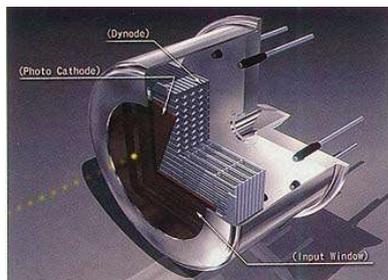
7.2 光电倍增管

光电倍增管

光电倍增管由光窗、光电阴极、电子光学系统、电子倍增系统和阳极五个主要部分组成。

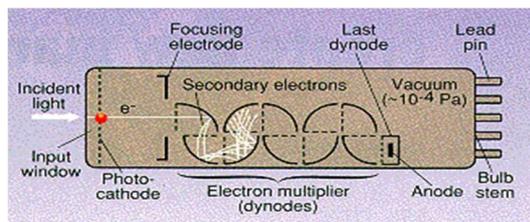
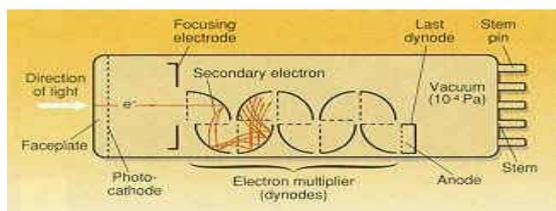
1. 光窗

光窗分侧窗式和端窗式两种，它是入射光的通道。一般常用的光窗材料有钠钙玻璃、硼硅玻璃、紫外玻璃、熔凝石英和氟镁玻璃等。由于光窗对光的吸收与波长有关，波长越短吸收越多，所以倍增管光谱特性的短波阈值决定于光窗材料。



2. 光电阴极

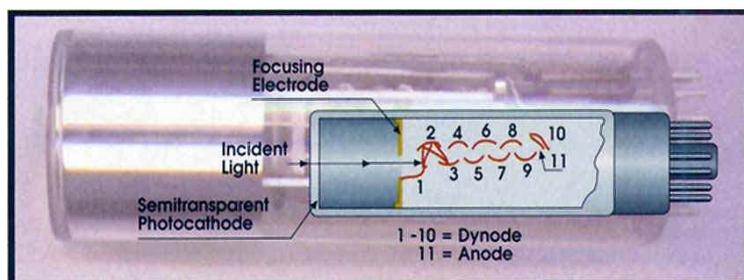
光电阴极多是由化合物半导体材料制作，它接收入射光，向外发射光电子。所以倍增管光谱特性的长波阈值决定于光电阴极材料，同时对整管灵敏度也起着决定性作用。



3. 电子光学系统

电子光学系统是适当设计的电极结构，使前一级发射出来的电子尽可能没有散失地落到下一个倍增极上，也就是使下一级的收集率接近于 1；并使前一级各部分发射出来的电子，落到后一级上所经历的时间尽可能相同，即渡越时间零散最小。

4. 倍增系统



倍增系统是由许多倍增极组成的综合体，每个倍增极都是由二次电子倍增材料构成，具有使一次电子倍增的能力。因此倍增系统是决定整管灵敏度最关键的部分。

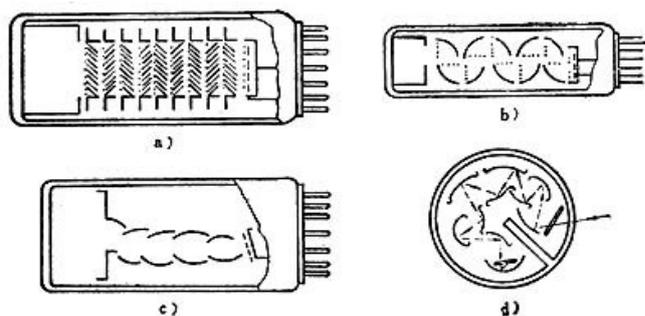
倍增极材料大致可分以下四类：

- 1) 含碱复杂面主要是银氧铯和铯铯两种，它们既是灵敏的光电发射体，也是良好的二次电子发射体。
- 2) 氧化物型，主要是氧化镁。
- 3) 合金型，主要是银镁、铝镁、铜镁、镍镁、铜铍等合金。
- 4) 负电子亲合势发射体。

这几类材料在低电压下有大的 δ 值，以便整管工作电压不致于过高；热发射小，以便整管的暗电流和噪声小；二次电子发射稳定，以便温度较高或一次电流较大时，长时间工作 δ 不下降；而且容易制备。

倍增极结构形式		特点
聚焦型	直瓦片式	极间电子渡越时间零散小，但绝缘支架可能积累电荷而影响电子光学系统的稳定性。
	圆瓦片式	结构紧凑，体积小，但灵敏度的均匀性差些。
非聚焦型	百叶窗式	工作面积大，与大面积光电阴极配合可制成探测弱光的倍增管，但极间电压高时，有的电子可能越级穿过，收集率较低，渡越时间零散较大。
	盒栅式	收集率较高（可达 95%），结构紧凑，但极间电子渡越时间零散较大。

所谓聚焦不是指使电子束会聚于一点，而是指电子从前一级倍增极飞向后一级倍增极时，在两电极间的电子运动轨迹，可能有交叉。非聚焦则是指在两电极间的电子运动轨迹是平行的。



各种倍增极的结构形式

a) 百叶窗式 b) 盒栅式 c) 直瓦片式 d) 圆瓦片式

5. 阳极

阳极是采用金属网作的栅网状结构，把它置于靠近最末一级倍增极附近，用来收集最末一级倍增极发射出来的电子。

光电倍增管的主要参量与特性

光电倍增管的主要参量与特性是区分管子质量好坏的基本依据。分为基本参数（静态参数）、应用参数（动态参数）、运行特性（例行特性）。

基本参数与管子工作原理、结构特征、材料性质、制造工艺有关。它包括灵敏度、量子效率、增益、暗电流、光谱响应等。

应用参数与管子应用方法和探测对象有关，反映某种应用的特殊要求。它包括闪烁计数中的脉冲幅度分辨率、噪声能当量、计数坪特性；光子计数中的暗噪声计数、单电子分辨率、峰谷比；快速光脉冲测量中的上升时间、半高宽、渡越时间、时间分辨率等。

运行特性与管子运行条件、运行环境有关。它表征管子承受的外部条件和使用极限，包括稳定性、温度特性、最大线性电流、抗电磁干扰特性、抗冲击振动特性等。

1. 灵敏度

倍增管灵敏度有阴极灵敏度与阳极灵敏度之分。每一种灵敏度对于入射光，又都有光谱灵敏度（对于单色光）与积分灵敏度（对于多色光或全色光）之分。

测试阴极灵敏度时，以阴极为一极，其它倍增极和阳极都连到一起为另一极，相对于阴极加 100~300V 直流电压，照射到光电阴极上的光通量约为 $10^{-2} \sim 10^{-5} \text{lm}$ 。

测试阳极灵敏度时，各倍增极和阳极都加上适当电压，因为阳极灵敏度是整管参量，与整管所加电压有关，所以必须注明整管所加电压。

积分灵敏度与测试光源的色温有关，一般用色温为 2856K 的白炽钨丝灯（A 光源）。色温不同时即使测试光源的波长范围相同，各单色光在光谱分布中的组分不同时，所得的积分灵敏度也不同。

灵敏度		公式	说明
阴极灵敏度	阴极光谱灵敏度	$S_K(\lambda) = I_{K\lambda} / \Phi_\lambda$	S: 灵敏度 λ : 波长 I: 光电流 Φ : 光通量 下标 K: 阴极 下标 A: 阳极
	阴极积分灵敏度	$S_K = I_K / \Phi$	
阳极灵敏度	阳极光谱灵敏度	$S_A(\lambda) = I_{A\lambda} / \Phi_\lambda$	
	阳极积分灵敏度	$S_A = I_A / \Phi$	

2. 电流增益 M

阳极电流与阴极电流之比，或阳极灵敏度与阴极灵敏度之比，即

$$M = I_A / I_K = S_A / S_K$$

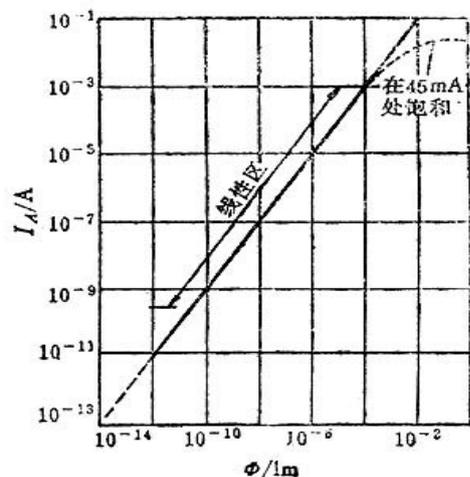
若倍增管有 n 个倍增极，并且每个倍增极的倍增系数均相等，则

$$M = \delta^n$$

因为 δ 是电压的函数，所以 M 也是电压的函数。

3. 光电特性

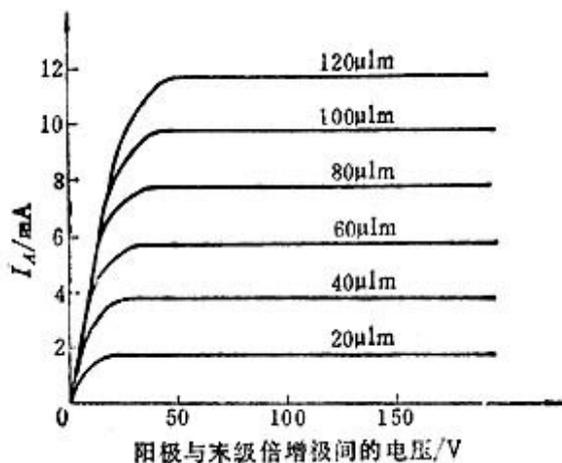
阳极光电流与入射于光电阴极的光通量之间的函数关系，称为倍增管的光电特性。对于模拟量测量，必须选取能保证光电流与光照在大范围内保持线性关系的那些型号的光电倍增管（工程上一般取特性偏离于直线 3% 作为线性区的界限）。



光电特性图

4. 伏安特性

光电倍增管的伏安特性曲线分为阴极伏安特性曲线（阴极电流与阴极电压之间的关系）与阳极伏安特性曲线（阳极电流与阳极和最末一级倍增极之间电压的关系）。在电路设计时，一般使用阳极伏安特性曲线来进行负载电阻、输出电流、输出电压的计算。



阳极伏安特性曲线

5. 暗电流

在各电极都加上正常工作电压并且阴极无光照情况下阳极的输出电流。它限制了可测直流光通量的最小值，同时也是产生噪声的重要因素，是鉴别管子质量的重要参量。应选取暗电流较小的管子。

光电倍增管中产生暗电流的因素较多，例如，阴极和靠近于阴极的倍增极的热电子发射；阳极或其它电极的漏电；由于极间电压过高而引起的场致发射；光反馈，以及窗口玻璃中可能含有的少量的钾、镭、钷等放射性元素蜕变产生的 β 粒子，或者宇宙线中的 μ 介子穿过光窗时产生的契伦柯夫光子等都可能引起暗电流。

6. 噪声与噪声等效功率

光电倍增管噪声主要是指由倍增管本身引起的输出偏离于平均值的起伏，主要来源是光电阴极、光电发射的随机性和各倍增极二次电子发射的随机性，同时也与背景光或信号光中的直流分量有关。

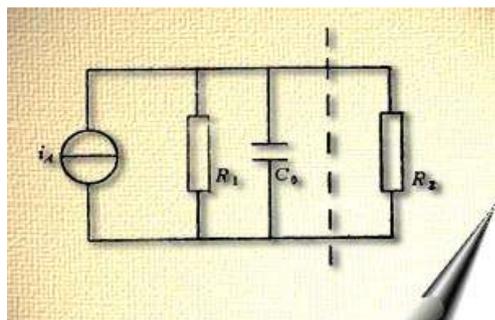
噪声等效功率（NEP）表述倍增管阳极信号与噪声有效值之比等于 1 时，入射于倍增管光电阴极的光功率（通量）的有效值。即

$$I_A/I_{nA}=1 \text{ 时, } NEP=I_{nA}/S_A$$

它是倍增管可能探测到的信号光功率（通量）的最小值。

微变等效电路

从倍增管阳极伏安特性曲线来看，最大光通量所对应的曲线拐点以右，基本上是平直均匀分布的，一般使用倍增管也都是利用这一区域的特性，因此在交流微变电路中可以把倍增管看成是电流源，并考虑阳极电路的电容效应。



光电倍增管的交流微变等效电路

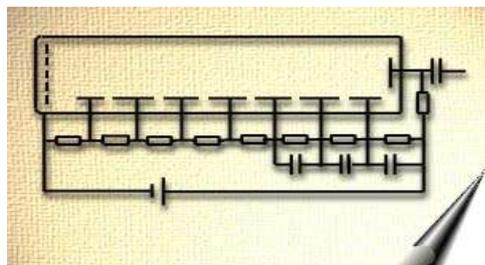
i_A -阳极电流 C_0 -等效电容 R_1 -直流负载

R_2 -下一级放大器的输入电阻

光电倍增管的使用

1. 供电电路

倍增管各电极要求直流供电，从阴极开始至各级的电压要依次升高，一般多采用电阻链分压办法来供电。一般情况下，各级电压均相等，约 80~100V，总电压约 1000~1300V。



光电倍增管供电电路图

1) 电源电压稳定性的要求

电源电压稳定性要求较高。如果电源电压不稳，会引起许多参量的变化，特别是电流增益变化，从而直接影响输出特性。目前已有光电倍增管专用的电源稳压块。

2) 电阻链分压电阻的确定

若电阻链为均匀分压，则每个分压电阻的阻值应相等。因倍增管中的电流与电阻链中的电流是并联关系，要保证阳极电流最大时流过电阻链的电流基本不变，这就要求流过电阻链的电流 I_R 至少要比阳极最大的平均电流 I_{Am} 大 10 倍以上。

一般说， I_R 越大（即 $R=U_b/I_R$ 越小）对稳定极间电压 U_b 越有利。但 I_R 也不能太大，因为 I_R 太大会增大电阻的功耗，加重电源负担。当 U_b 给定后，分压电阻 R 的最大值应取决于阳极的最大平均电流， R 最小值应取决于高压电源输出的功率。

以上讨论的是均匀分压情形，实际各倍增极间电压也可以不相等。这样，有可能使某单项指标得到提

高。例如，要提高管子的时间特性，可适当增大靠近于阴极的几级倍增极间的电压。因为电压高，可以提高收集率，减小极间电子渡越时间零散。如果第一级倍增极对阴极的收集率提高一点，整管的时间特性就会有较大改善，这点对于测量脉冲光是很重要的。但是，在选定的工作状态下，阳极灵敏度往往要降低，这是因为要使整管工作稳定，总电压不宜过高，在总电压基本保持不变的情况下，前几级电压高了，后几级电压就得相应降低，因此将导致阳极灵敏度下降。同理，要使阳极输出的线性范围宽一些，可适当增大阳极与最末一级倍增极的电压，但这时阳极灵敏度也要下降。

3) 并联电容的确定

倍增管的输出电流主要是来自于最后几级，探测脉冲光时，为了不使阳极脉动电流引起极间电压发生大的变化，常在最后几级的分压电阻上并联电容器。

4) 接地方式

倍增管供电电路与其后续信号处理电路必须要有一个共用的参考电位，即接地点。倍增管的接地方式有两种，即阴极接地或阳极接地。

阴极接地的特点是，便于屏蔽，光、磁、电的屏蔽罩可以跟阴极靠得近些，屏蔽效果好；暗电流小，噪声低。但这时阳极要处于正高压，会导致寄生电容大，匹配电缆连接复杂，特别是后面若接直流放大器，整个放大器都处于高电压，不利于安全操作；如果后面接交流放大器，则必须接一个耐压很高的隔直电容器，而一般耐压很高的电容器体积大而且价格高。

阳极接地的特点是，便于跟后面的放大器相接，操作安全，后面不仅可以通过一个低压耦合电容与交流放大器相接，也可以直接与直流放大器相接。但这时阴极要处于负高压，屏蔽罩不能跟阴极靠得很近，至少要间隔 1~2cm，因此屏蔽效果差一些，暗电流和噪声都比阳极接地时大，而且整个倍增管装置的体积也要大些。

Anywhere science goes, BURLE photomultiplier tubes take the lead.

From 450 kilometers above the earth, NASA's Gamma Ray Observatory is increasing our understanding of the universe. And on three of its telescopes, BURLE photomultiplier tubes are the primary components for gamma radiation detection.

BURLE photomultipliers have been an integral part of space science for years, and they were on board the Mars probe. Back on earth, they're used in high-energy physics experiments (including

many planned for the SSC), calorimeter experiments, geophysical exploration and medical analytical systems.

If your science calls for photomultiplier tubes, let BURLE help you take the lead. Call today in the United States at 1-800-366-2875 or fax: 717-295-6096. Visit us on the Web at www.burle.com or email us at burles@s@success.net

BURLE
ELECTRON TUBES

Photo: NASA

2. 使用注意事项

1) 使用前应了解器件的特性。真空光电器件的共同特点是灵敏度高、惰性小、供电电压高、采用玻璃外壳、抗震性差。

2) 使用时不宜用强光照。光照过强时，光电线性会变差而且容易使光电阴极疲劳（轻度疲劳经一段时间可恢复，重度疲劳不能恢复），缩短寿命。

3) 工作电流不宜过大。工作电流大会烧毁阴极面，或使倍增级二次电子发射系数下降，增益降低，光电线性变差，缩短寿命。

4) 用来测量交变光时，负载电阻不宜很大，因为负载电阻和管子的等效电容一起构成电路的时间常数，若负载电阻较大，时间常数就变大，频带将变窄。

部分光电倍增管特性

型号	倍增极级数	光窗材料	阴极类型	光谱范围 nm	峰值波长 /nm	阴极灵敏度 /($\mu\text{A}\cdot\text{Im}^{-2}$)	一定电压下的阳极灵敏度		暗电流 /nA	主要用途
							电压 /V	阳极灵敏度 /($\text{A}\cdot\text{Im}^{-2}$)		
GDB-106	9	透紫玻璃	Sb-K-Cs	200~700	400	30	360	30	7	光度测量
GDB-142	9	硼硅玻璃	Sb-K-Cs	300~700	400	30	1100	10	30	传真机
GDB-146	8	透紫玻璃	Sb-K-Cs	200~700	400	20	1100	10	80	光度测量
GDB-151	9	石英	Sb-Na-K-Cs	185~850	400	20	800	1	20	原子吸收分光光度计
GDB-152	9	石英	Fe-Cs	200~300	235	20mA/W	1000	1000A/W	7	测汞仪
GDB-153	10	硼硅玻璃	Ga-As(Cs)	200~910	340	150	1250	20	2	激光接收器
GDB-235	8	钠钙玻璃	Sb-Cs	300~650	400	40	1100	10	60	闪烁计数器
GDB-239	11	钠钙玻璃	Ag-O-Cs	400~1200	800	10	1500	1	1100	激光接收器
GDB-333	14	硼硅玻璃	Sb-Na-K-Cs	300~850	420	70	2200	500	500	核物理研究
GDB-404	9	硼硅玻璃	Sb-Na-K-Cs	300~850	450	90	1250	10	2	激光接收器
GDB-413	11	硼硅玻璃	Sb-K-Cs	300~700	400	40	1250	100	10	分光光度计, 扫描电镜
GDB-415	11	硼硅玻璃	Sb-Na-K	300~650	420	20	2000	10	30	放射性测井仪
GDB-423	11	硼硅玻璃	Sb-Na-K-Cs	300~850	420	60	1500	100	400	激光接收器
GDB-526	11	硼硅玻璃	Sb-K-Cs	300~700	400	30	950	10	50	同位素扫描仪
GDB-546	11	硼硅玻璃	Sb-Na-K-Cs	300~850	420	70	1800	200	100	激光接收器
GDB-567	11	硼硅玻璃	Sb-K-Cs	300~700	400	30	1000	10	50	伽玛照相机
GDB-576	11	硼硅玻璃	Sb-Cs	300~650	400	20	1200	10	500	同位素扫描仪

7.3 光电导器件

利用半导体光电导效应制成的器件。典型的光电导器件为光敏电阻。

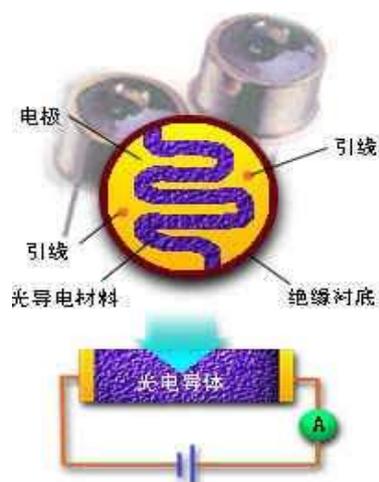
光敏电阻的原理是半导体的光电导效应，只有能量 ($h\nu$) 大于材料禁带宽度 (E_g) 的光子，才能使材料产生光电导效应。每一种半导体或绝缘体都有一定的光电导效应，但只有其中一部分材料经过特殊处理，掺进适当杂质，才有明显的光电导效应。现在使用的光电导材料有II-VI族、III-V族化合物，硅、锗等，以及一些有机物。

常用光电导材料

光电导器件材料	禁带宽度 /eV	光谱响应范围/nm	峰值波长/nm
硫化镉 (CdS)	2.45	400~800	515~550
硒化镉 (CdSe)	1.74	680~750	720~730
硫化铅 (PbS)	0.40	500~3000	2000
碲化铅 (PbTe)	0.31	600~4500	2200

硒化铅 (PbSe)	0.25	700~5800	4000
硅 (Si)	1.12	450~1100	850
锗 (Ge)	0.66	550~1800	1540
锑化铟 (InSb)	0.16	600~7000	5500
砷化铟 (InAs)	0.33	1000~4000	3500

光敏电阻的结构是在一块光电导体两端加上电极，贴在硬质玻璃、云母、高频瓷或其它绝缘材料基板上，两端接有电极引线，封装在带有窗口的金属或塑料外壳内。光敏面作成蛇形，电极作成梳状是因为这样可以保证有较大的受光表面，也可以减小电极之间距离，从而既可减小极间电子渡越时间，也有利于提高灵敏度。



光敏电阻

光敏电阻的主要特性

1. 光电导灵敏度 S_g

按灵敏度定义（响应量与输入量之比），可得

$$S_g = G_p / E$$

G_p : 光电导，单位为西门子 $S (\Omega^{-1})$ 。

E : 照度，单位为勒克斯 (lx)。

S_g : 单位为西门子/勒克斯 (S/lx) 或 Sm^2/W 。

2. 光电特性

光电流与照度的关系称为光电特性。光敏电阻光电特性如下：

$$I_p = S_g E^\gamma U^\alpha$$

I_p : 光电流

E : 照度

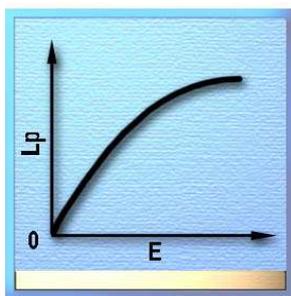
α : 电压指数

S_g : 光电导灵敏度

U : 光敏电阻两端所加的电压

γ : 光照指数

γ 与材料和入射光强弱有关，对于硫化镉光电导体，在弱光照下 $\gamma = 1$ ，在强光照下 $\gamma = 1/2$ ，一般 $\gamma = 0.5 \sim 1$ 。 α 与光电导体和电极材料之间的接触有关，欧姆接触时 $\alpha = 1$ ，非欧姆接触时 $\alpha = 1.1 \sim 1.2$ 。



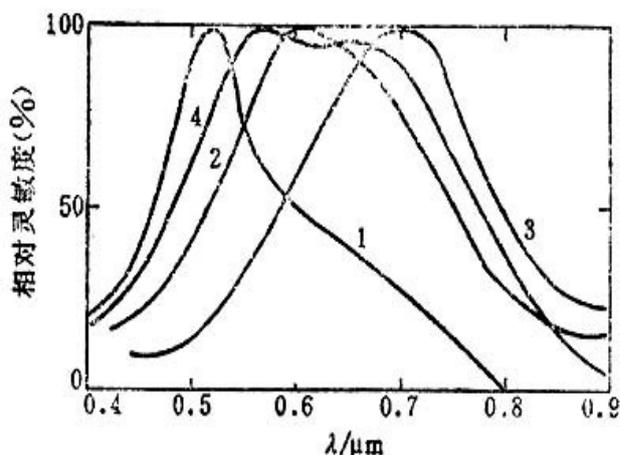
硫化镉光敏电阻的光电特性曲线

由图可见，硫化镉光敏电阻在弱光照下， I_p 与 E 具有良好的线性关系，在强光照下则为非线性关系，其它光敏电阻也有类似的性质。如果电压指数 $\alpha = 1$ ，在弱光照下有

$$I_p = S_g E U$$

3. 光谱特性

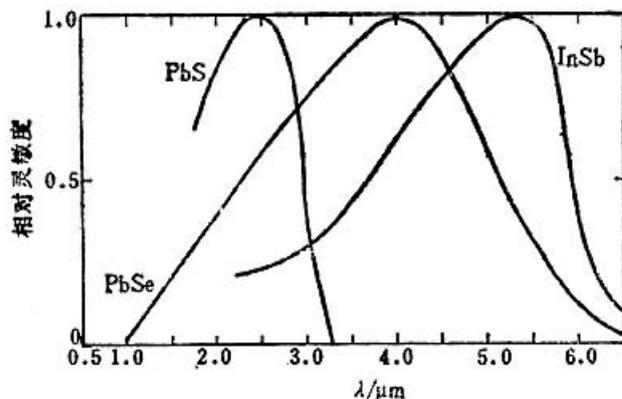
光谱特性多用相对灵敏度与波长的关系曲线表示。从这种曲线中可以直接看出灵敏范围、峰值波长位置和各波长下灵敏度的相对关系。



在可见光区灵敏的几种光敏电阻的光谱特性曲线

1-硫化镉单晶 2-硫化镉多晶 3-硒化镉多晶

4-硫化镉与硒化镉混合多晶



在红外区灵敏的几种光敏电阻的光谱特性曲线

由图可见，硫化镉单晶、硫化镉与硒化镉混合多晶，硫化镉多晶、硒化镉多晶等几种光敏电阻的光谱特性曲线覆盖了整个可见光区，峰值波长在 515~600nm 之间。这与人眼的光谱光视效率 $V(\lambda)$ 曲线的范围和峰值波长 (555nm) 是很接近的，因此可用于与人眼有关的仪器，例如照相机、照度计、光度计等。不过它

们的形状与 $V(\lambda)$ 曲线还不完全一致。如直接使用，与人的视觉还有一定的差距，所以必须加滤光片进行修正，使其特性曲线与 $V(\lambda)$ 曲线完全符合，这样即可得到与人眼视觉相同的效果。

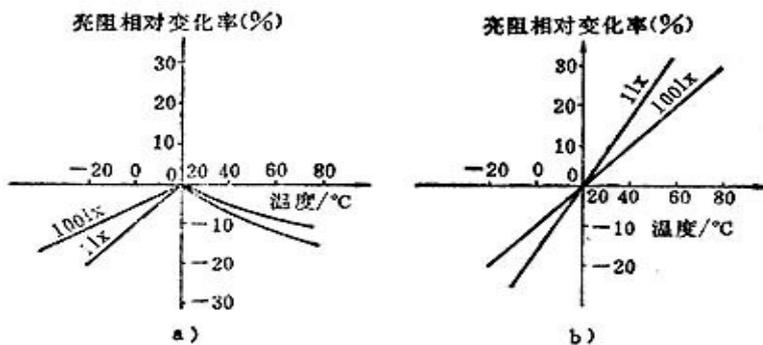
4. 温度特性

光敏电阻的温度特性很复杂，在一定的照度下，亮电阻的温度系数 α 有正有负，

$$\alpha = (R_2 - R_1) / R_1 (T_2 - T_1)$$

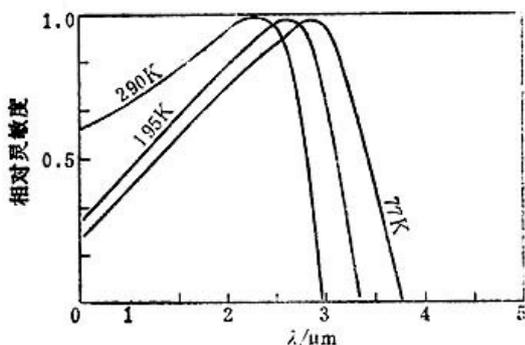
R_1 、 R_2 分别为与温度 T_1 、 T_2 相对应的亮电阻。

温度对光谱响应也有影响。一般说，光谱特性主要决定于材料，材料的禁带宽度越窄则对长波越敏感，但禁带很窄时，半导体中热激发也会使自由载流子浓度增加，使复合运动加快，灵敏度降低。因此，采取冷却灵敏面的办法来提高灵敏度往往是很有有效的。



硫化镉光敏电阻的温度特性曲线

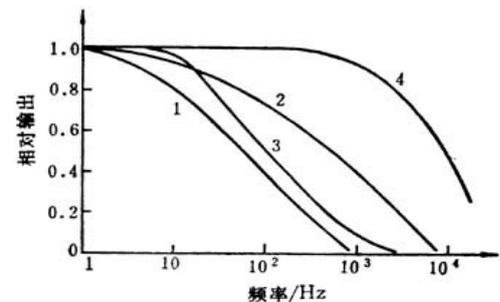
a) 硫化镉单晶 b) 硫化镉多晶



硫化铅光敏电阻在冷却情况下相对光谱灵敏度的变化

5. 频率特性

光敏电阻是依靠非平衡载流子效应工作的，非平衡载流子的产生与复合都有一个时间过程，这个时间过程在一定程度上影响了光敏电阻对变化光照的响应。光敏电阻采用交变光照时，其输出将随入射光频率的增加而减小。

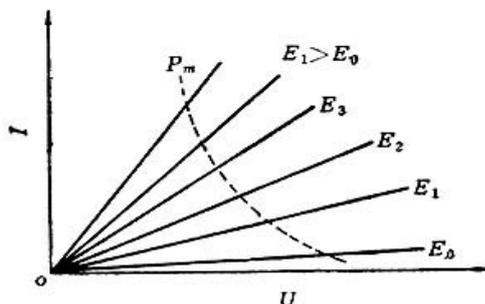


几种光敏电阻的频率特性曲线

1-硒 2-硫化镉 3-硫化铊 4-硫化铅

6. 伏安特性

在一定的光照下，加到光敏电阻两端的电压与流过光敏电阻的亮电流之间的关系称为光敏电阻的伏安特性，常用曲线表示。



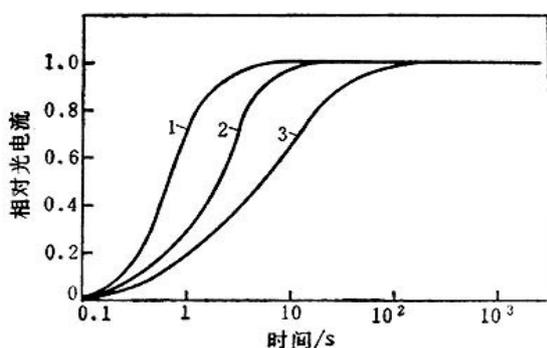
光敏电阻的伏安特性曲线

图中的虚线为额定功耗线。使用光敏电阻时，应不使电阻的实际功耗超过额定值。从图上来说，就是不能使静态工作点居于虚线以内的区域。按这一要求，在设计负载电阻时，应不使负载线与额定功耗线相交。

7. 前历效应

前历效应是指光敏电阻的时间特性与工作前“历史”有关的一种现象。前历效应有暗态前历与亮态前历之分。

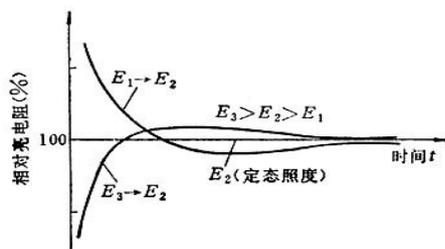
暗态前历效应是指光敏电阻测试或工作前处于暗态，当它突然受到光照后表现为暗态前历越长，光电流上升越慢，其效应曲线如下图所示。一般，工作电压越低，光照度越低，则暗态前历效应就越重。



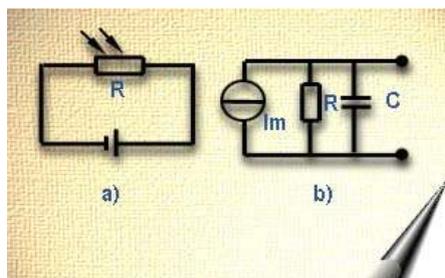
硫化镉光敏电阻的暗态前历效应曲线

- 1-黑暗放置 3 分钟后
- 2-黑暗放置 60 分钟后
- 3-黑暗放置 24 小时后

亮态前历效应是指光敏电阻测试或工作前已处于亮态，当照度与工作时所要达到的照度不同时，所出现的一种滞后现象，其效应曲线如下图所示。一般，亮电阻由高照度状态变为低照度状态达到稳定值时所需的时间要比由低照度状态变为高照度状态时短。



硫化镉光敏电阻亮态前历效应曲线
微变等效电路



光敏电阻的微变等效电路

设入射于光敏电阻的辐射为正弦形式，有

$$\Phi = \Phi_0 + \Phi_m e^{j\omega t}$$

Φ ：辐射通量

ω ：调制辐射的圆频率

则流过光敏电阻的光电流 I 为

$$I = I_0 + e^{j\omega t} \cdot I_m / (1 + j\omega \tau)$$

I_0 ：光电流 I 的直流分量， $I_0 = S_g \Phi_0 U$ 。

U ：加到光敏电阻两端的电压

$I_m / (1 + j\omega \tau)$ ：微变光电流的幅值， $I_m = S_g \Phi_m U$ 。

τ ：时间常数

如果令

$$i = I_m / (1 + j\omega \tau)$$

则可以看出，流过光敏电阻 R 的微变光电流 i ，相当于电流源 I_m 在 RC 并联电路中被电容 C 分流后流过电阻 R 的电流。图中的 C 是按光敏电阻的属性而引入的等效电容， $C = \tau / R$ ， R 为与直流辐射分量相对应的光电阻。

光敏电阻的使用

光敏电阻的重要特点是，光谱响应范围宽，测光范围宽，灵敏度高，无极性之分。但由于材料不同，在性能上差别较大。使用中应注意：

- 1) 当用于模拟量测量时，因光照指数 γ 与光照强弱有关，只有在弱光照下光电流与入射辐射通量成线性关系。
- 2) 用于光度量测试仪器时，必须对光谱特性曲线进行修正，保证其与入眼的光谱光视效率曲线符合。
- 3) 光敏电阻的光谱特性与温度有关，温度低时，灵敏范围和峰值波长都向长波方向移动，可采取冷却灵敏面的办法来提高光敏电阻在长波区的灵敏度。
- 4) 光敏电阻的温度特性很复杂，电阻温度系数有正有负，一般说，光敏电阻不适于在高温下使用，温度高时输出将明显减小，甚至无输出。
- 5) 光敏电阻频带宽度都比较窄，在室温下只有少数品种能超过 1000Hz，而且光电增益与带宽之积为一常量，如要求带宽较宽，必须以牺牲灵敏度为代价。
- 6) 设计负载电阻时，应考虑到光敏电阻的额定功耗，负载电阻值不能很小。
- 7) 进行动态设计时，应意识到光敏电阻的前历效应。

7.4 光伏器件

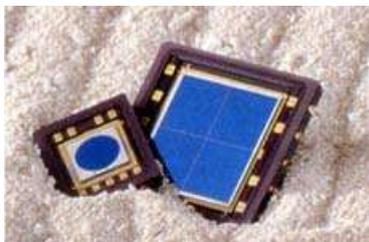
利用半导体 PN 结光伏效应制成的器件称为光伏器件，也称结型光电器件。这类器件品种很多，其中包

括各种光电池、光电二极管、光电晶体管、光电场效应管、PIN 管、雪崩光电二极管、光可控硅、阵列式光电器件、象限式光电器件、位置敏感探测器 (PSD)、光电耦合器件等。

光电池

光电池的基本结构就是一个 PN 结。按材料分, 有硅、硒、硫化镉、砷化镓和无定型材料的光电池等。按结构分, 有同质结和异质结光电池等。

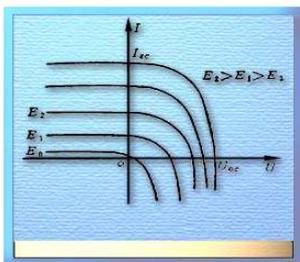
光电池中最典型的是同质结硅光电池。国产同质结硅光电池因衬底材料导电类型不同而分成 2CR 系列和 2DR 系列两种。



硅光电池结构示意图

2CR 系列硅光电池是以 N 型硅为衬底, P 型硅为受光面的光电池。受光面上的电极称为前极或上电极, 为了减少遮光, 前极多作成梳状。衬底方面的电极称为后极或下电极。为了减少反射光, 增加透射光, 一般都在受光面上涂有 SiO_2 或 MgF_2 , Si_3N_4 , $\text{SiO}_2\text{-MgF}_2$ 等材料的防反射膜, 同时也可以起到防潮, 防腐蚀的保护作用。

光电池在光照下能够产生光生电势, 光电流实际流动方向为, 从 P 端流出, 经过外电路, 流入 N 端, 光生电势与照度是对数关系。当光电池短路时, 短路电流 I_{sc} 与照度 E 成线性关系, $S = I_{sc}/E$ 称为灵敏度。



硅光电池的伏安特性曲线

在一定的照度下, 曲线在横轴的截距, 代表该照度下的开路电压 U_{oc} 。曲线在纵轴的截距, 代表该照度下的短路电流 I_{sc} 。硅光电池的 U_{oc} 一般为 $0.45\sim 0.6\text{V}$, 最大不超过 0.756V , 因为它不能大于 PN 结热平衡时的接触电势差。硅单晶光电池短路电流为 $35\sim 40\text{mA}/\text{cm}^2$ 。

几种国产硅光电池的特性

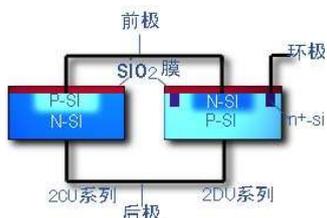
参数名称单位	开路电压 /mV	短路电流 /mA	输出电流 /mA	转换效率 (%)	面积 /mm ²
测试条件	T=30℃ E=1000W/m ²	T=30℃ E=1000W/m ²	T=30℃ E=1000W/m ² 输出电压 400mV下		
型号					
2CR11	450~600	2~4		> 6	2.5 × 5
2CR21	450~600	4~8		> 6	5 × 5
2CR31	450~600	9~15	6.5~8.5	6~8	5 × 10
2CR32	550~600	9~15	8.6~11.3	8~10	5 × 10
2CR41	450~600	18~30	17.6~22.5	6~8	10 × 10
2CR42	500~600	18~30	22.5~27	8~10	10 × 10
2CR51	450~600	36~60	35~45	6~8	10 × 20
2CR52	500~600	36~60	45~54	8~10	10 × 20
2CR61	450~600	40~65	30~40	6~8	$(\frac{17}{2})^2 \pi$
2CR62	500~600	40~65	40~51	8~10	$(\frac{17}{2})^2 \pi$
2CR71	450~600	72~120	54~120	> 6	20 × 20
2CR81	450~600	88~140	66~85	6~8	$(\frac{25}{2})^2 \pi$
2CR82	500~600	88~140	86~110	8~10	$(\frac{25}{2})^2 \pi$
2CR91	450~600	18~30	13.5~30	> 6	5 × 20
2CR101	450~600	173~288	130~288	> 6	$(\frac{35}{2})^2 \pi$

光电二极管

光电二极管和光电池一样，其基本结构也是一个 PN 结。它和光电池相比，重要的不同点是结面积小，因此它的频率特性特别好。光生电势与光电池相同，但输出电流普遍比光电池小，一般为数微安到数十微安。

按材料分，光电二极管有硅、砷化镓、铟化镓、铟化铅光电二极管等许多种。按结构分，也有同质结与异质结之分。其中最典型的还是同质结硅光电二极管。

国产硅光电二极管按衬底材料的导电类型不同，分为 2CU 和 2DU 两种系列。2CU 系列以 N-Si 为衬底，2DU 系列以 P-Si 为衬底。2CU 系列光电二极管只有两个引出线，而 2DU 系列光电二极管有三条引出线，除了前极、后极外，还设了一个环极。



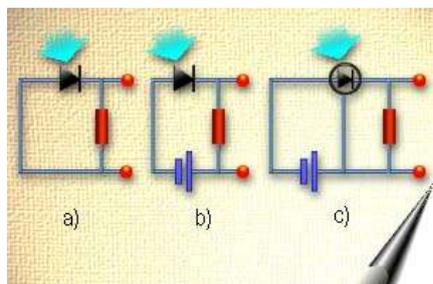
硅光电二极管结构示意图

2DU 管加环极的目的是为了减少暗电流和噪声。光电二极管的受光面一般都涂有 SiO₂ 防反射膜，而 SiO₂ 中又常含有少量的钠、钾、氢等正离子。SiO₂ 是电介质，这些正离子在 SiO₂ 中是不能移动的，但是它们的静电感应却可以使 P-Si 表面产生一个感应电子层。这个电子层与 N-Si 的导电类型相同，可以使 P-Si 表面与 N-Si 连通起来。当管子加反偏压时，从前极流出的暗电子流，除了有 PN 结的反向漏电子流外，还有通过表面感应电子层产生的漏电子流，从而使从前极流出的暗电子流增大。为了减小暗电流，设置一个 N-Si 的

环把受光面 (N-Si) 包围起来, 并从 N⁻-Si 环上引出一条引线 (环极), 使它接到比前极电位更高的电位上, 为表面漏电子流提供一条不经过负载即可达到电源的通路。这样, 即可达到减小流过负载的暗电流、减小噪声的目的。如果使用时环极悬空, 除了暗电流、噪声大些外, 其它性能均不受影响。

2CU 管子, 因为是以 N-Si 为衬底, 虽然受光面的 SiO₂ 防反射膜中也含有少量的正离子, 而它的静电感应不会使 N-Si 表面产生一个和 P-Si 导电类型相同的导电层, 从而也就不可能出现表面漏电流, 所以不需要加环极。

光电二极管的用法:

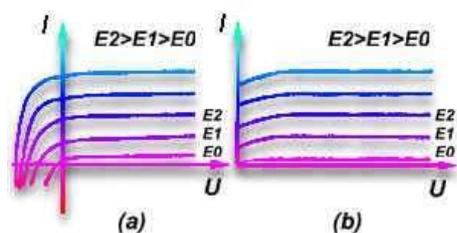


光电二极管的用法

a) 不加外电源 b) 加反向外电源 c) 2DU 环极接法

光电二极管的用法只能有两种。一种是不加外电压, 直接与负载相接。另一种是加反向电压, 如上图所示。实际上, 不是不能加正向电压, 只是正接以后就与普通二极管一样, 只有单向导电性, 而表现不出它的光电效应。

由于多数场合光电二极管都是加反向电压, 所以其伏安特性曲线常画成如下图所示的形式。

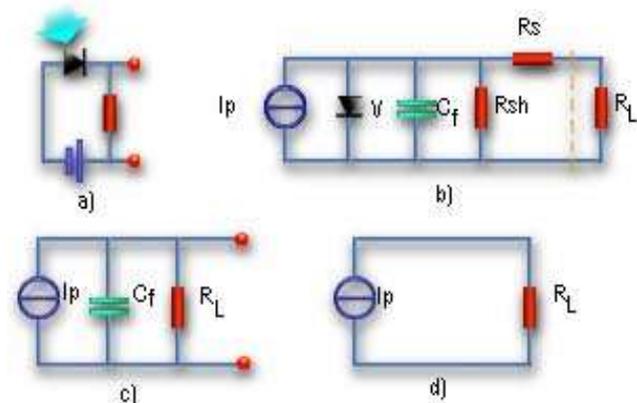


光电二极管的伏安特性曲线

上图的画法与硅光电池的伏安特性曲线图比较, 有两点不同。一是把硅光电池的伏安特性曲线图中 I、II 象限里的图线对于纵轴反转了一下, 变为上图 (a)。这里是以横轴的正向代表负电压, 这样处理对于以后的电路设计很方便。二是因为开路电压 U_{oc} 一般都比外加的反向电压小很多, 二者比较可略而不计, 所以实用曲线常画为上图 (b) 的形式。

微变等效电路与频率特性:

在一定的照度下, 光电二极管的伏安特性曲线几乎是平直的, 可把它看成是恒流源。



光电二极管的等效电路

其中图 a 为实际电路；图 b 为考虑到光电二极管结构、功能后画出的微变等效电路，其中 I_p 为光电流， V 为理想二极管， C_j 为结电容， R_{sh} 为漏电阻， R_s 为体电阻， R_L 为负载电阻；

图 c 是从图 b 简化来的，因为正常运用时，光电二极管要加反向电压， R_{sh} 很大， R_s 很小，所以图 b 中的 D 、 R_{sh} 、 R_s 都可以不计，因而有图 c 的形式；

图 d 又是从图 c 简化来的，因为 C_j 很小，除了高频情况要考虑它的分流作用外，在低频情况下，它的阻抗很大，可不计。因此具体应用时多用图 d 和图 c 两种形式。

流过负载的交变电流复振幅为

$$I_L = I_p \cdot 1 / (1 + j \omega \tau)$$

ω ：入射光的调制圆频率， $\omega = 2 \pi f$ ， f 为入射光的调制频率。

$$\tau = C_j R_L$$

I_L 的模量为

$$I_L = \frac{1}{\sqrt{1 + \omega^2 \tau^2}} |I_p|$$

可见， I_L 是频率的函数，随着入射光调制频率的增加而减小。当 $\omega = 1/\tau$ 时， $I_L = |I_p| / \sqrt{2} \approx 0.707 I_p$ ，这时 $f = 1/2 \pi \tau$

称为上限截止频率，或称为带宽。

几种国产 2CU 型硅光电二极管的特性

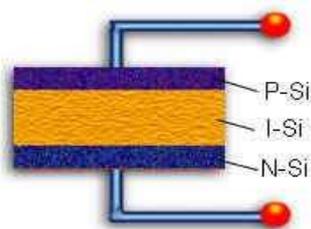
参数名称单位	最高工作电压 U_{max}/V	暗电流 I_D $/\mu A$	光电流 $/\mu A$	电流灵敏度 $/(\mu A \cdot \mu W^{-1})$	结电容 $/pF$	响应时间 $/s$
型号	测试条件 $I_D < 0.2 \mu A$ $E < 1 mW \cdot m^{-2}$	$U = U_{max}$	$U = U_{max}$	$U = U_{max}$ $\lambda = 0.9 \mu m$	$U = U_{max}$	$U = U_{max}$ $R_L = 100 \Omega$
2CU ₁ A	10	≤ 0.2	> 80	> 0.5	≤ 5	10^{-7}
2CU ₁ B	20	≤ 0.2	> 80	> 0.5	≤ 5	10^{-7}
2CU ₁ C	30	≤ 0.2	> 80	> 0.5	≤ 5	10^{-7}
2CU ₁ D	40	≤ 0.2	> 80	> 0.5	≤ 5	10^{-7}
2CU ₁ E	50	≤ 0.2	> 80	> 0.5	≤ 5	10^{-7}
2CU ₂ A	10	≤ 0.1	> 30	> 0.5	< 5	10^{-7}
2CU ₂ B	20	≤ 0.1	> 30	> 0.5	< 5	10^{-7}
2CU ₂ C	30	≤ 0.1	> 30	> 0.5	< 5	10^{-7}
2CU ₂ D	40	≤ 0.1	> 30	> 0.5	< 5	10^{-7}
2CU ₂ E	50	≤ 0.1	> 30	> 0.5	< 5	10^{-7}

几种国产 2DU 型硅光电二极管的特性

参数名称/单位	最高工作电压 U_{max}/V	中心暗电流 $I_{c}/\mu A$	环电流 $I_{r}/\mu A$	光电流 $I_{ph}/\mu A$	电流灵敏度 $I_{ph}/(A \cdot W^{-1})$	响应时间 $t_r/\mu s$	结电容 C_j/pF	正向压降 V_f/V
测试条件		$U = -50V$	$U = -50V$	$U = -50V$ 在 10^4lx 照度下	$U = -50V$ $\lambda = 0.9\mu m$	$U = -50V$ $R_L = 100\Omega$	$U = -50V$ $f = 1kHz$	正向电流 10mA
型号								
2DUAG	50	≤ 0.05	≤ 3	> 6	> 0.4	< 0.1	2~3	≤ 3
2DU ₁ A	50	≤ 0.1	≤ 5	> 6	> 0.4	< 0.1	2~3	≤ 6
2DU ₂ A	50	0.1~0.3	5~10	> 6	> 0.4	< 0.1	2~3	≤ 5
2DU ₃ A	50	0.3~1.0	10~30	> 6	> 0.4	< 0.1	2~3	≤ 5
2DUBG	50	≤ 0.05	≤ 3	> 20	> 0.4	< 0.1	3~8	≤ 3
2DU ₁ B	50	≤ 0.1	≤ 5	> 20	> 0.4	< 0.1	3~8	≤ 5
2DU ₂ B	50	0.1~0.3	5~10	> 20	> 0.4	< 0.1	3~8	≤ 5
2DU ₃ B	50	0.3~1.0	10~30	> 20	> 0.4	< 0.1	3~8	≤ 5

PIN 管

PIN 管是光电二极管中的一种。它的结构特点是，在 P 型半导体和 N 型半导体之间夹着一层（相对）很厚的本征半导体。这样，PN 结的内电场就基本上全集中于 I 层中，从而使 PN 结双电层的间距加宽，结电容变小。由式 $\tau = C_j R_L$ 与 $f = 1/2\pi\tau$ 知， C_j 小， τ 则小，频带将变宽。因此，这种管子最大的特点是频带宽，可达 10GHz。另一个特点是，因为 I 层很厚，在反偏压下运用可承受较高的反向电压，线性输出范围宽。由耗尽层宽度与外加电压的关系可知，增加反向偏压会使耗尽层宽度增加，从而结电容要进一步减小，使频带宽度变宽。所不足的是，I 层电阻很大，管子的输出电流小，一般多为零点几微安至数微安。目前有将 PIN 管与前置运算放大器集成在同一硅片上并封装于一个管壳内的商品出售。

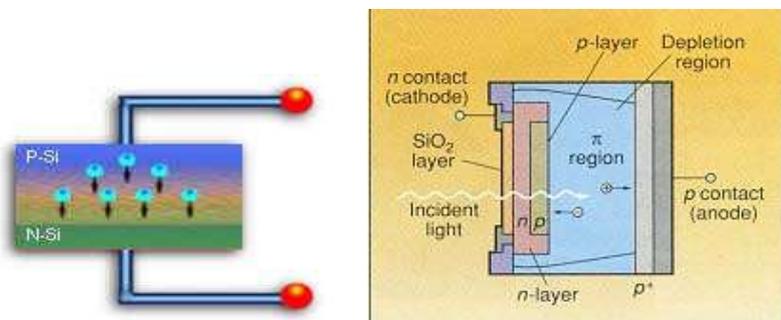


PIN 管结构示意图

几种国产 2CU 型 PIN 光电二极管的特性

雪崩光电二极管

雪崩光电二极管是利用 PN 结在高反向电压下产生的雪崩效应来工作的一种二极管。



雪崩光电二极管工作原理示意图

这种管子工作电压很高，约 100~200V，接近于反向击穿电压。结区内电场极强，光生电子在这种强电场中可得到极大的加速，同时与晶格碰撞而产生电离雪崩反应。因此，这种管子有很高的内增益，可达到几百。当电压等于反向击穿电压时，电流增益可达 10^5 ，即产生所谓的自持雪崩。这种管子响应速度特别快，

带宽可达 100GHz，是目前响应速度最快的一种光电二极管。

噪声大是这种管子目前的一个主要缺点。由于雪崩反应是随机的，所以它的噪声较大，特别是工作电压接近或等于反向击穿电压时，噪声可增大到放大器的噪声水平，以至无法使用。

**200 mm² ultra-sensitive APDs.
90% quantum efficiency.
Preconnected circuits.
Packaged to go.**



Our large-area avalanche photodiodes (LAAPDs) are the most sensitive solid-state detectors per unit area. Now, they are integrated with a high-voltage power supply, low-noise preamplifier, and the LAAPD of your choice — ideal for your initial evaluation. These modules feature AC or DC coupling depending on the application, while sensitivity at the input is better than 100 picowatts in bandwidths of up to 25 MHz. Featuring standard diameters of 5, 10, and 16mm, these devices are now available with improved quantum efficiency enhanced for UV, blue, and red spectra. A cooled version also cuts noise in half, and stabilizes the signal with changes in ambient temperature.

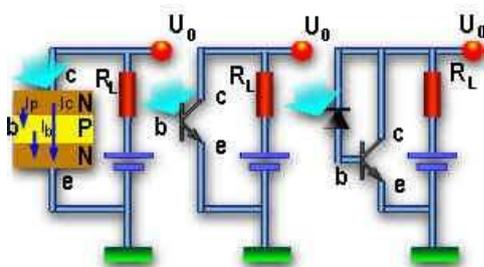
The compact LAAPD can be sold separately or in a convenient module package.

Advanced Photonix's LAAPD is the superior choice for many applications that formerly used PMTs or small APDs. With so many advantages, it's no wonder LAAPDs are now replacing conventional detectors for fluorescence measurements, chemical analysis, remote scanning, ranging, and scintillator applications. So, give us a call at (805) 987-0146 and order our large-area avalanche photodiodes — separately or packaged to go — and put them to work for you.

Advanced Photonix also offers standard and customized silicon PIN photodiodes. Call today for our optoelectronics catalogue.

光电晶体管

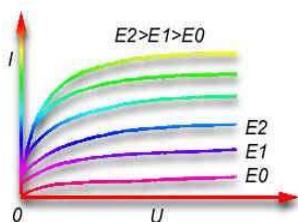
光电晶体管和普通晶体管类似，也有电流放大作用。只是它的集电极电流不只是受基极电路的电流控制，也可以受光的控制。所以光电晶体管的外形，有光窗、集电极引出线、发射极引出线和基极引出线（有的没有）。制作材料一般为半导体硅，管型为 NPN 型，国产器件称为 3DU 系列。



光电晶体管原理性结构图

正常运用时，集电极加正电压。因此，集电结为反偏置，发射结为正偏置，集电结为光电结。当光照到集电结上时，集电结即产生光电流 I_p 向基区注入，同时在集电极电路即产生了一个被放大的电流 I_c ($= I_e = (1 + \beta) I_p$)， β 为电流放大倍数。因此，光电晶体管的电流放大作用与普通晶体管在上偏流电路中接一个光电二极管的作用是完全相同的。

光电晶体管的灵敏度比光电二极管高，输出电流也比光电二极管大，多为毫安级。但它的光电特性不如光电二极管好，在较强的光照下，光电流与照度不成线性关系。所以光电晶体管多用来作光电开关元件或光电逻辑元件。



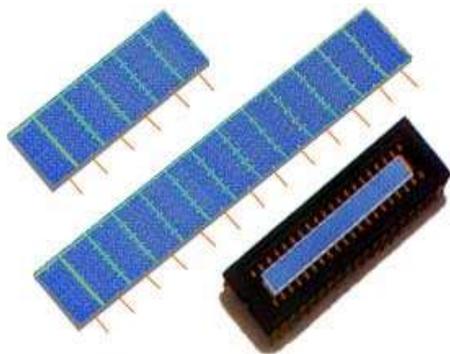
光电晶体管的伏安特性曲线

几种国产 3DU 型光电三极管的特性

参数名称单位	最高工作电压 U_{max}/V	暗电流 μA	光电流 $/mA$	结电容 $/pF$	响应时间 $/\mu s$	集电极最大电流 $/mA$	最大功耗 $/mW$
型号	测试条件	$U_{cc}=U_{max}$	$U_{cc}=U_{max}$ $E=10^{11}lx$	$U_{cc}=U_{max}$ $f=1kHz$	$U_{cc}=10V$ $R_L=100\Omega$		
3DU11	10	≤ 0.3	> 0.5	≤ 10	10	20	150
3DU12	30	≤ 0.3	> 0.5	≤ 10	10	20	150
3DU13	50	≤ 0.3	> 0.5	≤ 10	10	20	150
3DU21	10	≤ 0.3	> 1.0	≤ 10	10	20	150
3DU22	30	≤ 0.3	> 1.0	≤ 10	10	20	150
3DU23	50	≤ 0.3	> 1.0	≤ 10	10	20	150
3DU31	10	≤ 0.3	> 2.0	≤ 10	10	20	150
3DU32	30	≤ 0.3	> 2.0	≤ 10	10	20	150
3DU33	50	≤ 0.3	> 2.0	≤ 10	10	20	150
3DU51A	15	≤ 0.2	> 0.3	< 5	10	10	50
3DU51B	30	≤ 0.2	> 0.3	< 5	10	10	50
3DU51C	30	≤ 0.2	> 1.0	< 5	10	10	50

阵列式或象限式结型光电器件

利用集成电路技术使 2 个至几百个光电二极管或光电池排成一行，集成在一块集成电路片子上，即成为阵列式的一维光电器件，也可以使光电二极管或光电池制成象限式的二维光电器件。



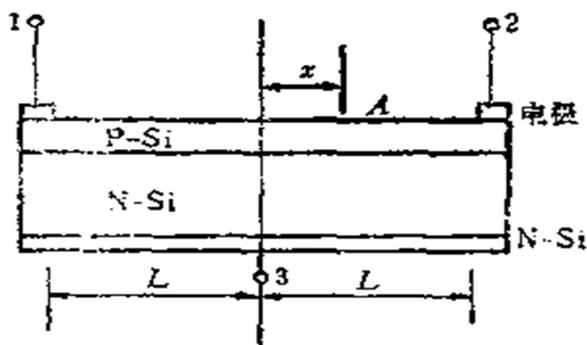
这两种器件中，衬底是共用的，而各光敏元都是独立的，分别有各自的前极引出线。这种器件的特点是，光敏元密集度大，总尺寸小，容易作到各单元多数一致，便于信号处理。就目前的应用看，两个并列的光电二极管或光电池，可用来辨别光点移动的方向。2~4 个并列的光敏元，可用来收集光点移动的相位信息。几十个至几百个或更多并列的光敏元，可用来摄取光学图象或用作空间频谱分析。象限式光电器件可用来确定光点在二维平面上的位置坐标。多用于准直、定位、跟踪或频谱分析等方面。

光电位置探测器 (PSD, Position Sensitive Detectors)

PSD 是利用离子注入技术制成的一种可确定光的能量中心位置的结型光电器件，有一维的和二维的两种。当入射光是一个小光斑，照射到光敏面时，其输出则与光的能量中心位置有关。这种器件和象限光电器件比较，其特点是，它对光斑的形状无严格要求，光敏面上无象限分隔线，对光斑位置可连续测量。



1) 一维 PSD



如图所示，PSD 的受光面为 P-Si，同时也是个均匀电阻层。设 1、2 两电极间距离为 2L。如果入射光点位于 A 点，则电极 1、2 输出的光电流与 A 点至电极 1、2 的距离成反比，有

$$I_1 = I_0 \cdot (L-x) / 2L$$

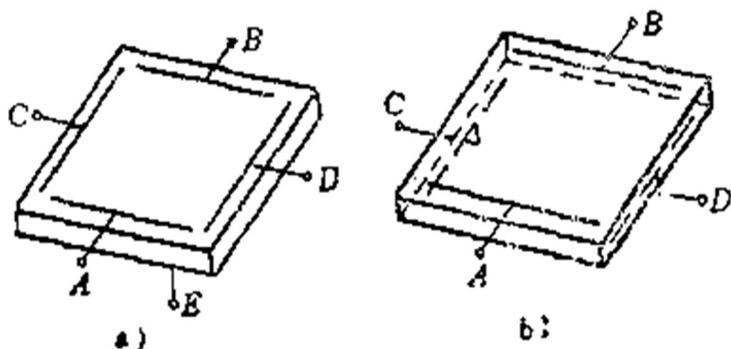
$$I_2 = I_0 \cdot (L+x) / 2L$$

$$x = L \cdot (I_2 - I_1) / (I_2 + I_1)$$

式中， $I_0 = I_1 + I_2$ 。

利用上式即可确定光斑能量中心对于器件中心的位置。

2) 二维 PSD



二维 PSD 有两种形式，一种是单面型的如图 a 所示，在受光面上设有两对电极，A、B 为 x 轴电极，C、D 为 y 轴电极，E 为背面衬底共用电极，用它可对正面各电极进行反偏置。设 $I_A \sim I_D$ 为电极 A~D 的光电流，则光点能量中心的位置坐标为

$$x = (I_A - I_B) / (I_A + I_B)$$

$$y = (I_C - I_D) / (I_C + I_D)$$

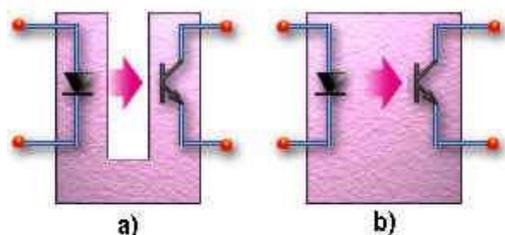
另一种是双面型的，如图 b 所示，正面与背面之间是一个 PN 结，正面和背面都是均匀电阻层。x 轴电

极 A、B 安在正面的受光面上，y 轴电极 C、D 垂直于 x 轴安在背面。光点产生的光电流分为正面与背面两部分。对于这种结构，反偏压是加在正面电极与背面电极之间（信号电极与偏置电极不独立）。设 $I_A \sim I_D$ 分别为电极 A~D 的光电流，代入上式即可求得光点能量中心位置。

需要指出，上面的式子都是近似式，在器件中心附近是正确的，而距离器件中心较远接近边缘部分时误差较大。对于二维 PSD，双面型的比单面型的位置误差小。这可能是因为 x、y 轴分开设置减少了彼此干扰的缘故。

光电开关与光电耦合器

光电开关和光电耦合器都是由发光端和受光端组成的组合件。光电开关不封闭，发光端与受光端之间可以插入调制板。光电耦合器则是把发光元件与受光元件都封闭在一个不透光的管壳内。



光电开关与光电耦合器结构示意图

a) 光电开关 b) 光电耦合器

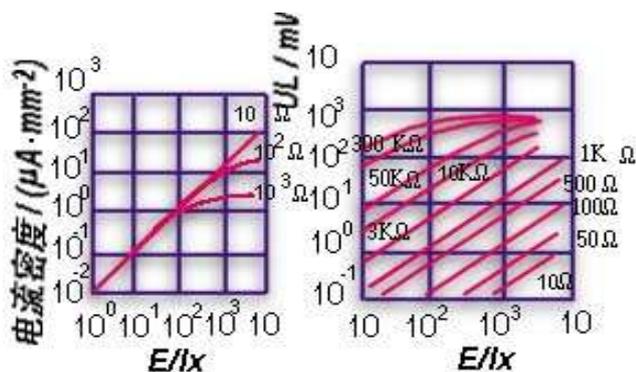
发光端与受光端彼此独立，完全没有电的联系，两端之间的电阻一般都在 $10^{11} \Omega$ 以上。光电开关多用于光电计数、报警、安全保护、无接触开关，及各种光电控制等方面。光电耦合器多用于电位隔离、电平匹配、抗干扰电路、逻辑电路、模/数转换、长线传输、过流保护，及高压控制等方面。

各种结型光电器件的类似特性

这类器件在原理上都相同，所以在性质上都类似。

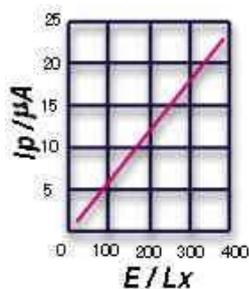
1. 光电特性

器件的光电特性一般是指光电流与照度之间的函数关系，不过有时也表示为其它输出量与照度的函数关系。器件的光电特性主要决定于材料，同时也与结构和使用条件（负载大小、所加电压高低）有关。

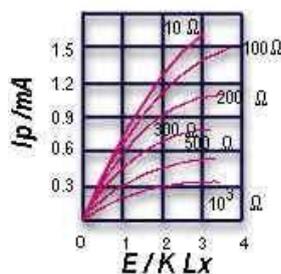


硅光电池的
光电特性曲线

2DUA 硅光电二
极管在零偏压下的光电特性曲线



硅光电二极管在反偏为 15V 时的光电特性曲线

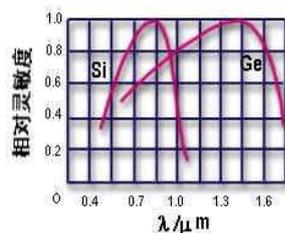


硒光电池光电特性曲线

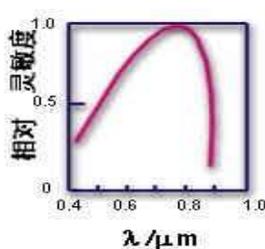
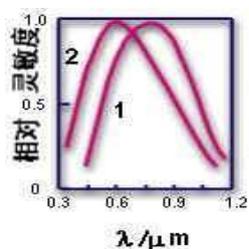
由图可见，器件的光电特性不仅与材料有关，同时也与光照范围、负载大小、外加电压等条件有关。一般说，负载电阻一定时，照度低时光电流与照度有良好的线性关系，照度高时线性关系则变差；在相同照度范围内，负载电阻小比负载电阻大时线性关系好；在其它条件都相同的情况下，加反偏压比不加反偏压时，线性范围宽。

2. 光谱特性

器件的光谱特性多用相对灵敏度与波长的关系曲线表示。



硅、锗光电二极管的光谱特性曲线



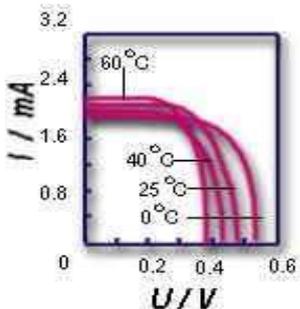
硅光电池的光谱特性曲线 砷化镓光电池的光谱特性曲线

1-普通硅光电池 2-浅结硅光电池

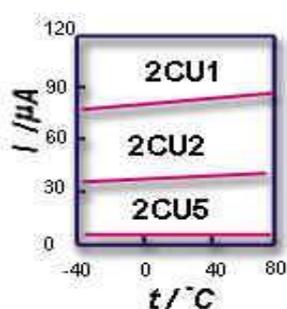
由图可见，器件的光谱特性主要决定于材料，同时与结深和温度（见温度特性）略有关系。

3. 温度特性

温度对器件的开路电压 U_{oc} 、短路电流 I_{sc} 、暗电流 I_d 、光电流 I_p 及单色光灵敏度都有影响。



在不同温度下硅光电池的



2CU 光电二极管光电流伏安特性曲线（在 1000lx 时）

随温

度的变化曲线

4. 噪声、信噪比与噪声等效功率

1) 噪声

结型光电器件的噪声主要是电流散粒噪声和电阻的热噪声。电流散粒噪声是结区释放非平衡载流子的随机性引起的，其噪声电流均方值的计算式为

$$I_{In}^2 = 2qI \Delta f$$

q: 电子电量

I: 通过 PN 结电流的平均值

Δf : 测试系统带宽

电阻的热噪声是由于电阻中电子的热运动引起的，其噪声电流均方值的计算式为

$$I_{Rn}^2 = (4kT/R) \cdot \Delta f$$

k: 波耳兹曼常数

T: 温度

R: 电阻值

Δf : 测试系统的带宽

需要指出的是器件在实际应用时，后面总要接放大器，所以式中的 R 应为 PN 结的漏电阻和负载电阻的并联值。因为 PN 结的漏电阻远大于负载电阻，所以 R 值实际是 PN 结的负载电阻值。

2) 信噪比与噪声等效功率 (NEP)

一般情况下，电流的散粒噪声都比电阻的热噪声大，如果只考虑电流的散粒噪声，则器件的信号与噪声有效值之比为

$$\frac{I_p}{\sqrt{I_n^2}} = \frac{S\Phi}{\sqrt{2qI\Delta f}}$$

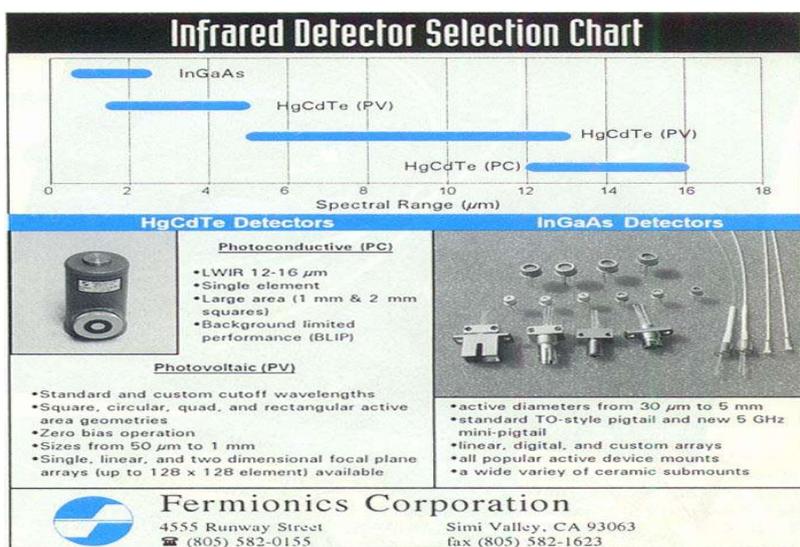
I_p : 光电流

S: 电流灵敏度

Φ : 入射于器件的辐射通量

噪声等效功率 (NEP) 为

$$NEP = \frac{\sqrt{2qI\Delta f}}{S}$$



使用要点

1) 极性结型器件都有确定的极性，如要加电压使用时，光电结必须加反向电压，即 P 端与外电源的低

电位相接。

2) 使用时对入射光强范围的选择应视用途而定。用于开关电路或逻辑电路时光照可以强些。用于模拟量测量时,光照不宜过强。因为一般器件都有这样的性质:光照弱些,负载电阻小些,加反偏压使用时,光电线性好,反之则差。

3) 灵敏度主要决定于器件,但也与使用条件和方法有关,例如光源和接收器在光谱特性上是否匹配;入射光的方向与器件光敏面法线是否一致等。

4) 结型器件的响应速度都很快。它主要决定于负载电阻和结电容所构成的时间常数($\tau = RC$)。负载电阻大,输出电压可以大,但 τ 会变大,响应变慢。相反,负载电阻小些,输出电压要减小,但 τ 会变小,响应速度变快。

5) 灵敏度与频带宽度之积为一常数的结论,对结型光电器件也适用。

6) 器件的各种参量差不多都与温度有关,但其中受温度影响最大的是暗电流。暗电流大的器件,容易受温度变化的影响,而使电路工作不稳定,同时噪声也大。

7) 除了温度变化,电、磁场干扰可引起电路发生误动作外,背景光或光反馈也是引起电路误动作的重要因素,应设法消除。

第八章 光电成像器件

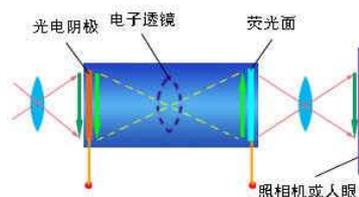
光电成像器件是指能够输出图象信息的一类器件,例如使不可见光图象变为可见光图象的器件,使光学图象变为电视信号的器件等。这类器件的应用领域很广,例如,夜视技术、电视技术、工件的图象测量、精密零件的微小尺寸测量、产品外观检测、应力应变场分析、机器人视觉、交通管理与指挥、定位、跟踪等。

光电摄像器件应具有三种基本功能:光电变换、光电信号存储和扫描输出。光学图象投射到器件光敏面后,称为像素的独立的光敏单元分别完成光电转换,在光敏面上形成电量的潜象。扫描装置形成的扫描线按一定的轨迹串行、逐点地采集这些转换后的电量形成输出信号。扫描线经过某一个像素的时间只占扫描整个光敏面所需周期的极小部分,为了提高检测灵敏度,每个像素在扫描周期内应不间断地对转换后的电量进行积累,这种功能称为光电信号的积分存储。

8.1 象管

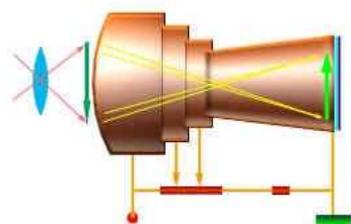
象管是变象管和图象增强管的统称。变象管是指能够把不可见光图象变为可见光图象的真空光电管。图象增强管是指能够把亮度很低的光学图象变为有足够亮度图象的真空光电管。象管和摄像管的主要区别是,象管内部没有扫描机构,不能输出电视信号,对它的使用就跟使用望远镜去观察远处景物一样,观察者必须通过它来直接面对着景物。

象管的结构和原理



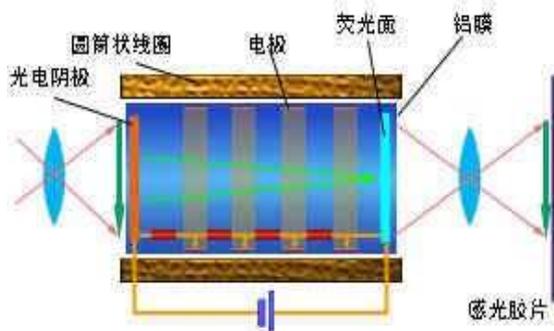
象管的结构原理示意图

象管有三个基本部分。一是光电变换部分，即光电阴极，它可以使不可见光图象或亮度很低的光学图象，变成光电子发射图象。二是电子光学部分，即电子透镜，有电聚焦和磁聚焦两种形式，它可以使光电阴极发射出来的光电子图象，在保持相对分布不变的情况下进行加速。三是电光变换部分，即荧光屏，它可以使打到它上面的电子图象变成可见光。



静电聚焦型象管结构示意图

上图为静电聚焦型象管的基本结构，其中几个圆筒形的电极可形成对光电子聚焦和加速的电场，使电子打在荧光屏上呈倒立的象。电聚焦型象管各电极电压之比保持不变时，即使总电压稍有变化，电子轨迹也基本不变，因此，各电极电压多用电阻链分压的办法供给。这样，同时也减小了整个装置的重暴和体积，但电聚焦的球面象差较大，画面的中心部分和边缘部分放大率不等，图象有失真。因此，上图中的光电阴极多作成曲面状，以补偿电聚焦引起的象差。但曲率大时，焦距又要变小，使边缘部分的分辨率降低。因此，近年来多采用光纤面板，使其外侧为平面，内侧为球面，以解决光学透镜和电透镜的象差问题。



电磁聚焦型象管结构示意图

上图为电磁聚焦型象管的基本结构。其中圆筒形电极用来形成电透镜和加速，管外的线圈用来使管内产生平行于管轴的磁场，以形成磁透镜。如果光电子有偏离于管轴的速度分量，磁场会使它呈螺旋状前进。电子每旋一圈所需的时间与初速度无关，所以光电面上某一点发射出来的电子，不管起初是沿什么方向发射，最终都可以被会聚于一点。

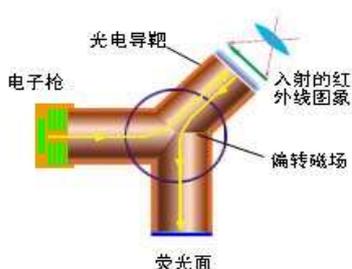
设电子所旋的圈数为 n ，电力线和磁力线方向均平行于管轴，则 $n = c(HL/U)$ ，其中 H 为磁场强度， L 为电极间距离， U 为电极间电压， c 为比例常数。一般象管 $n = 1 \sim 3$ 。

以被加速和聚焦了的电子去轰击荧光面时，荧光面可产生明亮的光学图象，一般 1 个入射的光电子，可使荧光屏产生多个光子。例如，用 1 个 10keV 的电子去轰击 P-20 荧光体时，可产生 300 个光子；1 个 20keV 的电子，可产生 1 千多个光子。光电面的响应时间极短，可忽略不计，而荧光面的响应时间约为数毫秒至数秒，所以整管的响应时间主要决定于荧光面。另外，为了避免光反馈，增加荧光面的光输出，常在荧光面上对着光敏面一边涂有铝膜。



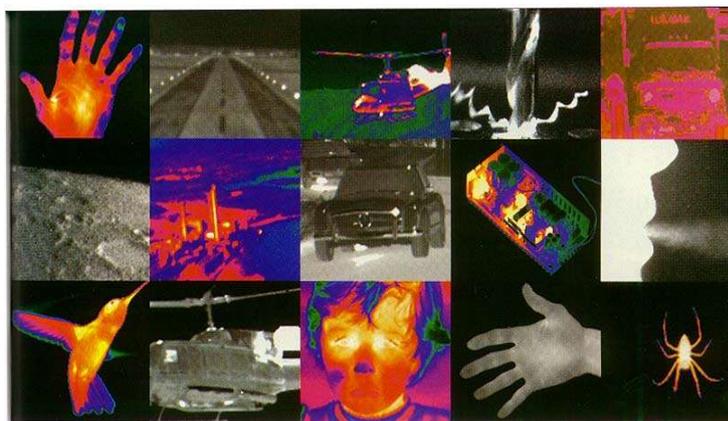
变象管

红外变象管的光电阴极多为 S-1、Ag-O-Cs 阴极，它可以使波长小于 $1.15 \mu\text{m}$ 的红外光变成光电子。对于波长大于 $1.15 \mu\text{m}$ 的红外光，采用负电子亲和势阴极，有的变象管也采用光电导技术，使红外光成像到光电导靶面上，在靶的另一边形成电势分布图象，使入射的电子流受到调制，利用返回的电子流使荧光面发光。



装有光电导靶的反射式变象管

红外变象管多应用于军事、公安等方面，供夜间侦察用。在民用方面，可用于暗室管理、物理实验、激光器校准和夜间观察生物活动等。另外，温度高于 400°C 的物体都会发出大量的红外线，可通过红外变象管观察到它的象。如果与标准光源的亮度比较，即可求出它的强度。这就是夜视温度计的原理。



紫外变象管的窗口材料为石英玻璃，光电发射材料为 S-11， Sb-Cs 阴极。它可以使波长大于 200nm 的紫外光变成光电子。紫外变象管与光学显微镜结合起来，可用于医学和生物学等方面的研究。

图象增强管

图象亮度增益可按下式计算：

光通量增益

$$K_\Phi = \Phi_o / \Phi_i$$

光亮度增益

$$K_L = M_o / E_i = K_\Phi (A_i / A_o)$$

光电子增益

$$K_p = n_{op} / n_{ip}$$

Φ_o 、 Φ_i ：输出、输入光通量

M_o 、 E_i : 输出面的出射度、输入面的照度

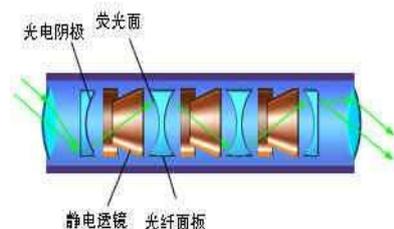
A_i 、 A_o : 输入、输出面的面积

n_{op} 、 n_{ip} : 输出、输入的光子数

这里所有的参量都是对可见光而言的，光源为色温 2856K 的白炽钨丝灯。

对象管的亮度增益，可作如下的估计。对于 S-20 的光电阴极，量子效率约为 10%，入射于荧光面的电子能量，由于技术原因最大只能加大到 20keV，平均 1 个高能电子可产生 500~1000 个光子，因此整管的亮度增益为 50~100。如果象管后面是一个照相光学系统，并考虑到透镜对光的吸收，这样小的亮度增益是不能使感光胶片产生清晰的图象的。因此管内必须有使亮度进一步增益的措施。

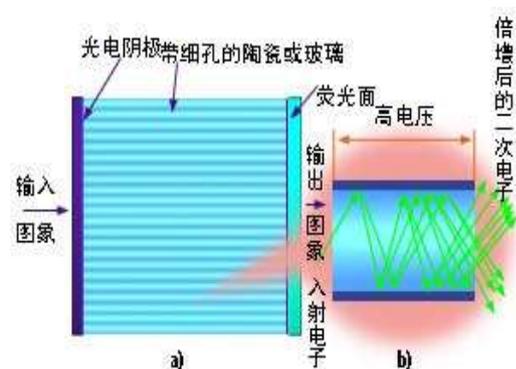
1. 级联式图象增强管



级联式图象增强管

为了增强图象的亮度，现在采取的方法之一是，使几个独立的象管串联起来，使亮度逐级增益。这种 3 管式的图象增强管，亮度增益可达 10^5 。

2. 微通道板式图象增强管



微通道板式图象增强管

该管的核心部分是微通道板，它是由若干个极细的空心管道组成，管径约十几微米。微管道是由高阻材料制成的，微管道的内壁为二次电子发射系数 $\delta > 1$ 的材料。微通道板的厚度约数毫米，在它的两端加上高的直流电压（约数千伏）后，在每个微管道内即形成极强的电场。这时，当光电面发射的电子进入微管道后，在强电场作用下经过和管壁的多重碰撞，而得到电子倍增。一般直流电压为 10kV 的微通道板，可得到 $10^5 \sim 10^6$ 的电子增益。这种增强管和级联式增强管比较，它在输入面和输出面之间没有电子光学系统，所以整管可作得很短，体积很小，很便于与其它光电器件配用。

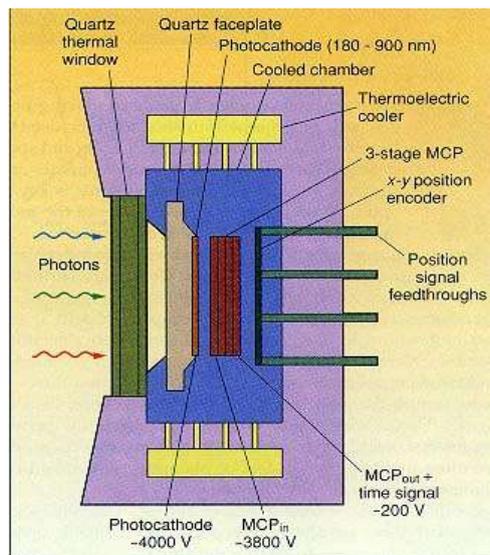
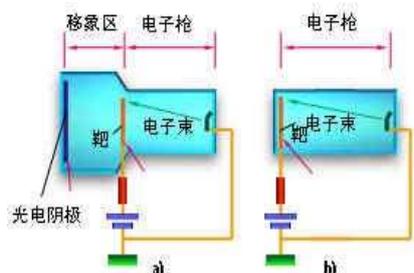


FIGURE 1. Mepsicon imager replicates an image photon by photon on the microchannel plate (MCP) surface.

8.2 摄像管

能够输出视频信号的一类真空光电管称为摄像管。按结构分，常把有移象区的摄像管称为光电发射式摄像管，它的光电变换部分和光信息存储部分是由两部分来完成的，彼此分离，总称为移象区。把没有移象区的摄像管称为光电导式摄像管或视象管，它的两部分功能全由一个靶来完成。电子枪部分二者基本相同。这两类管子比较，光电发射式摄像管历史最早，信号质量也最高，但体积大，结构复杂，调整麻烦，所以目前除特殊场合（微光摄像领域）外一般用得较少。光电导式摄像管比前者体积小，结构简单，信号质量有的也接近于前者，所以电视领域中用得较为普遍。



摄像管示意图

a) 光电发射式摄像管 b) 光电导式摄像管

摄像管的主要参量

1. 光电特性

摄像管输出的光电流与入射的光照度之间的函数关系。常表示为

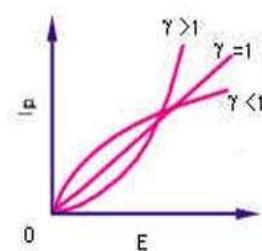
$$I_p = kE^\gamma$$

I_p : 光电流

E : 照度

γ : 光照指数

k : 比例系数



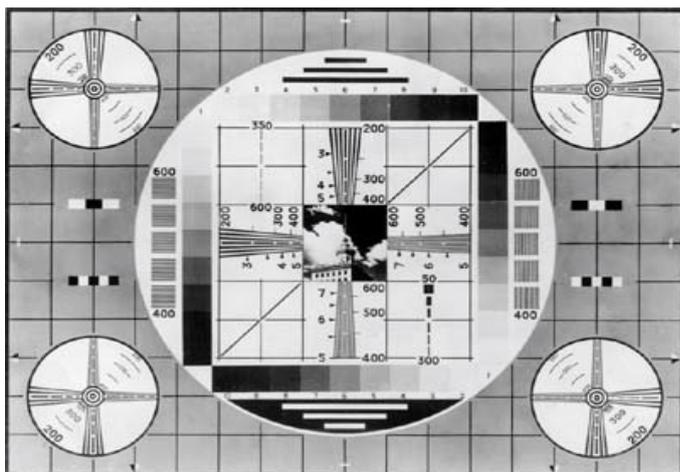
摄像管光电转换特性曲线

由图可见，在光电转换特性中最重要的参量是 γ 。 $\gamma = 1$ 表示 I_p 与 E 成正比例关系，这时，电视信号的灰度等级均匀。 $\gamma < 1$ ，电视信号有均匀的灰度畸变，但此时在低照度下的灵敏度有相对的增加，高照度下的光电特性呈一定的饱和状态。 $\gamma = 1$ 有利于提高暗场时的信噪比， $\gamma < 1$ 有利于扩展动态范围， $\gamma > 1$ 是不适用的。

2. 分辨率

摄像管的分辨率是指摄像管对于光学图象细节的鉴别能力。一般有两种表示法：

1) 极限分辨率



电视分辨率图

在最佳照度下，使高对比度的黑白相间条形图案投射到摄像管的光敏面上，然后在监视器上去观察可分辨的最高空间频率数。在电视中，通常是指在光栅高度范围内可分辨的最多电视行数 (TVL/H)。有时也采用“线对/mm”的单位，它等于可分辨的电视行数一半除以靶的有效高度 (mm)。例如 25mm 的视象管，靶面的有效高度约为 10mm，若可分辨的最多电视行数为 400 时，相当于 20 线对/mm。按这种方法表示的分辨率称为极限分辨率。

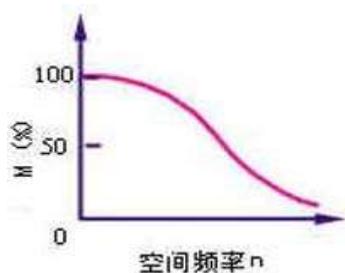
2) 调制传递函数 (MTF)

极限分辨率是依靠观察者的眼睛来分辨的，因而带有一定的主观性，同时也不能反映摄像系统各部分对分辨力的影响。因此，多采用调制传递函数 (MTF) 的概念。MTF 的定义是，输出调制度 M_o 与输入调制度 M_i 之比的百分数，即

$$MTF = M_o / M_i \times 100\% = (A - B) / (A + B) \times 100\%$$

调制度是无线电学中的概念，引用到光学中来可以说它是对比度。M 的定义是，光信息的最大值 A 与光信息最小值 B 之差，对 A、B 和的百分比。

图象在传送过程中，调制度 M 是随空间频率的增大而减小的。如果把调制度的损失程度以百分数表示 (以零频时的调制度为 100%)，则调制度与空间频率的关系曲线，就是调制传递函数。MTF 能够用仪器测量，因此它能客观地反映摄像管的分辨力。

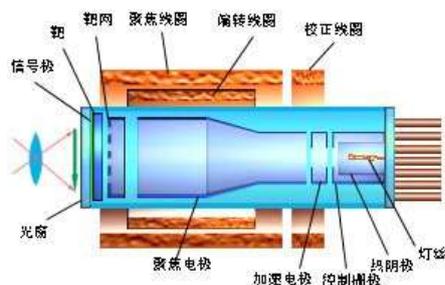


调制传递函数曲线

3. 惰性

摄像管的光电流输出滞后于光信息输入的一种现象。常用三场后残余信号的百分比表示。这里所说的场，是指电视场，按我国的电视制式，场周期为 20ms，三场后残余信号的百分比，即光信息变化 60ms 后的输出电信号，对于 60ms 前原输出电信号的百分比。

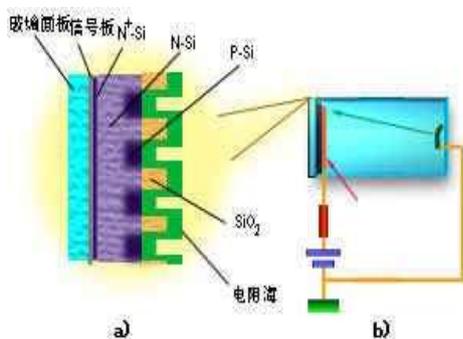
光电导摄像管 (视象管)



视象管结构示意图

视象管的基本结构包括两大部分：光电靶和电子枪。光电靶由光窗、信号极和靶组成，靶面的光敏层可进行光电转换，它的轴向电阻小，横向电阻大，有利于保持光电转换形成的电量潜象，并在扫描周期内实现积分存储。电子枪部分包括灯丝、热阴极、控制栅极、各加速电极和聚焦电极、靶网电极和管外的聚焦线圈、偏转线圈、校正线圈等，它的作用是产生热电子，并使它聚焦成很细的电子射线，按着一定的轨迹扫描靶面。

1. 硅靶摄像管



硅靶结构图与硅靶摄像管简图结构：

硅靶摄像管窗口玻璃内表面涂有一层很薄的既可透光也可导电的金属膜，在它上面接有引线可同负载相连，称为信号板。挨着信号板的是一块N型硅片（硅靶）。硅片朝着电子枪一边的表面，先生成一层氧化层（ SiO_2 ），接着利用光刻技术在 SiO_2 上光刻成几十万个小孔，再通过掺杂使每个小孔都变成 P-Si。这样，许多个小的 P-Si 被 SiO_2 隔离，故称为 P 型岛。每个 P 型岛与 N 型衬底之间即形成一个 PN 结（光电二极管）。最后再在 SiO_2 和 P 型岛表面上蒸涂上一层电阻率适当的材料，即成为硅靶。

硅靶的光电变换：

硅靶的光电变换过程是，信号板通过引线、负载电阻与靶电源的正极相连。电子枪的热阴极接地，扫描电子束即具有地的电位。当电子束扫描到每个 P 型岛时，P 型岛的 PN 结即被反偏置，结电容被充电到靶电源电压。无光照时，由于 PN 结有反向漏电流（暗电流），在负载电阻上要产生少量的电压降，靶的两边——成象面（信号板一边的靶面）和扫描面（朝着电子枪一边的靶面）之间的电压，略低于靶电源电压。有光照时，光进入到每个 PN 结区将产生电子空穴对，它们被结的内电场分离以后，光生的电子通过信号板等外电路入地，光生的空穴则被积累于 P 型岛上。如果光照是均匀的，靶的扫描面电位只是均匀地升高。如果光照不均匀，是一幅光学图象，则扫描面上各 P 型岛的电势分布，将正比于入射光学图象的亮度分布，亮度高的点，所对应的 P 型岛的电势也高。

硅靶摄像管产生视频信号的过程：

扫描电子束按一定的制式去扫描靶面。例如先从左上角开始，从左向右扫，然后再一行挨着一行地从上向下扫，当扫到最右下角时，再返回到左上角，接着扫下一帧。这相当于用一条软导线，按着一定的次序去接通每个 P 型岛。当电子束与每个 P 型岛接触时，上靶电子数的多少，正比于各 P 型岛电势的高低。因此，在输出回路（靶-负载电阻-电源-热阴极-靶）中即产生与之对应的电子流，在负载电阻上即得到与之对应的

视频电压信号。与此同时，被扫到的 P 型岛，即被拉回到地的电位，为下一帧光积分作准备。当下一帧再扫到它时，它就可以把一个帧周期时间内积累起的信息电荷，瞬时地放出来，就这样反复下去，即可得到与光学图象亮度分布相对应的时序视频信号。

硅靶的特点：

1) 抗烧蚀性

耐强光、耐高温是硅靶的突出特点，曾作过这样的实验，用 $f:1.5$ 镜头使硅靶直接承受正午的阳光直射而未见损坏；硅靶还能够经得起 400°C 高温的烘烤。硅靶管的寿命主要决定于电子枪，一般都可以工作 1500 小时以上。

2) 光谱特性及灵敏度

光谱特性主要决定于材料，硅靶的光谱响应范围为 $0.4\sim 1.1\ \mu\text{m}$ ，缺陷波长在 $0.65\sim 0.85\ \mu\text{m}$ 。对于色温 2856K 的标准光源，积分灵敏度为 $4350\ \mu\text{A}/\text{l m}$ ，若滤去红外光，则为 $1100\ \mu\text{A}/\text{l m}$ 。

3) 光电特性与分辨率

硅靶的光照指数 $\gamma \approx 1$ 。分辨率因受二极管尺寸限制，一般偏低些，平均为 $600\text{TVL}/\text{H}$ ，中心可达 $700\text{TVL}/\text{H}$ 。

4) 惰性

硅靶的惰性较小，主要是放电惰性，三场后的残余信号约为 8%。

5) 暗电流

硅靶的暗电流较大，约为 10nA ，这是它的一个缺点。暗电流与温度有关，温度每升高 9°C ，暗电流约增加一倍。

6) 斑点

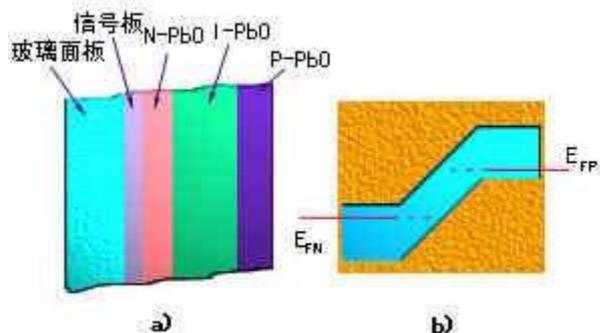
用硅靶管摄制的图象，在显示器上常有若干个黑点或白点。这是因为靶面上几十万个光电二极管中，总难免有个别失效（如光刻时的漏刻或连通等）而造成的。

7) 开花

开花现象有反型开花和过荷开花两种。反型开花是高亮度区二极管不能正常工作，使亮区范围扩展的一种现象。其原因是高亮度区部分， SiO_2 下面的 N 型硅表面产生了反型层，使邻近的二极管连通造成的。过荷开花是在黑暗背景中局部有强光，使显示器上光点逐渐扩展的一种现象。其原因是硅靶上的每个 P 型岛所能够收集的空穴数是有限度的，如果某点的入射光特别强，以至所产生的空穴数超过了该处 P 型岛所能收集的限度，它就要向周围扩展，使邻区信号增强，亮点尺寸逐渐扩大而造成“开花”。

2. 氧化铅靶摄像管

PbO 靶摄像管的结构和工作过程都与硅靶管类似，所不同的是，它的靶是由 PbO 材料制成的。



PbO 靶结构及能带图

窗口玻璃内壁是一层金属膜作为信号板，接着就是 PbO 靶，靶的成像面一边为 N- PbO ，扫描面一边为 P- PbO ，两者之间夹着一层（相对）很厚的本征氧化铅 I- PbO ，因而具有 PIN 结构。工作时，信号板通过负载和靶电源的正极相接，电子枪的热阴极接地，当扫描电子束扫描靶面时，相当于对 PIN 进行反偏置。靶电源电压约 45V 左右。 PbO 靶也有光信息的存储功能，它的轴向电阻较小，横向电阻很大，扫描面上的电势起伏可保持较长时间不变。

PbO 管的特点:

PbO 管是目前广播电视和彩色电视中用得较多的一种摄像管。

1) 灵敏度

灵敏度较高, 可达 $400 \mu A/1m$ 。因本征层电阻率高, 能承受强电场, 所加的反向电压几乎全降落在本征层上。光生载流子在强电场中漂移速度很大, 渡越时间极短, 因此复合率小, 故光生载流子几乎全都能构成信号电流。

2) 暗电流

暗电流小, 低于 $1nA$ 。PI 结与 IN 结的势垒在反偏压下阻挡了多数载流子的注入, 暗电流仅限于 I 层的热生载流子, 并且 I 层的电阻又很大, 故暗电流极小, 而且靶压增高时, 暗电流有饱和现象。

3) 光电特性

光电特性好, $\gamma \approx 1$ 。因光生载流子几乎全能达到两极, 伏安特性曲线在工作区呈饱和状态。增加靶压不能使信号电流增大, 所以光电流正比于靶面照度, 光照系数 $\gamma \approx 1$ 。这点对彩色电视特别有利, 可使摄像管中的 γ 校正电路大为简化。另外, 因信号电流不随靶压变化, 所以灵敏度也就与靶压无关, 这也是个很重要的性质。

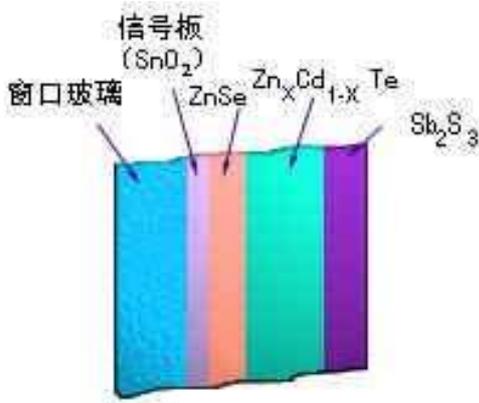
4) 惰性

惰性小, 三场后残余信号不大于 4%。摄像管中的惰性有光电惰性和放电惰性两种, PbO 靶的光电惰性极小, 可不计, 主要是放电惰性。因为靶层中电场极强, 光生载流子渡越时间极短, 光生载流子被 PbO 中陷阱能级俘获的机会少, 从而也就减少了光电流对光照变化的滞后效应。

3. 碲化镉靶摄像管

碲化镉靶摄像管(又称 Newvicon 管)的灵敏度很高, 可在星光下获得可用的图象, 多用于微光电视。整管的结构和工作过程也与硅靶管类似, 只是靶材料不同。

Newvicon 管采用 II-VI 族化合物半导体异质结蒸发靶。化学成分为碲化镉 $Zn_xCd_{1-x}Te$ 。



Newvicon 管靶的结构示意图

由图可见, 窗口玻璃内壁涂有一层 SnO_2 薄膜为信号板, 接着就是靶。靶有三层结构, 第一层为 $ZnSe$, 属于 N 型半导体, 厚 $50 \sim 100nm$ 。第二层为碲化镉和碲化锌的固溶体 ($Zn_xCd_{1-x}Te$), 属于 P 型半导体, 厚 $3 \sim 5 \mu m$ 。第三层是无定形三硫化二锑 Sb_2O_3 , 厚 $100nm$ 。第一层与第二层之间形成异质结。第二层与第三层之间不形成结。第一层 $ZnSe$ 无光电效应, 它的作用是增强对短波光的吸收, 提高整个可见光区的灵敏度, 另外它还可以阻止光生空穴向成像面一边扩散, 有提高灵敏度, 减小暗电流的作用。光电效应主要发生在第二层, x 值的大小对灵敏度、暗电流和光谱特性都有较大的影响, x 值小, 灵敏度高, 体内暗电流增大, 光谱特性的峰值波长向长波方面移动。第三层的作用是减小扫描电子束的电子注入效应(图象上的浮雕现象), 减小暗电流和惰性。

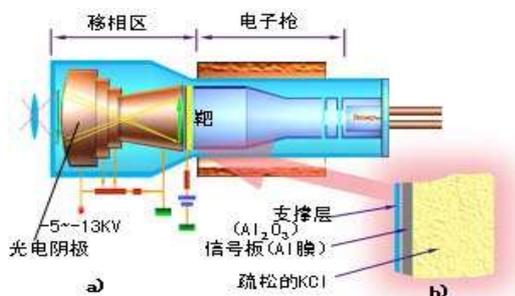
Newvicon 管的特点是, 在可见光范围内灵敏度比硅靶管高, 约为硅靶管的 $1.5 \sim 2$ 倍, 晕光现象比硅靶管少, 工艺简单, 成本低, 但在红外灵敏度, 抗烧蚀性和惰性等方面都不如硅靶管好。靶电压为 $10 \sim 25V$ 。

Newvicon 管适于在低照度下监视, X 射线电视摄像及显微电视摄像等方面应用。

光电发射式摄像管

1. 二次电子导电摄像管 (SEC)

SEC 管是 60 年代初出现的一种高灵敏度摄像管, 适于微光摄像。SEC 管的基本结构分三部分: 光电阴极, 靶和电子枪。光电阴极和靶合在一起统称为移象区。



SEC 管的基本结构简图

SEC 管的工作原理是: 光学图象投射到光电阴极上, 产生相应的光电子发射, 在加速电场和聚焦线圈所产生的磁场共同作用下, 使它打到 SEC 靶上, 产生二次电子发射。于是在靶的扫描面即形成与之对应的电势起伏, 从而使光电阴极上的光学图象转移到 SEC 靶, 变为靶扫描面上的电势分布图象。靶的二次电子发射系数越大, 移象后靶上的电势起伏也就越大。以后, 视频信号的输出过程, 就与光电导式摄像管完全相同了。

SEC 靶有三层结构。第一层是 Al_2O_3 膜, 起机械支撑作用, 厚 $50\sim 70\text{nm}$; 第二层是铝膜, 厚 50nm , 起信号板作用, 通过负载电阻和靶电源的正极相接; 第三层是疏松的 KCl , 其密度只有固体 KCl 的 $1\%\sim 2\%$, 厚 $15\sim 20\ \mu\text{m}$ 。它是 KCl 在氩气中蒸发形成的纤维状薄层, 在慢电子束扫描下, 靶的扫描面稳定在零电位, 因此靶的两面承受着一定的靶电压。当光电阴极发出来的光电子被加速后打到靶上时, 靶将产生二次电子发射。这些二次电子除了极小一部分与正离子复合外, 绝大部分在靶电场的作用下流向信号板, 而在靶的扫描面留下一个正电势图象。读出时, 扫描电子束补充靶上失去的电子, 同时把被扫点电势拉回到零。这样, 在输出回路中即产生视频电流。由此可见, 这种靶的导电, 不是利用材料导带中的电子或价带中的空穴, 而是依靠二次电子导电的, 故称它为二次电子导电摄像管。实验证明, 靶的电子增益与一次电子的加速电压有关。加速电压为 8kV 时, 电子增益最大。如果大于 8kV , 一部分一次电子就要穿透靶层, 对产生二次电子发射无贡献, 从而使电子增益降低。若小于 8kV , 一部分一次电子又可能因没有足够的能量穿过 Al_2O_3 层和 Al 层而达不到靶层, 或因能量损失较多, 靶的电子增益也要降低。如果加速电压为 8kV 时, 约有 2keV 的能量消耗于穿透 Al_2O_3 和 Al 层, 约有 2keV 的能量消耗于穿透 SEC 靶而未发挥作用, 实际被靶吸收起发射二次电子作用的能量只约为 4keV 左右。对 KCl 靶来说, 每产生一个二次电子约要消耗 30eV 的能量, 因此靶的二次电子增益约为 120 左右。

SEC 管的主要特性:

1) 灵敏度

SEC 管整管的灵敏度与光电阴极的灵敏度、加速电压和靶有关, 一般可达 $20000\ \mu\text{A}/\text{lm}$ 。

2) 光电特性

照度小于 10^{-2}lx 时, $\gamma = 1$ 。照度大于 10^{-1}lx 以后, γ 变小, 接近于 0.6。照度大于 1lx 时, 光电流与照度的关系曲线出现拐点, 输出电流开始饱和, 如果长时间工作在这种状态下, 可使靶面烧伤。

3) 分辨率

25mm 管中心分辨率约为 $600\text{TVL}/\text{H}$

4) 惰性

三场后残余信号小于 5%。

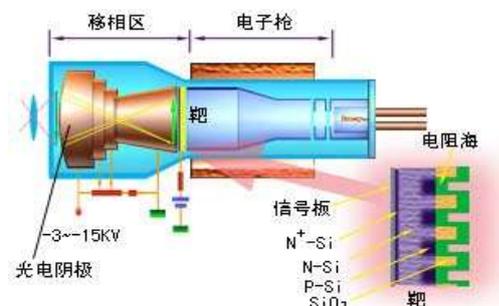
5) 存储性能

SEC 靶的电阻率大于 $10^{16}\ \Omega \cdot \text{cm}$, 漏电极小, 暗电流小于 $0.1\ \mu\text{A}$, 因此信息电荷可在靶上长时间存储而

不致泄漏。摄象管曝光数分钟后，仍可保持原清晰度。甚至 24h 后仍可扫描读出信号。利用这一特点，在摄取相对静止的弱亮度对象时，可通过长时间连续曝光，积累信号，然后再用慢电子扫描或间歇扫描读出，可得到较高的信噪比图象。存储性能好，是 SEC 管突出的特点，多被用于天文学、原子核物理、医疗和安全检查等方面。

2. 硅增强靶摄象管 (SIT)

SIT 管的结构和 SEC 管类似，只是靶不同。SIT 管中的靶和硅靶摄象管中的靶基本相同。



SIT 管基本结构简图

从光电阴极发射出来的电子，在高的电压作用下去轰击硅靶，使靶内产生电子空穴对。经理论计算，每个入射的光电子能量为 3.4~3.5eV 时，可产生 1 个电子空穴对。如果移象区的电压为 10kV，则每个入射光电子可产生 2800~2900 个电子空穴对。但由于表面和体内的复合和收集率等原因，实际的电子增益约为 2000 倍。通过改变移象区的加速电压可改变靶的电子增益。

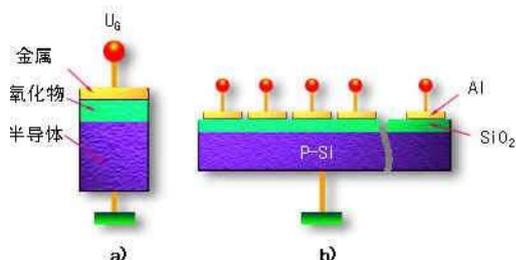
各种摄象管的性能比较

管种 \ 性能	灵敏度/ ($\mu\text{A}\cdot\text{lm}^{-1}$)	极限 分辨率/ (TVL·H ⁻¹)	惰性 (三场后) (%)	暗电流 / μA	γ 值	动态 范围	抗烧伤 能力	晕光现象	工作温度 /°C
硫化镉管	170	800	20~25	20	0.65	350:1	差	很小	+10~+40
氧化铅管	350~400	750	~3	2	0.95	60:1	较差		-30~+50
硅靶管	4350	700	7~10	10	1	50:1	极好	严重	-40~+50
硫化镉管	2600	800	10~20	<1	0.95	60:1	很好	小	+25~+35
砷化镓管	350	900	<2	<1	1		较好		-20~+60
硫化锌镉管	4300	800	<20	10	1		较好	很小	<60
二次电子导电摄像管(SEC)	2×10^4	600	5	<0.1		40:1	最差		
硅增强靶管(SIT)	3×10^5	700	5~7	<10		50:1	最好		

8.3 CCD 固体摄像器件

CCD 结构与原理

1) CCD 的单元结构



CCD 单元与线阵列结构的示意图

a) CCD 单元 b) CCD 线阵列

CCD 单元部分,就是一个由金属-氧化物-半导体组成的电容器,简称 MOS(Metal -Oxide -Semiconductor) 结构。如果衬底接地,突然给金属极板加一个正的电压 U_g (栅极电压),则金属极板和衬底之间就会产生一个电场。这个电场就要迫使半导体表面部分的空穴离开表面入地,从而在表面附近形成一个带负电荷的耗尽区,这个耗尽区也称为表面势阱。表面势阱的深度,近似地与极板上所加的电压成正比(在形成反型层之前)。这时,电子在表面处的势能为 $E_p = -qU_s$, 其中的 U_s 称为表面势,即半导体表面对于衬底的电势差。如果以某种方式(电注入或光注入)向势阱中注入电子,则这些电子将要聚集于表面附近,称为电荷包。

2) 电荷包的储存

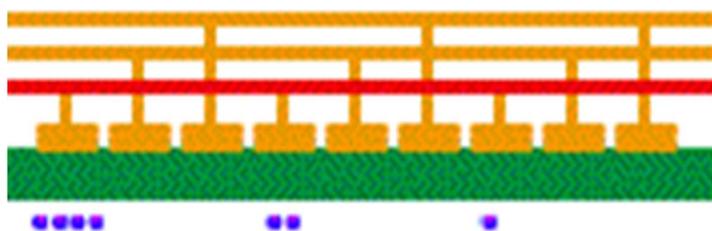
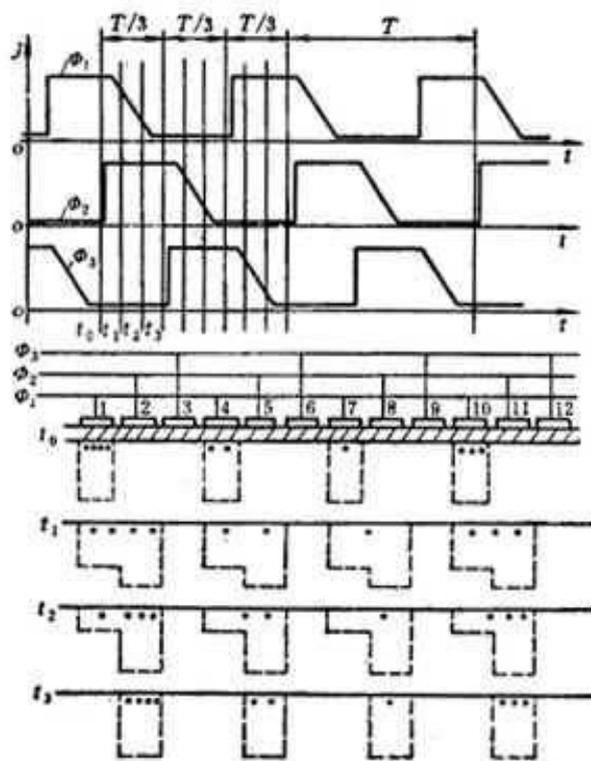
因为每个 CCD 单元都是一个电容器,所以它能储存电荷。但是,当有电荷包注入时,势阱深度将随之变浅,因为它始终要保持极板上的正电荷总量恒等于势阱中自由电荷加上负离子的总和。每个极板下的势阱中所能储存的最大信息电荷量 Q 为

$$Q = C_{ox} U_g$$

C_{ox} : 单位面积氧化层的电容

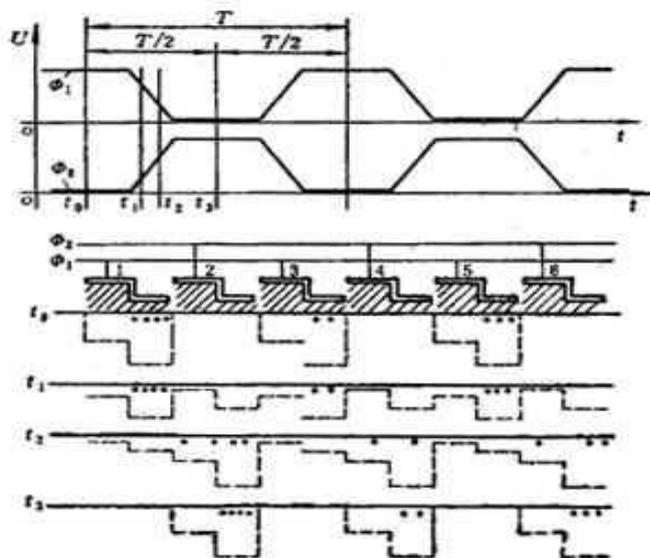
3) 电荷包的转移

CCD 中电荷包的转移是由各极板下面的势阱不对称和势阱耦合引起的。将线阵列各极板分为三组,然后分别加以相位不同的时钟脉冲驱动,这即是所谓的三相 CCD。这时,由于同一时刻三相脉冲的电平不同,各极板下面所造成的势阱深度也就不同。从而电荷包就要沿着表面从电势能高的地方向电势能低的地方流动。



三相 CCD 的电荷包转移过程

三相 CCD 的时钟波形刚好互相错开 $T/3$ 周期，因此时钟电压波形每变化 $T/3$ 周期，电荷包就要转移过一个极板，每变化一个周期，即转移过三个极板。同理，除了有三相 CCD 外，还有二相的、四相的 CCD。



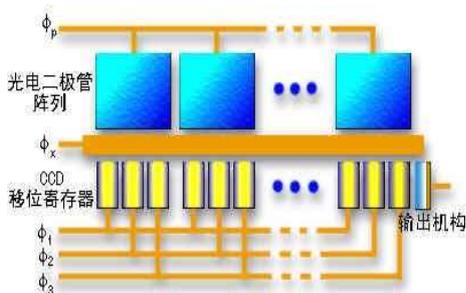
二相 CCD 的电荷包转移过程

二相 CCD 的时钟波形对称，但氧化层 (SiO_2) 厚度不均匀，从而极板下面的势阱也不均匀。因此电荷包也会沿着表面从电势能高的地方向电势能低的地方流动。对于二相 CCD，时钟电压波形每变化 $T/2$ ，电荷包将转移过一个极板，每变化一个周期，则转移过二个极板。由此可见，CCD 具有移位寄存器的功能。

CCD 摄像器件

实用固体摄像器件都是在一块硅片上同时制作出光电二极管阵列和 CCD 移位寄存器两部分。光电二极管阵列专门用来完成光电变换和光积分，CCD 移位寄存器专门用来完成光生电荷转移。因为这种转移不是借助于外来的扫描，而是依靠驱动脉冲来完成的，故也称为自扫描。根据光敏像素的排列方式，CCD 摄像器件分为线阵列和面阵列两大类。

1) 线阵列 CCD 摄像器件



线阵列固体摄像器件基本结构简图

对于线阵列 CCD 摄像器件来说，不论是三相的还是二相的，都有单侧传输和双侧传输两种结构形式。单侧传输的特点是结构简单，但电荷包转移所经过的极板数多，传输效率低。双侧传输的特点是结构复杂一些，但电荷包转移所经过的极板数只是单侧传输的一半，所以损耗小，传输效率高。一般光敏元位数少的片子，多采用单侧传输结构，而位数多的片子，则多采用双侧传输结构。

光电二极管阵列和 CCD 移位寄存器统一集成在一块半导体硅片上，分别由不同的脉冲驱动。设衬底为 P-Si，光电二极管阵列中各单元彼此被 SiO_2 隔离开，排成一行，每个光电二极管即为一个像素。各光电二极管的光电变换作用和光生电荷的存储作用，与分立元件时的原理相同。如图中 Φ_p (行扫描电压) 为高电平时，各光电二极管为反偏置，光生的电子空穴对中的空穴被 PN 结的内电场排斥，通过衬底入地，而电子则积存于 PN 结的耗尽区中。在入射光的持续照射下，内电场的分离作用也在持续地进行，从而即可得到光生电荷的积累。

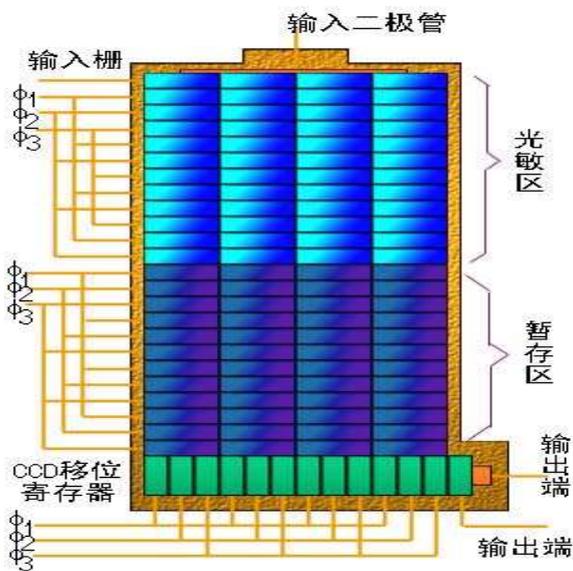
转移栅 (Φ_x) 由铝条或多晶硅构成，转移栅接低电平时，在它下面的衬底中将形成高势垒，使光电二极管阵列与 CCD 移位寄存器彼此隔离。转移栅接高电平时，它下面衬底中的势垒被拆除，成为光生电荷 (电荷包) 流入 CCD 的通道。这时，电荷包并行地流入 CCD 移位寄存器，接着，在驱动脉冲的作用下，电荷包按着它在 CCD 中的空间顺序，通过输出机构串行地转移出去。

对于二相 CCD，时钟电压波形，每变化 $T/2$ ，电荷包将转移过一个极板，变化一个周期，则转移过二个极板。因为二相 CCD 是二个极板对应着一个光敏元，所以时钟波形变化一个周期，电荷包所转移过的空间距离也是一个光敏元的中心距。对于三相 CCD，时钟电压波形每变化 $T/3$ 周期，电荷包就要转移过一个极板，每变化一个周期，即转移过三个极板，时钟电压波形变化一个周期，电荷包所转移过的空间距离，正好是一个光敏元的中心距。这就是线阵列固体摄像器件大致的工作过程。

2) 面阵列 CCD 摄像器件

二维固体摄像器件中，电荷包转移情况与线阵列器件类似，只是它的形式较多。有的结构简单，但摄像质量不好，有的摄像质量好些，但驱动电路复杂，目前比较常用的形式是帧转移结构。光敏区是由光敏 CCD 阵列构成的，其作用是光电变换和在自扫描正程时间内进行光积分，暂存区是由遮光的 CCD 构成的，它的位数和光敏区一一对应，其作用是在自扫描逆程时间内，迅速地将光敏区里整帧的电荷包转移到它里面暂存

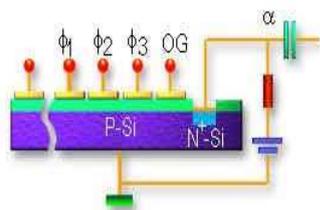
起来。然后，光敏区开始进行第二帧的光积分，而暂存区则利用这个时间，将电荷包一次一行地转移给 CCD 移位寄存器，变为串行信号输出。当 CCD 移位寄存器将其中的电荷包输出完了以后，暂存区里的电荷包再向下移动一行给 CCD 移位寄存器。当暂存区中的电荷包全部转移完毕后，再进行第二帧转移。



二维固体摄像器件电荷包帧转移结构图

电荷包的输出机构

1) 利用二极管的输出机构

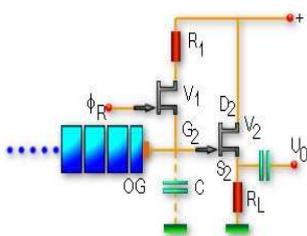


利用二极管的输出机构

CCD 电荷包输出机构的形式很多，其中最简单的是利用二极管的输出机构。

图中，与 $\Phi_1 \Phi_2 \Phi_3$ 相连的电极称为栅极，与 OG 相连的电极称为输出栅，输出栅的右边就是输出二极管。输出栅和其它栅极一样，加正电压时，它下面的半导体表面也产生势阱。它的势阱介于 Φ_3 的势阱和输出二极管耗尽区之间，能够把二者连通起来，因此可以通过改变 OG 上所加的电压来控制它下面的通道。例如，电荷包已由 Φ_2 转入 Φ_3 ，当 Φ_3 下的势阱由深变浅的同时，OG 下的势阱正好也比较深，这时 Φ_3 势阱中的电荷包就能够通过 OG 下的势阱流入输出二极管的耗尽区。因输出二极管是反偏置的，内部有很强的自建电场，因此电荷包一进入二极管的耗尽区，即可被迅速地拉走，成为输出回路的电子流。因此，在没有电荷包输出时，a 点为高电平，而有电荷包输出时，因为电子流通过负载电阻要产生电压降，a 点则为低电平，a 点电压降低的程度正比于电荷包所携带的电量，所以这个电压变化即是输出信号。

2) 选通积分型输出机构



选通积分型输出机构

图中, V_1 为复位管, R_1 为限流电阻, V_2 为输出管, R_2 为负载电阻, C 为等效电容。电荷包输出前, 要先给 V_1 的栅极加一窄的复位脉冲 Φ_R , 这时, V_1 导通, C 被充电到电源电压, V_2 管的源极 S_2 的电压也跟随上升接近于电源电压。 Φ_R 变为低电平以后, V_1 截止, 但 V_2 在栅极电压的控制下仍为导通状态。当电荷包经过输出栅 OG 流过来时, C 被放电, V_2 的源极电压也跟随下降, 下降的程度则正比于电荷包所携带的电量, 即构成输出信号。

CCD 的主要参量

1) 转移效率 η 和损耗率 ε

电荷包从一个势阱向另一个势阱中转移, 不是立即的和全部的, 而是有一个过程。为了描述电荷包转移的不完全性, 引入转移效率的概念。在一定的时钟脉冲驱动下, 设电荷包的原电量为 Q_0 , 转移到下一个势阱时电荷包的电量为 Q_1 , 则转移效率 η 定义为

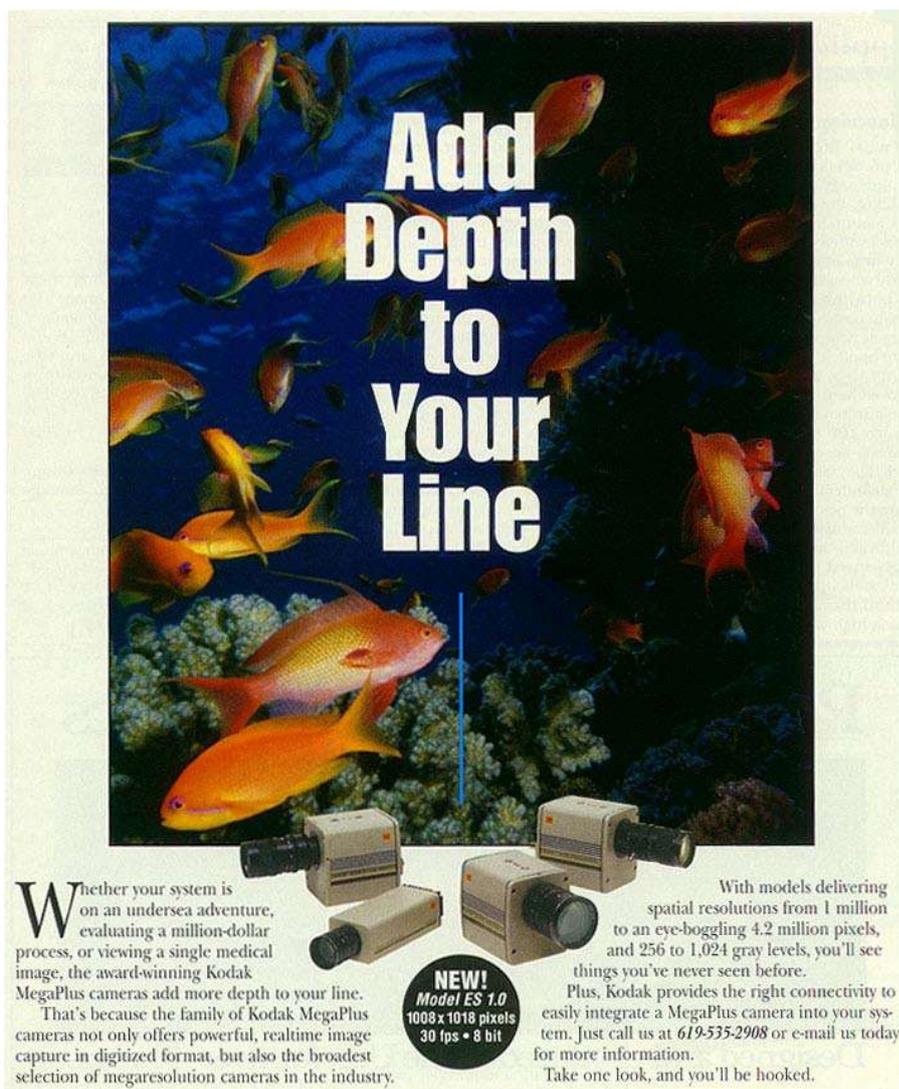
$$\eta = Q_1 / Q_0$$

ε 表示残留于原势阱中的电量与原电量之比, 故

$$\varepsilon = 1 - \eta$$

如果线阵列 CCD 共有 n 个极板, 则总效率为 η^n 。

引起电荷包转移不完全的主要原因是表面态对电子的俘获和时钟频率过高, 所以表面沟道 CCD 在使用时, 为了减少损耗, 提高转移效率, 常采用偏置电荷技术, 即在接收信息电荷之前, 就先给每个势阱都输入一定量的背景电荷, 使表面态填满。这样, 即使是零信息, 势阱中也有一定量的电荷。因此, 也称这种技术为“胖零 (fat zero)”技术。另外, 体内沟道 CCD 采取体内沟道的传输形式, 有效避免了表面态俘获, 提高了转移效率和速度。



Add Depth to Your Line

Whether your system is on an undersea adventure, evaluating a million-dollar process, or viewing a single medical image, the award-winning Kodak MegaPlus cameras add more depth to your line. That's because the family of Kodak MegaPlus cameras not only offers powerful, realtime image capture in digitized format, but also the broadest selection of megaresolution cameras in the industry.

With models delivering spatial resolutions from 1 million to an eye-boggling 4.2 million pixels, and 256 to 1,024 gray levels, you'll see things you've never seen before. Plus, Kodak provides the right connectivity to easily integrate a MegaPlus camera into your system. Just call us at 619-535-2908 or e-mail us today for more information. Take one look, and you'll be hooked.

NEW!
Model ES 1.0
1008 x 1018 pixels
30 fps • 8 bit

2) 时钟频率的上、下限

CCD 是利用极板下半导体表面势阱的变化来储存和转移信息电荷的，所以它必须工作于非热平衡态。时钟频率过低，热生载流子就会混入到信息电荷包中去而引起失真，时钟频率过高，电荷包来不及完全转移，势阱形状就变了，这样，残留于原势阱中的电荷就必然多，损耗率就必然大。因此，使用时，对时钟频率的上、下限要有一个大致的估计。

(a) 时钟频率的下限 f_{\downarrow}

f_{\downarrow} 决定于非平衡载流子的平均寿命 τ ，一般为毫秒量级。电荷包在相邻两电极之间的转移时间 t ，应小于 τ ，对于三相 CCD，电荷包从前一个势阱转移到后一个势阱所需的时间为 $T/3$ ，所以

$$f_{\downarrow} > 1/3 \tau$$

对于二相 CCD，

$$f_{\downarrow} > 1/2 \tau$$

(b) 时钟频率的上限 f_{\uparrow}

f_{\uparrow} 决定于电荷包转移的损耗率 ϵ ，就是说，电荷包的转移要有足够的时间，电荷包转移所需的时间应使之小于所允许的值。时钟频率上限 f_{\uparrow} 可作如下估算，设 τ_D 为 CCD 势阱中电量因热扩散作用衰减的时间常数，与材料和极板的结构有关，一般为 10^{-8} s 级。若使 ϵ 不大于要求的 ϵ_0 值，则对于三相 CCD 有

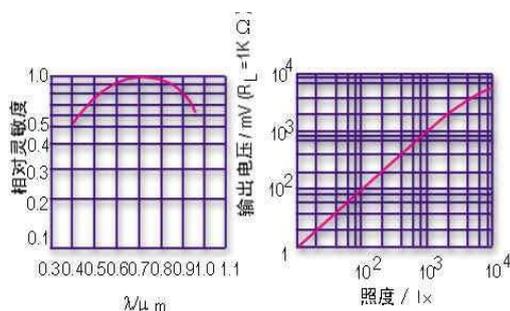
$$f_{\uparrow} \leq -1/(3 \tau_D \ln \epsilon_0)$$

对于二相 CCD 有

$$f_{\perp} \leq -1 / (2 \tau_D \ln \epsilon_0)$$

3) 光谱特性

现在固体摄像器件中的感光元件都是用半导体硅材料来作的，所以灵敏范围为 0.4~1.15 μm 左右，但光谱特性曲线不象单个硅光电二极管那么锐利，峰值波长为 0.65~0.9 μm 左右。



CCD 的光谱特性与光电特性曲线

4) 光电特性

在低照度下，CCD 的输出电压与照度有良好的线性关系。照度超过 100lx 以后，输出有饱和现象。

CCD 摄像器件的应用

CCD 摄像器件为固体器件，它的体积小、重量轻、工作电压低、惰性小、功耗小、输入光的动态范围大、扫描无畸变、机械抗震性能好。线阵 CCD 多用于文字字符识别、传真和尺寸检测，面阵 CCD 主要用于小型摄像机。

第九章 非相干光电信号变换

在光电系统的信号形成、传送、检测和处理过程中，通常要借助于几何光学的、物理光学的或光电子学的方法对信号的组编形式和能量状态进行人为的变换，包括将一种光量转换为另一种光量，将非光量转换为光量或将连续光量变换为脉冲光量等。

这种变换的目的在于：将待处理信息载荷到光载波上进而形成光电信号；改善系统的时间空间分辨能力和动态品质，提高传输和检测精度；改善系统的检测信噪比，提高工作可靠性和对环境的适应能力。

典型光电信号变换方法的光学原理和应用

变换方法	光学原理	应用
几何光学	直进、遮光；反射、折射；光学成像等非相干光的光学现象。	光开关、光学编码、光学扫描、瞄准定位、准直定向、测长测角、测距、成像测量等。

物理光学	衍射、干涉、光谱、波长变换、光学拍频等相干光的光学现象。	莫尔条纹、干涉计量、外差通讯、光谱分析、散斑全息测量等。
光电子学	电光效应、磁光效应、声光效应、空间光调制、光束纤维传输传感等光电现象。	光调制、光束偏转、光通讯、光记录、光存储、光显示、传光、传象、传感等。

几何变换的光电方法是基于将光学现象看作是直线光束传输的结果，是在几何光学的意义上利用光束传播的直线性、透光、反射、折射、成象等光学方法进行的信号变换。

物理变换的光电方法是将光学现象看作是电磁波振荡传播的结果，是在物理光学的意义上利用光的衍射、干涉、光谱、波长变换等光学现象和参量进行的信号变换。

随着光电子学的发展，出现了各种新型的光控制器件，它们为光信号的实时变换和光路的集成化提供了广阔的发展前景。这些变换方法常常单独或者若干组合后应用于不同目的的光电系统中。

本章将介绍对非相干光采用几何变换的各种光电方法。

根据光电信号的时空特点，在光电系统中常用的各种信号大致可分作下面的两种类型：

1) 随时间变化的光电信号。这类信号的变换通常集中发生于局部的空间范围，它的特征参量只随时间缓慢地，或周期性或瞬时地变化，因此是时间的一维函数而与空间坐标无关，可用函数 $F(t)$ 表示。

2) 随空间变化的光电信号。这类信号发生在一定的空间坐标范围内，光电信号的参量随空间位置而改变。因此是二维或三维空间的函数，用函数 $F(x, y, z)$ 表示，其中随时间变化的信号称作时空变化的光电信号，用函数 $F(x, y, z, t)$ 表示。

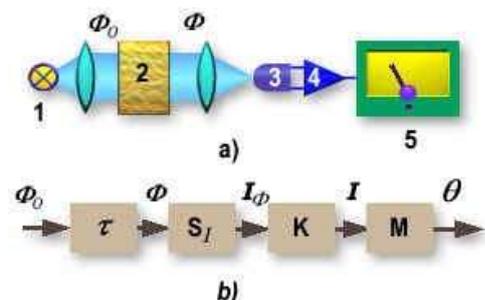
9.1 时变光信号的直接测量

光电系统所处理的输入信号有一类是辐射光通量直接随被测信息变化的。这种情况下的被测信息是载荷在光通量的幅度大小、变化频率或变化相位上，检测这些参量可以测量出所需要的信息。

例如，直接测量来自被测物体的光辐射进行的光谱分析和辐射测温；通过检测被测物体的透光率测量气体液体的浓度或薄膜厚度；以及通过测量反射率检测物体的表面状态等。下面介绍直接随信息改变的光通量测量方法。

光通量的幅度测量

1. 单通道测量系统



单通道测量系统和直读法

a) 原理示意图 b) 方框图

1-辐射源 2-被测样品（或标准样品） 3-检测器

4-光电信号放大器 5-指示仪表

被测光通量沿单一光学通道传送到光电接收器，经由读出装置进行测量的系统称作单通道测量系统。图中，由辐射源 1 产生的光通量 Φ_0 ，经布置于光路上的被测样品或标准样品 2 的部分吸收或散射，投射到检测器 3 的光敏表面上。利用光电信号放大器 4 可以测量出载有信息的光通量 Φ ，最后由指示仪表 5 显示记录下被测量的结果。单通道测量系统常用的测量方法是直读法和指零法。

1) 直读法

在单通道系统中，直接测量载荷信息的光通量值，并通过与标准样品的比对测定被测量的方法称作直读法。例如为了测量样品的透过率，可以在图 a 所示的光路中事先分别放置透过率已知的标准样品，每次对电表值进行标定。经过这样的处理后再将被测样品放在光路中，即可由电表的示值直接得到被测的透过率值。

直读法测量系统的方框图示于图 b 中。设照明样品的光通量为 Φ_0 ，样品透过率为 τ ，透过样品并为光电检测器接收的光通量为 Φ ，检测器的电流灵敏度为 S_I ，放大器增益为 K ，读出装置的传递系数为 M ，输出值为 θ ，则直读系统的测量值可表示为

$$\theta = S_I K M \Phi_0 \tau = K_0 \tau$$

或

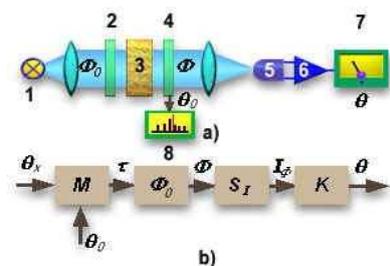
$$\tau = \theta / K_0$$

式中 $K_0 = S_I K M \Phi_0$ 。

当 K_0 保持恒定不变时，被测值 τ 和读出示值 θ 之间存在着一一对应的关系，可根据 θ 确定被测的 τ 值。但实际上 K_0 值不是完全不变的。特别是 Φ_0 因子与照明状态有关，存在着长期工作的不稳定性 and 不重复性。比例因子 K_0 的波动会直接造成测量示值的偏差。此外定标所用基准的误差或在定标过程中引入的偏差也会直接影响测量结果，因此直读法在实际应用中受到一定的限制。

2) 指零法

提高单通道测量系统测量精度的简单方法是指零法或零示法。这种方法利用标定好的读数装置补偿被测光通量值，使测量系统在输出光通量为零的状态下读数。它可以消除由于辐射光通量、光检测器和放大器工作参数不稳定造成的影响。



采用指零法的单通道测量系统

a) 原理示意图 b) 方框图

1-光源 2-起偏器 3-被测物 4-检偏器

5-检测器 6-放大器 7-指示仪表 8-读数装置

上图表示了采用指零法测量偏振物质偏振角的光电装置原理示意图。在如图的光路中布置起偏器 2 和检偏器 4。事先使检偏器的偏振面相对起偏器转 90° ，此时检测器 5 的光敏面上的光通量为零。在起偏器 2 和检偏器 4 之间放置具有偏光性质的被测样品，它引起偏振面的旋转，增大了透过光的数值。借助于指示仪表 7 可以观察到相对于零位的偏差。若转动检偏器 4，重新使指示仪表示值为零，则检偏器的转角等于由被测样品引起的偏振面的转角。用高精度的角度基准对检偏器转角逐点标定后，只要测量出检偏器相对零位的转角即可得到所需的偏振面旋转值。

利用图 b 中的方框图可以计算出系统的测量关系。设被测偏振面转角为 θ_x ，检偏器的读出转角为 θ_0 ，偏振器组合的相对转角为 $\theta_x - \theta_0$ ，对于透光率 τ 的转换因子为 M ，则检偏-起偏器组合和指示仪表的输出值 θ 分别可表示为：

$$\tau = M(\theta_x - \theta_0)$$

和

$$\theta = KS_1\Phi_0\tau$$

由以上二式有

$$\theta = KS_1\Phi_0M(\theta_x - \theta_0) = K_0(\theta_x - \theta_0)$$

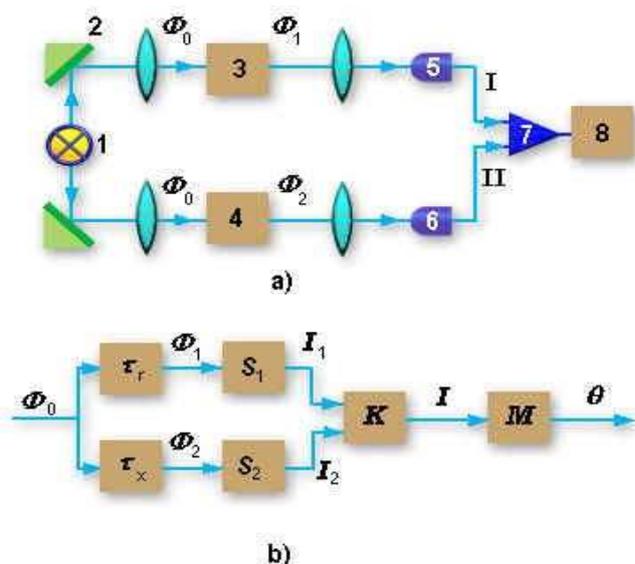
式中， $K_0 = KS_1\Phi_0M$ ， Φ_0 为入射照明光通量， S_1 为光检测器灵敏度， K 为放大器增益。在无试样的校准状态下，即 $\theta_x = 0$ 时，调整 θ_0 ，使 $\theta = 0$ 。随后加入试样则 $\theta_x \neq 0$ ，因而 $\theta \neq 0$ 。重新调整 θ_0 ，使 θ 再次为零，此时有

$$\theta_x = \theta_0$$

上式表明：在指示仪表指零状态下，检偏器的转角示值 θ_0 代表了被测偏振面的旋转值 θ_x 。

在上述系统中，检测电路没有直接用于读数，而只是用作指示仪表的指零。指零的结果使被测量和读数装置的补偿量相互平衡，此时通过读数装置读出的即为被测值。它的测量精度取决于读数装置的精度和指零装置的零位漂移。光通量的不稳定性对误差影响不大，因为系统的读数是在零光通量的情况下进行的。

2. 双通道测量系统



双通道幅度测量系统及差动法方框图

a) 原理示意图 b) 方框图

1-光源 2-反射镜 3-标准样品 4-被测样品

5、6-光电检测器 7-放大器 8-指示仪表

单通道测量的主要缺点是入射光通量的波动会直接影响测量结果。为了克服这一缺点，采用了具有两个光学通道的测量方法。图 a 给出了用作测量透过率的双通道测量系统示意图。图中，光辐射源 1 借助于反射镜 2 和透镜分别沿着基准通道 I 和测量通道 II 并行传送。在基准通道 I 的光路中放置具有固定透过率的标准样品 3，通过光电检测器 5 监视入射光通量 Φ_0 的变化。在测量通道 II 中放置被测样品 4，透过被测样品的光通量 Φ_2 经光电检测器 6 后在放大器 7 中和基准通道的参考信号相比较，差值信号通过不同形式的比较处理最终得到测量结果。双通道系统不仅能完全消除入射光辐射起伏的影响而且可在一定程度上消除杂光引起的扰动，是光通量幅度测量中常用的方法。双通道系统的信号处理方法主要包括差动法、补偿法和交替比较法等。

1) 差动法

在双通道系统中，直接利用测量通道和基准通道中两束光通量的差值确定被测值与标准值间偏差的方法称作差动法。系统的组成如图所示。图中，设入射光通量为 Φ_0 ，标准样品透过率为 τ_r ，被测样品透过率为 τ_x ，光探测器灵敏度分别为 S_1 和 S_2 ，放大器增益为 K ，指示装置的传递系数为 M ，则输出指示值 θ 可表示为

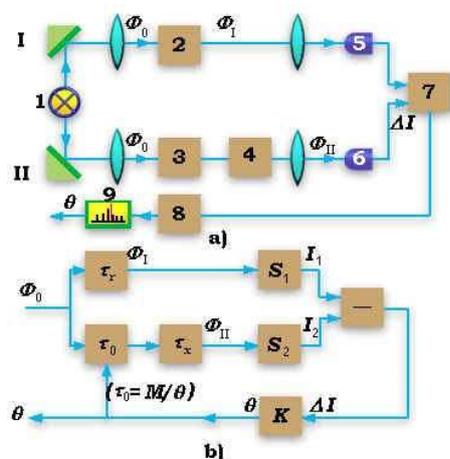
$$\theta = KM\Phi_0(S_2\tau_x - S_1\tau_r)$$

对于对称的光探测器有 $S_1 = S_2 = S$ ，则上式变成

$$\theta = KM\Phi_0S(\tau_x - \tau_r)$$

上式表明，采用双通道差动法能直接得到被测值和标准值间的偏差，故适用于二个光量的比较测量。对于较小的偏差，二个通道的增益可以很高，这有助于提高系统的测量精度，差动法对共模形式引入的外来光干扰有抑制能力，可以消除杂散光的影响。

2) 补偿法



差动补偿法双通道测量透过率的示意图和方框图

- 1-光源 2-标准样品 3-可变透过率板 4-被测样品
- 5、6-光电检测器 7-放大器 8-伺服电动机 9-读出装置

上图给出了补偿法测量透过率的双通道系统示意图。这个系统的基本组成和双通道测量系统相似。在它的测量通道中装有可变透过率板 3，它是一块玻璃平板，沿着各截面镀有吸收率不同的膜层。在装上被测样品 4 之前，透过率板处于最大吸收位置，并使二通道的输出光通量相等，处于平衡状态。当插入被测样品 4 之后，测量通道的光通量减少。此时若移动板 3 改变透过率值，使其上透过率的增大恰好等于被测样品的吸收值，这就可以使二个通道重新达到平衡。透过率板的转动由与之相连的读数机构 9 显示，它的示值与被测样品的透过率相对应。这样，在二通道的输出光通量相等时，透过率板的位置或读数机构示值就是被测透过率的量度。这个过程通常是自动完成的，这时二个通道的光通量分别由光电检测器 5 和 6 接收，由差动放大器 7 得到增益。放大器的输出电压接到伺服电机 8 的控制绕组上，当两个通道的光通量不相等时，放大器的输出端产生控制电压，使电机轴带动板 3 移动，一直到输出电压为零停止。控制电机的转动方向是根据光通量 Φ_I 和 Φ_{II} 的比较关系，按照负反馈的连接方法确定的。

对差动补偿法双通道系统，根据图 b 的方框图可以计算出差动放大器的失调信号为

$$\Delta I = I_I - I_{II} = S_1\Phi_0\tau_r - S_2\Phi_0\tau_0\tau_x$$

式中， τ_r 是基准通道标准样品的透过率， τ_x 是被测透过率， τ_0 是可变透过率，它是输出读数 θ 的函数。并有关系式

$$\tau_0 = M(1/\theta)$$

式中， M 为比例因子。

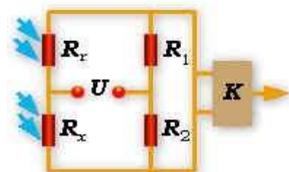
当两个通道平衡时， $\Delta I = 0$ 。有

$$S_1\Phi_0\tau_r = S_2\Phi_0\tau_0\tau_x$$

或

$$\tau_x = (S_1/S_2) (\tau_r / \tau_0) = (S_1/S_2) (\tau_r/M) \theta = K_0 \theta$$

式中, $K_0 = (S_1/S_2) (\tau_r/M)$, 上式表示在两通道平衡时被测透过率与读数示值的对应关系。可以看出比例因子 K_0 与入射光通量无关, 从而减少了辐射照明不稳定对测量误差的影响。双通道系统的这种补偿作用是由于它的入射光通量是以同样的效果影响两个通道, 而它们的差值在准确平衡的条件下接近于零, 因而对测量结果没有影响。



比较补偿法双通道系统的光敏电阻电桥

在前述的双通道系统中, 如果不采用差动放大器接收而用比较测量法也可以得到同样的补偿效果。比较测量中的光电接收器可以采用组成电桥的光敏电阻。图中参考用光敏电阻 R_r 接于基准通道, 测量用光敏电阻 R_x 接于测量通道。设光敏电阻灵敏度为 S_1 、 S_2 , 则其电导值 G_r 和 G_x 可表示为

$$G_r = 1/R_r = S_1 \tau_r \Phi_0$$

和

$$G_x = 1/R_x = S_2 \tau_x \tau_0 \Phi_0$$

当电路平衡时有

$$R_2/G_r = R_1/G_x$$

$$\tau_x = (S_1/S_2) (R_1/R_2) (\tau_r / \tau_0)$$

$$= (S_1/S_2) (R_1/R_2) (\tau_r/M) \theta = K_1 \theta$$

式中, $K_1 = (S_1/S_2) (R_1/R_2) (\tau_r/M)$ 。

上式表示被测透过率 τ_x 和读数示值 θ 间的定量关系。比例因子 K_1 与入射光通量 Φ_0 无关, 因此, 比较补偿法和差动补偿法一样, 具有抑制共模扰动的作用。

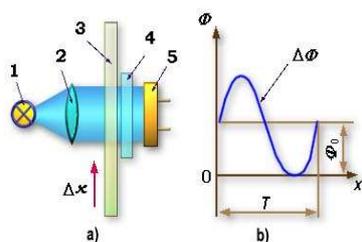
光通量的频率测量

某些光电系统所处理的光通量是随被测信息的规律呈周期性变化的 (正弦的或方波的等)。这时, 被测信息载荷在光通量变化的次数多少和频率快慢之中。光信号经过光电转换后, 采用常规电信号的波数和频率测量方法就可以检测出所需要的信息。

使光通量的频率或波数随被测信息变化的方法是多种多样的。例如, 简易光电转速表中采用的透光孔圆盘或反射式标志是用几何光学的透光、反光的方法使光通量随转角或转速变化; 工业制品的计数检测是使光通量随制品个数变化; 在计量技术中广泛采用的莫尔条纹或干涉条纹测量是用物理光学方法使光通量随机械位移而周期性变化。

和光通量的幅度测量相比较, 频率测量有较高的测量精度, 这是由于它的测量基准可以达到很高的水平。例如, 利用专门的稳频装置, 激光器的频率稳定度可达 10^{-14} 。通常的电信号时钟基准也可达到 $10^{-6} \sim 10^{-8}$ 的相对精度, 而幅度测量的基准通常只能达到 $10^{-3} \sim 10^{-4}$ 。这就决定了在一般情况下光通量的幅度测量只能有 $10^{-2} \sim 10^{-3}$ 数量级的相对精度, 而频率测量其相对精度却不低于 $10^{-5} \sim 10^{-7}$ 。此外频率测量是数字式的, 测量结果易于和计算机连接, 使用较为方便。因此在现代光电测量中常优先考虑采用频率测量法。

1. 波数测量



莫尔条纹计数和条纹光强分布

a) 光栅头示意图 b) 波形图

1-照明灯 2-聚光镜 3-主光栅尺 4-指示光栅 5-光电器件

图中给出了常见的计量光栅读数头示意图。当光栅主尺位移 Δx 时引起光敏元件上光通量 $\Delta \Phi$ 的变化 (如图 b), 它们的关系可近似表示为

$$\Delta \Phi = \Phi_0 \cos(\Delta x \cdot 2\pi / T)$$

式中, T 是光栅的栅距, Φ_0 是光通量变化的幅度。由图和上式可知, 当 Δx 移动 T 长度时引起 $\Delta \Phi$ 一个周期的变化, 用波数 N 表示 $\Delta \Phi$ 变化的周期数, 可以得到位移 x 和波数 N 的关系为

$$x = TN$$

利用上式, 只要测得波数值 N 即可定量计算出位移值 x 。

通过测量光通量随信息变化的周期数来检测被测值的方法称作波数测量或波形计数。在一般情况下波数测量的作法是将由被测值 x 引起的周期变化的光通量转换为电脉冲, 再用电子计数器来计算波数的变化 N , 最后计算出测量结果 x 为

$$x = m_0 N$$

式中, m_0 称脉冲当量, 表示单位波数变化对应的被测值。在前述莫尔条纹计数装置中 $m_0 = T$, 脉冲当量可通过计算和实验标定得到。波形内的精确测量通常采用条纹细分技术。

2. 频率测量

前述的莫尔条纹装置还可以用来测量光栅主尺的运动速度。当将被测物体安装在主尺上并以速度 v 移动时, 莫尔条纹的光强分布将随时间周期性变化。为了确定接收光通量的变化频率 f 与运动速度 v 的关系有

$$dx/dt = T \cdot dN/dt$$

或

$$v = Tf$$

式中, $f = dN/dt$ 是波数的时间变化率。

由于 T 是常数, 因此运动速度与光通量的变化频率成正比。通过测量光通量变化的频率来测定被测参量的方法称作频率测量法。频率测量主要应用于各种物理量的速率测量中。此时, 应首先将光通量的变化速率转换为电脉冲的频率, 采用电子频率计测量脉冲频率 f , 最后根据脉冲当量值 m_0 可计算出被测速率 v

$$v = m_0 f$$

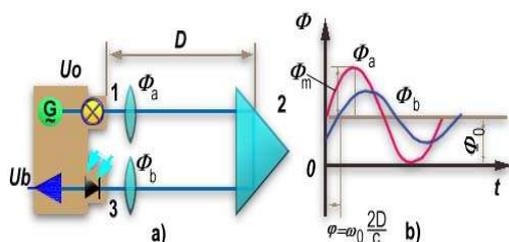
式中, v 是线速度、角速度或其它变量的时间变化率。脉冲当量 m_0 可通过计算和实验标定。

波形计数和频率测量法能有十几位的读数能力和示值精度, 并且光电接收器和光源系统的不稳定性实际上对测量结果影响不大。信号幅度的改变, 将不影响到计数和测频的准确性, 这是该种方法的另一重要特点。

光通量的相位和时间测量

1. 相位测量

在光电检测系统中辐射光通量随被测信息进行周期性变化。若在某一瞬间信号的变化相位代表了被测信息的瞬时值, 那么检测到这个相位值即能确定被测值, 这种方法称作光通量的相位测量。典型的光通量相位测量的实例是光电光波测距。



相位测量法光波测距

a) 测距原理示意图 b) 波形图

1-辐射源 2-协作靶 3-接收器

相位法光波测距系统用半导体激光器或发光二极管作辐射源，它的辐射通量近似地与激励电流成正比。若在光源的供电电路中施加正弦电压 U_0 ，就能得到光强按正弦规律变化的辐射光波。设辐射光通量为 Φ ，则有

$$\Phi = \Phi_0 + \Phi_m \sin(\omega_0 t + \phi_0)$$

式中， Φ_0 为直流分量； Φ_m 是交变分量的振幅； $\omega_0 = 2\pi f_0$ 是光源激励电压即辐射光通量的变化角频率， f_0 是相应的频率；设初始相位 ϕ_0 和激励电压的相位相同，即 $\phi_0 = 0$ 。

光辐射经光学系统发射到称作协作靶的反射镜 2 上，后者被事先装置在距离测距仪为 D 的位置上，反射的光波重新折向测距仪并为光电接收器 3 接收。由于光波传播速度 c 恒定，传播距离 D 与传播时间 t 成正比，所以光波传送到协作靶再返回到测距仪的时间 t 为

$$t = 2D/c$$

在这段时间里光源激励电压 U_0 的相位改变了 ϕ 值

$$\phi = \omega_0 t = \omega_0 \cdot 2D/c$$

这样，光电接收器上的信号电压 U_b 和光源激励电压 U_0 间产生了 ϕ 的相位差（如图 b 所示）。借助于电测技术来测量角 ϕ ，即可以计算出待测距离 D

$$D = \phi \cdot c / 2\omega_0$$

设辐射光通量变化一周时光波传播距离为 L_0 ，则有关系式

$$L_0 = c/f_0$$

同时考虑空气折射率 n 的影响，有

$$D = (L_0/2n) (\phi/2\pi) = (c/2nf_0) (\phi/2\pi)$$

式中， $L_0/2n$ 称作测尺长度，是测距仪的测长单位，它与相对相位差 $\phi/2\pi$ 的乘积即为被测距离。

上式是相位法光波测距的测距公式。由式可知最小可测距离 ΔD_{\min} 为

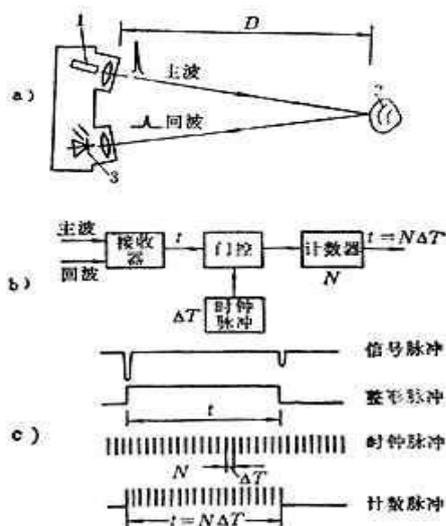
$$\Delta D_{\min} = \Delta \phi_{\min} \cdot c / 4\pi n f_0$$

式中， $\Delta \phi_{\min}$ 是最小相位测量分辨率。

上式表明最小可测距离 ΔD_{\min} 与 $\Delta \phi_{\min}$ 成正比，和光源激励频率 f_0 成反比。

2. 时间测量

利用周期性光通量的周期或单个光脉冲间的时间间隔来测量信息的时间测量法在光电检测中同样有着广泛应用。时间测量和相位测量有许多相似之处。最简单的测时方法是在确定时间起止点之间用时钟脉冲填充计数，这种方法可以得到 10^{-10} s 以上的测时精度。



时间法脉冲激光测距

a) 原理示意图 b) 方框图 c) 波形图

1-激光器 2-目标 3-接收器

脉冲激光测距仪和激光雷达是时间测量法的典型应用，它的原理示意图如图 a。由固体激光器 1 产生的能量为几兆瓦、作用时间为几纳秒、发射角为几毫弧度的激光巨脉冲经光学系统射向被测目标 2，经目标漫反射之后返回到接收系统 3。通过测量主波脉冲和回波脉冲的时间差即可测量出被测距离。

图 b 和 c 给出了脉冲测时的原理框图和波形图。主波脉冲和回波脉冲经光电转换分别形成电脉冲，控制触发器产生确定脉宽 t 的定时脉冲。在此脉冲持续时间内由频率为 $f = 1/\Delta T$ 的时钟脉冲填充计数，若计数数值为 N ，则脉宽持续时间为

$$t = N \Delta T = N/f$$

得

$$D = ct/2 = N \cdot c/2f = KN$$

式中， $K = c/2f$ 是测距脉冲当量，表示单位计数脉冲对应的被测距离。若 $c \approx 299900 \times 10^3 \text{ m/s}$ ，取 $f = 149.900 \text{ MHz}$ ，则有 $K \approx 1 \text{ m/脉冲}$ 。

由上式，在利用脉冲计数法完成时间测量后可根据实际计数值 N 和脉冲当量 K 得到被测距离。最小可测距离 D_{\min} 由时钟脉冲频率 f 决定， f 愈高， D_{\min} 愈小，通常 f 可达 300 MHz ，对应的 D_{\min} 值为 0.5 m 。

9.2 时变光信号的调制和解调

调制的基本原理

在光电系统中，光通量是信息的载体。光通量载荷信息的方法有多种形式。例如物质燃烧过程所形成的光辐射本身就包含着物质内部结构成分的信息。更多的情况是通过人为的变换把被测信息载荷到光通量上。例如在透过率测量系统中，使恒定的光通量通过被测介质，随介质吸收情况的不同输出光通量的数值有所改变，于是被测介质透过率的信息即被载荷到光通量上。这里，光通量作为信息的物质载体称作载波。使光载波信号的一个或几个特征参数按被传送信息的特征变化，以实现信息检测传送目的的方法称作调制。

光载波可以利用复合的非相干光波，也可以利用窄带单色、有确定初相位的相干光波。许多种光学参量都可以作为载波的特征参数，例如非相干光辐射能量的幅度、变化频率和相位以及脉冲时间；相干光的波振

幅、光频、相位、偏振方向以及光束的传播方向或光学介质的光学常数等。众多的可调制参量增加了光载波信号处理的灵活性和多样性。

光载波可以是将信息直接调制到光载波上的广义的调制，在更多的情况下是人为地使载波光通量随时间或空间变化，形成多变量的载波信号，然后再使其特征参数随被测信息改变。这种方法称作二次调制，因为它是对已随时间调制的光通量特征参数的再调制。使光载波参数按确定的时间或空间规律变换，这样做虽然看起来似乎增加了信号的复杂性，但是它有助于传输过程的信号处理和传输能力的提高，能更好地从背景噪声和干扰中分离出有用信号，提高信噪比和测量灵敏度。此外调制信号还能简化检测系统的结构，改善系统的工作品质，有时利用调制还可以扩大目标定位系统的视场和搜索范围。因此调制技术是光电检测系统中常用的方法。光学调制按时空状态和载波性质可分类如下：

按时空状态分类：

- 1) 时间调制：载波随时间和信息变化。
- 2) 空间调制：载波的空间状态按信息规律调制。
- 3) 时空混合调制：载波随时间空间和信息调制变化。

按载波波形和调制方式分类：

- 1) 直流载波：载波不随时间而只随信息变化。
- 2) 交变载波：载波随时间周期变化。

连续载波——调幅波、调频波、调相波。

脉冲载波——脉冲调宽、调幅、调频等。

光通量的调制可以在辐射源或光路系统中进行，能实现调制作用的装置称作调制器。从已调制信号中分离提取出有用信息的过程称作解调。本节将主要介绍有关光信号时间调制和解调的基本原理。

典型的调制方法

1. 连续波调制

连续波调制的光载波通常具有谐波的形式，可用函数描述为

$$\Phi(t) = \Phi_0 + \Phi_m \sin \omega t$$

式中， Φ_0 是光通量的直流分量，一般不载任何信息。 Φ_m 和 ω 是载波交变分量的振幅和角频率。由于光载波不可能是负值，所以载波的交变分量总是叠加在直流分量之上。被测信息可以对交变分量的振幅、频率或者初相位进行调制。

一般情况下，调制后的载波具有下列形式：

$$\Phi(t) = \Phi_0 + \Phi_m [V(t)] \sin \{ \omega [V(t)] t - \phi [V(t)] \}$$

式中， $V(t)$ 是由被测信息决定的调制函数。根据调制参量的不同，连续波调制可分为：

振幅调制 (AM)：调制参量为光通量振幅 $\Phi_m [V(t)]$ ；

频率调制 (FM)：调制参量为载波频率 $\omega [V(t)]$

相位调制 (PM)：调制参量为载波的初始相位 $\phi [V(t)]$ 。

1) 光信号的振幅调制

光载波信号的振幅瞬时值随调制信息成比例变化，其频率和相位保持不变的调制方法称振幅调制或调幅。此时，式中的 $\Phi_m [V(t)]$ 表示为

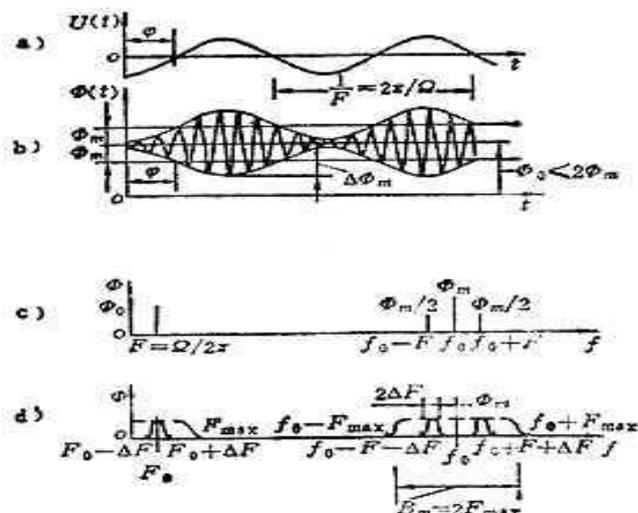
$$\Phi_m [V(t)] = [1 + mV(t)] \Phi_m$$

原式变成

$$\Phi(t) = \Phi_0 + [1 + mV(t)] \Phi_m \sin \omega t$$

式中， $V(t)$ 是调制函数，规定 $|V(t)| \leq 1$ 。 m 是调制度或调制深度，表示 $V(t)$ 对载波幅度的调变能力。并有

$$m = \Delta \Phi_m / \Phi_m = \text{被调制波的最大幅度变化} / \text{载波幅度} \leq 1$$



调幅波的波形和频谱

- a) 正弦调制函数 b) 对应的调幅波形
c) 调幅波频谱 d) 复杂调幅波频谱

下面分析调制函数为简单正弦变化情况下调幅波的形式及其频谱分布。此时，被传送信息按单一谐波规律变化（如图 a 所示），即

$$V(t) = \sin(\Omega t + \phi)$$

式中， $\Omega = 2\pi F$ 是被测信息的谐波变化角频率。 F 、 ϕ 是相应的频率和初相位。对应的载波信号表达式为

$$\Phi(t) = \Phi_0 + [1 + m \sin(\Omega t + \phi)] \Phi_m \sin \omega t$$

相应的信号波形表示在图 b 中。将上式展开，可以得到单一正弦调制函数的调幅波频谱：

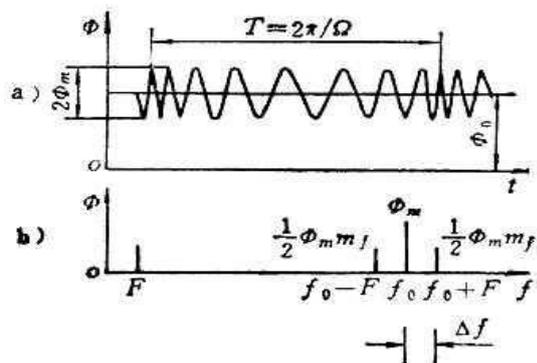
$$\Phi(t) = \Phi_0 + \Phi_m \sin \omega t + m \Phi_m \{ \cos[(\omega - \Omega)t - \phi] - \cos[(\omega + \Omega)t + \phi] \} / 2$$

相应的频谱图表示在图 c 中。由图中和上式可见，正弦调制函数的调幅信号除零频分量 Φ_0 外尚包含有三个谐波分量：以载波频率为中心频率的基频分量和振幅为基波振幅之半、频率分别为中心频率与调制频率之和频和差频的两个边频分量。和正弦调制函数本身的单一谱线相比，调幅波的频谱由低频移向高频，并且增加了两个边频。

可以证明，对于频谱分布在 $F_0 \pm \Delta F$ 范围的任意复杂函数 $V(t)$ ，其对应调幅波的频谱由以载波频率 f_0 为中心的一系列边频组成。这些边频是 $f_0 \pm F_1$ 、 $f_0 \pm F_2$ 、 \dots 、 $f_0 \pm \Delta F$ ，其中 F_1 、 F_2 、 \dots 是 ΔF 内的频谱分量。频谱图如图 d 所示。若调制信号具有连续的带宽 F_{max} ，则调幅波的频带是 $f_0 \pm F_{max}$ ，带宽为 $B_m = 2F_{max}$ ，其中 F_{max} 是调制函数的最高频率（如图中虚线所示）。

确定调制载波的频谱是选择检测通道带宽的依据。例如若载波频率为 $f_0 = 5\text{kHz}$ ，调制信号频率为 $f_s = 100\text{Hz}$ ，则调幅后的载波频谱分布在 $f_l = (5 - 0.1)\text{kHz} = 4.9\text{kHz}$ 和 $f_h = (5 + 0.1)\text{kHz} = 5.1\text{kHz}$ 之间，带宽为 0.2kHz 。使通道有选择性地滤波可减小噪声和干扰的影响，有利于提高信噪比。

2) 光信号的频率调制



调频信号波形和频谱

a) 调频信号波形 b) 相应频谱

光载波的频率按调制信号的幅度改变，偏离原有的载波频率，其瞬间偏离值与调制信号瞬时值成正比，载波的振幅保持不变，这种调制方法称频率调制或调频。此时，原式中的调制项可表示成

$$\omega[V(t)] = \omega_0 + \Delta\omega V(t)$$

式中， $V(t)$ 是调制函数，规定 $|V(t)| \leq 1$ 。

$\Delta f = \Delta\omega / 2\pi$ 是调制后载波频率相对于中心频率 f_0 的最大频率偏差，称作频偏。当调制函数 $|V(t)| = 1$ 时，载波频率的变化最大，为 $\omega_0 \pm \Delta\omega$ 。

这种情况下可得

$$\Phi(t) = \Phi_0 + \Phi_m \sin[\omega_0 t + \Delta\omega \int_0^t V(t) dt]$$

对于最简单的余弦调制函数的情况，有

$$V(t) = \cos(\Omega t + \phi)$$

式中， $\Omega = 2\pi F$ 是调制角频率。此时有

$$\Phi(t) = \Phi_0 + \Phi_m \sin[\omega_0 t + m_f \sin(\Omega t + \phi)]$$

式中， $m_f = \Delta\omega / \Omega = \Delta f / F$ 称作调制指数。 Δf 和 F 分别是频偏和调制频率。 m_f 表示了单位调制频率的变化引起的频偏变化的大小， m_f 值在设计时选定， $m_f > 1$ 时称宽带调频， $m_f \ll 1$ 时称窄带调频。调频信号的波形表示在图 a 中。将上式展开得

$$\Phi(t) = \Phi_0 + \Phi_m \{ \sin \omega_0 t \cos [m_f \sin(\Omega t + \phi)] + \cos \omega_0 t \sin [m_f \sin(\Omega t + \phi)] \}$$

在 $m_f \ll 1$ 窄带调频的简单情况下，可以认为

$$\cos [m_f \sin(\Omega t + \phi)] \approx 1$$

$$\sin [m_f \sin(\Omega t + \phi)] \approx m_f \sin(\Omega t + \phi)$$

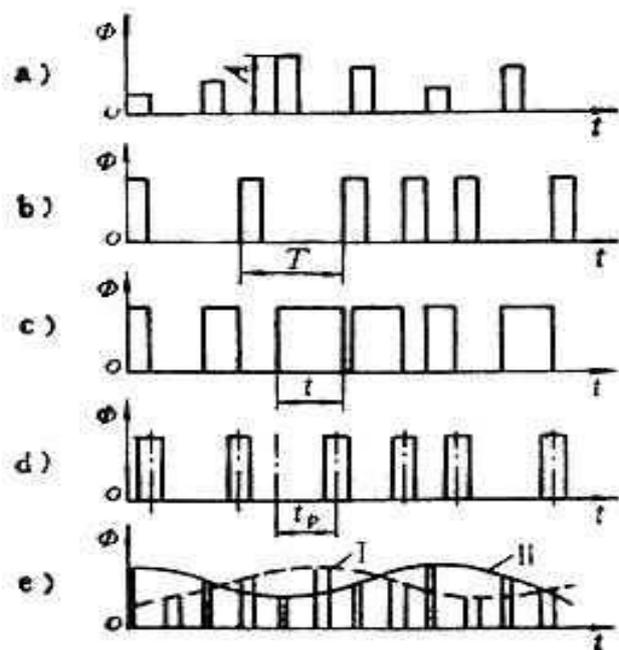
所以有

$$\begin{aligned} \Phi(t) &= \Phi_0 + \Phi_m [\sin \omega_0 t + m_f \cos \omega_0 t \sin(\Omega t + \phi)] \\ &= \Phi_0 + \Phi_m \sin \omega_0 t + \Phi_m m_f \{ \sin[(\omega_0 + \Omega)t + \phi] - \sin[(\omega_0 - \Omega)t - \phi] \} / 2 \end{aligned}$$

由式可知当频率调制指数 m_f 很小，载波频率变化范围 $\Delta\omega$ 不大时，按余弦规律调频的载波信号的频率和振幅调制一样，是由三个谐波分量组成的（如图 b）。其中包括振幅为 Φ_m 的载波频率 ω_0 和振幅为 $\Phi_m m_f / 2$ 的两个组合频率 $\omega_0 + \Omega$ 和 $\omega_0 - \Omega$ 分量。

在一般情况下调制信号形式复杂时，调频波的频谱是以载波频率为中心的一个带域，带宽随 m_f 而异。对于 $m_f \ll 1$ 的窄带调频，带宽 $B_f = 2F$ ；对于 $m_f > 1$ 的宽带调频，带宽 $B_f = 2(\Delta f + F) = 2(m_f + 1)F$ 。

2. 脉冲调制



脉冲调制的类型

- a) 脉幅调制 b) 频率调制 c) 脉宽调制
d) 相位调制 e) 复合调制

将直流光通量用诸如斩光盘的方法调制，使光通量间歇通断可以得到连续的光脉冲载波。若使载波脉冲的幅度、重复频率、脉宽、相位等参量或它的组合按调制信息改变，可以得到如图所示的各种类型的脉冲调制方式。下表列出了这些方法中各调制参数的变化。图中“—”表示脉冲参数保持不变，“0”表示脉冲参数随信息改变。

不同脉冲调制方式的参数变化

调制类型 \ 脉冲参数	幅度 A	周期 T	脉宽 t	相位 t _p
脉幅	0	—	—	—
频率	—	0	—	—
脉宽	—	—	0	—
相位	—	—	—	0
复合	0	—	0	—

调制方法不仅能提高光电系统的测量灵敏度，而且能在同一光学通路中实现多个信息的多路传输。上图 e 表示了两个同样采用脉幅调制的光信号，它们的脉宽不同。将此二信号放在同一信道中传送时，只要在接收端设置脉宽鉴别电路，利用光脉冲宽度不同的特点就能将它们分离到两个通路上。在各自的通路里从已调幅的脉冲序列中解调出信号包络线即可得到被传送的信息，这种方法已被应用于光纤多路通讯中。

可以实现单色光波或复合光波调制作用的装置或机构称作调制器。调制器包括机电调制器、辐射源调制器、光学调制器和光电子调制器等。

调制信号的解调

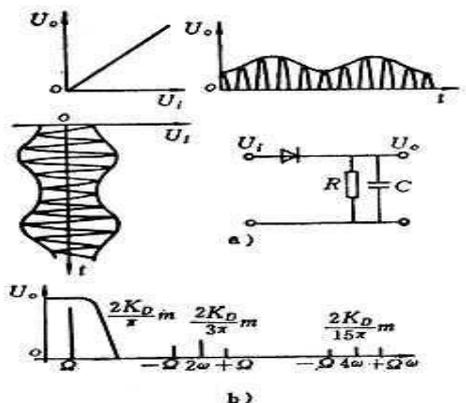
从已调制信号中分离提取出有用信息的过程称作解调或检波，实现解调作用的装置是解调器。解调或检波是信号调制的相反过程。

在时域分析中，调制是将有用信息及其时间变化载荷到载波的特征参量之上，而解调则是从这些调制了的特征参量上再现出有用信息。

从频域分析的角度，调制是将信号的频谱向以载波频率为中心频率的高频方向变换，而解调则是将变换了的频谱分布复原或反变换为初始的信号频谱分布。

不同的调制信号有不同的解调方法，我们来介绍和调幅波解调有关的二种方法，即：直线律检波和乘积检波。

1. 直线律检波和调幅信号的解调



检波二极管的伏安特性、基本检波电路和检波调幅解调信号的频谱图

a) 伏安特性和基本电路 b) 频谱图

解调是信号变换的非线性过程，需要利用非线性元件来实现。光载波的调幅信号通过光电变换后的隔直处理，通常具有包络线对称的双极性性质。

为得到单边包络线，再现调制信号，最适合于采用具有单向导电特性的二极管检波器。图 a 给出了检波二极管的伏安特性和基本电路。这是一种典型的单向导电特性。对于正极性信号幅度较大（1V 或更大）的情况，伏安特性可看作是通过原点的理想直线，其输出信号和输入信号成正比。工作于大幅度信号输入状态下的检波称作直线律检波。单向导电特性可表示为

$$U_o = K_D U_i \quad U_i > 0$$

$$0 \quad U_i \leq 0$$

式中， U_o 、 U_i 表示输出和输入信号， K_D 为比例因子。

假设按正弦规律调幅的光载波信号经光电变换及隔直处理后具有形式

$$U_i = (1 + m \sin \Omega t) \sin \omega t$$

代入 U_o 式，并做傅里叶级数展开，得

$$U_o = K_D (2/\pi) \left\{ (1 + m \sin \Omega t) - \sum_{n=1}^{\infty} \frac{1}{4n^2 - 1} [m \cos 2n \omega t + (m/2) \sin(2n \omega + \Omega) t - (m/2) \sin(2n \omega - \Omega) t] \right\}$$

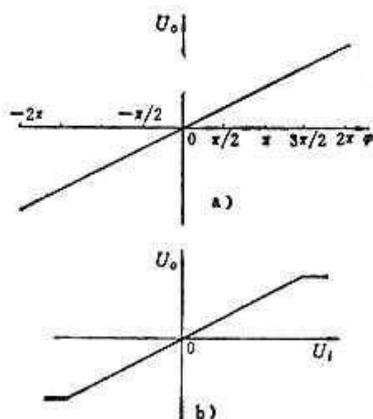
式中， n 是高次谐波的阶次。

由式可见，输出信号中除 { } 括号中的第二项是低频信号，是希望提取的调制信号之外，其余各项都是高频项，并且高频幅值逐次衰减。频谱图如图 b 所示。当 $\Omega \ll \omega$ 时，利用低通滤波器滤除高频分量即可得到有用的信号波形为

$$U_o = K_D (2/\pi) (1 + m \sin \Omega t)$$

检波器的输出信号和输入信号的包络线成正比例，实现了调幅波的解调。

2. 相敏检波和调相信号的解调



相敏检波的两工作特性

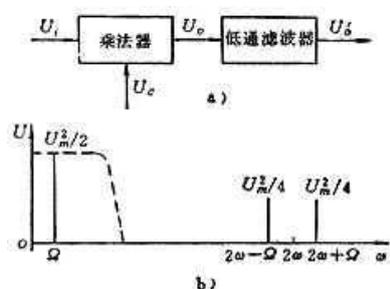
a) 鉴相整流 b) 极性检波

对于相位调制的载波信号，载波和参考信号间的相位差随被测信息改变。这种信号的解调采用相敏检波，这时检波器的输出电压应能反映出调制相位的变化。解调器的输出特性如图 a 所示。

在有些测量场合，常常不仅要求检测变量变化的方向而且希望确定变化的方向或极性。对这种有极性变量的调制，通常可用载波的幅度大小表示变量的数值而用载波信号和参考信号的相位差是同相或反相表示变量的极性，这称作极性检波。

显然为处理这种调幅信号也需要有对极性敏感的解调方法。这种解调器的输出特性具有图 b 的形式。

上述两种不仅能检测出调制信号的幅度，而且能确定载波相位数值的解调称作相敏检波或鉴相解调。它的基本原理是乘积检波。



乘积检波的原理框图和调幅波解调频谱

a) 原理框图 b) 调幅波解调频谱

乘积检波的原理框图表示于上图 a 中，它由解调器和低通滤波器串联而成。这里，解调器被看成是已调制信号 U_i 和参考信号 U_c 间的模拟乘法器，所以称作乘积检波。

设载波受单一频率余弦波调幅。调幅信号 U_i 为

$$U_i = U_m \cos \Omega t \cos \omega t$$

用作相位比较的参考信号 U_c 表示为

$$U_c = U_{cm} \cos(\omega t + \phi)$$

此处，二信号间的相位差 ϕ 将作为变量。为简单起见，设 $U_m = U_{cm}$ 。在解调器中，信号 U_i 和 U_c 相乘，输出信号 U_o 为

$$\begin{aligned} U_o &= U_i U_c \\ &= U_m^2 (\cos \Omega t \cos \omega t) [\cos(\omega t + \phi)] \\ &= (U_m^2/2) \cos \Omega t [\cos(2\omega t + \phi) + \cos \phi] \\ &= (U_m^2/2) \cos \phi \cos \Omega t + (U_m^2/4) \cos[(2\omega + \Omega)t + \phi] \\ &\quad + (U_m^2/4) \cos[(2\omega - \Omega)t + \phi] \end{aligned}$$

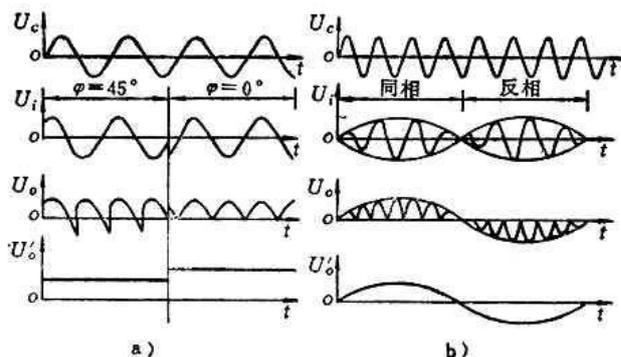
上式表明，乘法器的输出信号包括： $\cos \Omega t$ 、 $\cos(2\omega + \Omega)t$ 和 $\cos(2\omega - \Omega)t$ 三项。频谱分布如图 b。当 $\omega \gg \Omega$ 时，用低通滤波器可以滤除 U_o 中的高频项 $(2\omega \pm \Omega)$ ，得到相敏检波器的最终输出 U_o' 为

$$U_o' = (U_m^2/2) \cos \Omega t \cos \phi$$

该式表明：相敏检波器消除了高次谐波的影响，输出信号幅度与调幅信号的幅度成正比，因此能解调和再现出调制信号 $U_m \cos \Omega t$ ，相敏检波器的输出信号和载波与参考信号间的相位差 ϕ 有关。在载波信号幅度不变的条件下能单值地确定载波信号和参考信号间的相位差。

相敏检波的这些性质使之在调制信号的处理中得到广泛地应用。它具有下列的功能：

1) 对于幅度不变而单纯进行相位调制的情况，能解调出相位调制信息，得到如图 a 的输出特性。其解调信号的波形表示在下图 a 中。这种解调方法在诸如用调制盘检测光学目标方位角的测量中得到应用。



相敏检波的工作波形

a) 鉴相整流 b) 极性检波

2) 对有极性的调幅信号，载波信号的相位差只取 $\phi = 0$ 和 $\phi = 180^\circ$ 两种状态。这时有下列两种情况

$$\phi = 0, \quad U_o = (U_m^2/2) \cos \Omega t$$

$$\phi = 180^\circ, \quad U_o = (-U_m^2/2) \cos \Omega t$$

这表明，当载波与参考信号是同相位时输出信号为正值；反之为负值。解调信号的波形如图 b 所示。解调器的输出特性如图 7-13a 所示。这个性质被广泛应用于光电法几何位置偏差和运动方向的极性判断中。

9.3 简单光学目标的空间定位

下面介绍光强随空间分布的光信号的检测，先从简单光学目标的定位开始讨论。所谓光学目标就是不考虑被测对象的物理本质，只把它们看作是背景间有一定光学反差的几何形体或图形景物。

根据光强空间分布的复杂程度和测量目的，光学目标可以分成复杂图形景物和简单光学目标。

前者的图形分布结构细密、空间频率高、密度等级丰富，测量的目的在于确定图形的细节和层次、分析图形的内容等。

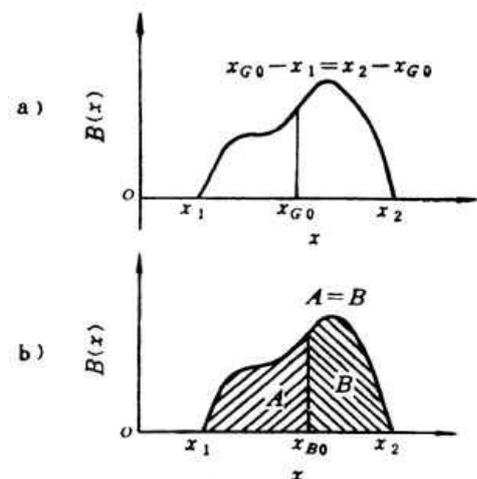
后者通常由点、线、平面等简单规则图形组成，信号处理的目的是确定目标相对基准坐标的角度或位置偏差，称作空间定位。

在实际的光学工程中，很多对象可以制作或简化成简单的光学目标。特别是工业图形大多是简单和规则的，如刻线、狭缝、十字线、光斑、方框窗口等。即使是复杂的物体，在确定它们相对于广阔背景的空间坐标时也可以将其看作是孤立的辐射光斑或通过稍许离焦的成像系统得到的弥散圆。因此，简单光学目标的空间定位是空间光电信号变换的基本内容。

光学目标的信息采集通常是通过光学成像系统完成的，根据辐射物体大小远近的不同分别可利用望远、照相、投影、显微等光路。这些光学系统的作用是将物空间的亮度分布转换为象空间的照度分布。因此光学目标形状位置的检测可归结为检测其象空间照度的分布及其随时间的变化。

在光学系统确定的情况下，对象面上象位置的分析代表了对物空间坐标的分析，它们之间由固定的光学变换常数来联系。

简单光学目标的位置检测大致分作几何中心与亮度中心两种检测方法。



光学目标的几何中心和亮度中心

a) 几何中心 b) 亮度中心

光学目标和其衬底间的光学反差构成了物体的外形轮廓。轮廓图形的中心位置称作它的几何中心，用 G_0 表示，如图 a 所示。几何中心的位置坐标 x_{G0} 可表示为

$$x_{G0} = (x_1 + x_2) / 2$$

式中， x_1 、 x_2 是物体边缘轮廓的坐标。通过测量和目标的轮廓分布相对应的象空间轮廓分布来确定物体中心位置的方法称作几何中心检测法。主要的处理方法有差分法、调制法、补偿法和跟踪法等。这些方法的主要依据是象分析。

光学目标的亮度分布是光辐射能量沿空间的分布。将物体按辐射总能量相等的标准分割为两部分，其中中心位置称作亮度中心，用 B_0 表示，如图 b 所示。亮度中心的位置 x_{B0} 满足关系

$$\int_0^{x_{B0}} B(x) dx = \int_{x_{B0}}^{\infty} B(x) dx$$

式中， $B(x)$ 是目标的亮度分布曲线。通过测量和目标物空间的亮度分布相对应的象空间照度分布来确定目标能量中心位置的方法称作亮度中心检测法。主要的处理方法有光学象分解和多象限检测等，这些方法的依据是象限分割。

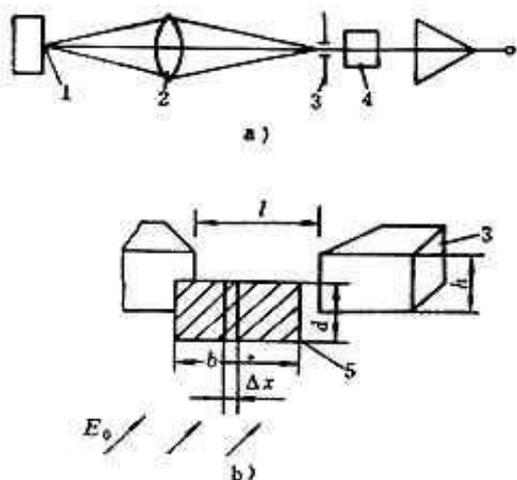
几何中心检测法

1. 象分析和几何参量调制

通过分析被测物体在象面上的几何中心相对于象面上某个选定基准的偏移情况，确定该物体在空间位置的方法称作象分析。能实现这种作用的装置称作象分析器。

象分析器的某本工作原理是将被测物体的光学象相对于象面基准的几何位置信息变换为通过该基准某一取样窗口的光通量，通过检测该光通量的变化来确定物体的坐标位置。因此象分析器是一种能将几何形位信息调制到载波光通量上形成光学信号的调制器，是几何量转换成光学量的 G/O 变换器。

1) 简单的象分析器及其定位特性

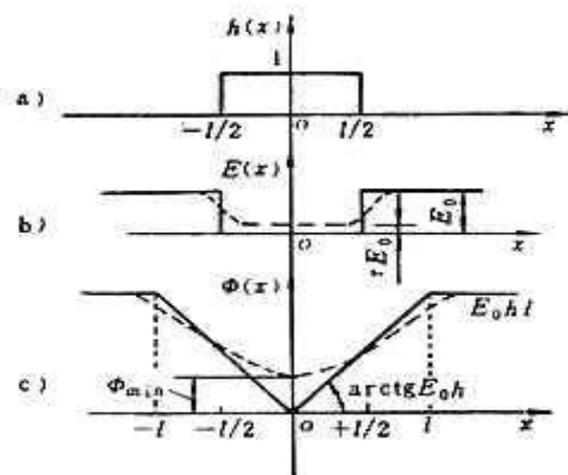


狭缝分析器的工作示意图

a) 装置示意图 b) 放大示意图

1-被测光学线纹 2-物镜 3-狭缝

4-光电接收器 5-线纹象



狭缝分析器象面上的照度分布和分析器定位特性

a) 狭缝窗口分布 b) 象面照度分布 c) 定位特性

我们用狭缝法测量光学线纹的对准和偏移状态的例子来说明一维象分析器的工作。这种测量一维坐标的方法在如同光电测微显微镜一类的光电装置中经常采用。图 a 表示了该测量装置的原理图。被测量的光学线纹 1 与其相对衬底的光学反差经物镜 2 成象于两边平行的透光狭缝 3 上，透过狭缝的光通量由光电接收器件 4 接收。

狭缝和线纹象组成的最简单的一维象分析器表示在图 b 的放大示意图中。图中 3 是狭缝，宽 j ，高 h ；5 是成象于狭缝面上的线纹象，宽 b ，高 d 。象面处于狭缝面上。狭缝的窗口分布 $h(x)$ 如图 a。象面上的照度分布是亮背景上的暗线如图 b。图中的实线表示理想照明和成象的情况。此时满足下列条件：象面照度 E_0 是空间均匀的；线象的透过率 $\tau = 0$ ，并且空间均匀一致；线象与狭缝等宽即 $l = b$ ；以及象高不小于缝高即 $d \geq h$ ，线象完全贯穿狭缝。由图可见满足理想成象条件的象面照度分布呈矩形的轮廓。在非理想的情况下象面照度分布如图中虚线所示。此时由于照明和光路散射、衍射、象差和离焦等影响，线象的透过率 $\tau \neq 0$ 并且是空间不均匀的。设最小照度 $E_{\min} = \tau E_0$ ，则线象与背景的照度差 ΔE 为

$$\Delta E = E_0 - E_{\min} = (1 - \tau) E_0 = D E_0$$

式中， $D = 1 - \tau$ 称作反差系数。

我们来计算狭缝分析器的定位特性（或传递特性），它表示透过狭缝的光通量 Φ 和线象中心线相对狭

缝中心线的相对偏移量 Δx 之间的关系。定位特性通常利用取样窗口的窗口函数 $h(x)$ 和目标象的照度函数 $E(x)$ 间的卷积积分求得，即通过窗口的光通量 $\Phi(x)$ 可表示为

$$\Phi(x) = h(x) * E(x) = \int_{-x}^{+x} h(\xi) E(x - \xi) d\xi$$

理想成像条件下狭缝分析器的定位特性如图 c 中的实线所示。它可以写成

$$\Phi = E_0 \cdot h \cdot |x| \quad |x| < 1$$

$$\Phi_{\max} = E_0 h l \quad |x| \geq 1$$

这是一组双向比例特性，相对纵轴对称分布。在缝宽 l 范围内是过原点的直线，斜率为 $E_0 h$ ；在缝宽范围外呈饱和状态。

对于非理想成像情况，由于象面照度的畸变，定位特性发生了较大的变化，如图 c 中的虚线。与理想情况下的曲线比较，其直线性破坏，有较大的零位输出并且曲线斜率变缓。相应的定位特性可以近似写成

$$\Phi = \Phi_{\min} = (1-D) E_0 l h \quad x = 0$$

$$DE_0 h |x| + \Phi_{\min} \quad |x| \leq 1$$

$$\Phi_{\max} = E_0 l h \quad |x| > 1$$

由上述的说明可以得到如下的结论：

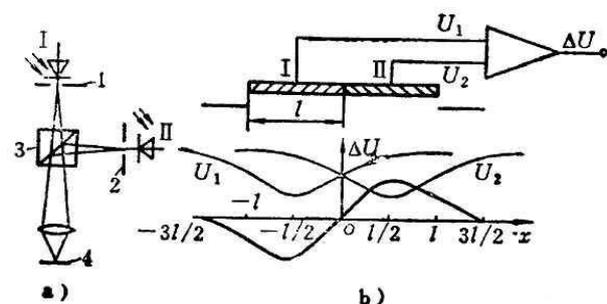
a) 组成简单象分析器的基本要求是，首先要形成被测目标的象。即在象面上形成有一定对比度的照度分布。其次，在象面上要设置测光窗口，透过窗口由光电元件接收目标象。核窗口可以是狭缝、劈尖的棱边或机械刀口等。窗口是定位的基准，应和光路的光轴保持确定的相对位置。测光窗口的尺寸和形状应与目标象保持严格的关系以便形成确定的定位特性。

b) 象分析器的定位特性给出了该装置输出光通量和输入目标象偏移量间的定量关系。通过检测光通量的极性和数值可以判断出目标的对准和偏移状态。也就是说利用象分析器将目标的位置信息调制到光通量的幅值变化中，因此象分析器本质上是调制器或位移——光通量的变换器。

c) 影响调制器定位特性的因素是很多的。这些因素的波动引起检测结果的示值扰动，破坏了系统工作的可靠性。

需要强调指出，对简易的单通道装置，定位特性存在许多缺点，例如图 c 中所示，在特性的零位处曲线很平坦，斜率较平缓、灵敏度不高；曲线有明显的非线性区域；对于两个方向的偏移该特性是对称的，因此没有极性反应来判断偏移方向。此外输出光通量是不随时间调制的，属于直流检测方式，由于这些不足之处，限制了简单象分析器的实际应用。

2) 双通道差分调制式象分析器



双通道差分调制式狭缝象分析器工作原理

a) 光路原理 b) 差分定位特性

1、2-狭缝 3-分束镜 4-被测线纹

这是一种常用的象分析器变换方式。这种方式采用了有两个接收光路的双通道结构。光路布置如图 a 所示。在象面上对称地放置两个狭缝 1 和 2，通过分束镜 3 得到的两个线象分别成象于狭缝面上。狭缝中心距等于线象宽 l 。两个狭缝的定位特性 $\Phi_1 = f_1(x)$ 和 $\Phi_2 = f_2(x)$ 具有类似的形状，表示在图 b 中。将两狭缝后的光电接收器件差分连接，则检测器的差分输出 ΔU 可利用图解法得到，如图 b 中粗线所示，这就是差分

型象分析器的定位特性。在线性区内，定位特性可表示为

$$\Delta U = U_1 - U_2 = E_0hx - (-E_0hx) = 2E_0hx \quad |x| \leq 1/2$$

可以看出，定位特性曲线有了明显的改进。首先具有了双极性的形式，可以根据差分光电流的极性判断偏移量 x 的方向。其次在 $x = \pm 1/2$ 范围内，曲线的线性情况改善，变化是单值的。此外曲线斜率 $d\Delta U/dx$ 提高了一倍，使得零位附近的分辨能力明显改善。为了进一步解决光通量调制问题，使系统变成交流测量系统，可采用光源调制方法使象面照度 E_0 按正弦形式变化，即

$$E_0 = E_m \sin \omega t$$

有

$$\Delta U = 2hE_m x \sin \omega t$$

假设 x 按正弦规律振动，即

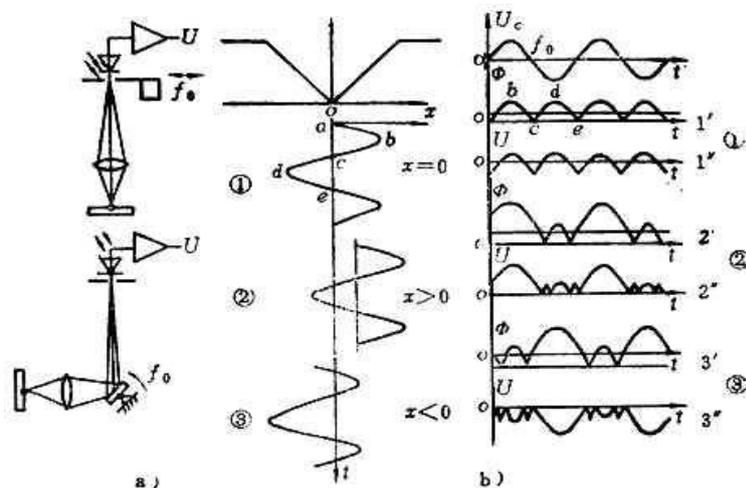
$$x = x_m \sin \Omega t$$

则

$$\Delta U = 2hE_m x_m \sin \Omega t \sin \omega t$$

此即为典型的带双极性的调幅波。利用前述的相敏检测方法可解调出被测的位置偏移 $x \sin \Omega t$ 。

3) 单通道扫描调制式象分析器



扫描调制型狭缝象分析器工作原理

a) 光路原理 b) 信号波形

将调制测量的技巧用于象分析中，可以得到扫描调制式象分析器。仍以狭缝分析器的例子来说明它的工作原理。在这类装置中（如图 a），成像光路增设了周期振动的光学零件（如棱镜、反射镜、光楔等）或者振动狭缝，使象面上的目标象相对狭缝周期振动。这时透过狭缝的光通量将随时间进行周期性的变化。这种使目标象和测光窗口之间有相对扫描运动的象分析器称作扫描调制型象分析器。

上图 b 给出了利用非线性图解法划出扫描调制型狭缝分析器的输出信号波形的过程。图中曲线①表示了当线象相对狭缝中心的偏移量为 $x=0$ 时，由于扫描运动形成的正弦运动轨迹（图中的 abcde）。经过理想成像条件下狭缝定位特性 $\phi = f(x)$ 的传递，得到如曲线 1' 的输出光通量时间波形。这个波形有正弦信号的全波整流形状，其基波分量是扫描频率的二倍频（如图中 1''），因此滤波后的输出直流信号 U 为零，这表明线象处于对准状态。曲线②和③分别表示当调 $x>0$ 和 $x<0$ 情况下调制光通量的变化情况。这种调制是幅度、宽度的复合调制，对应不同的偏移方向输出光通量有 180° 的相位变化。采用前述的极性检波方法可以得到 2'' 和 3'' 的整流输出电信号。其直流分量的极性和数值反映了线象的偏离方向和大小。调制信号的解调也可以采用测量脉冲宽度比或幅度比的方法进行。

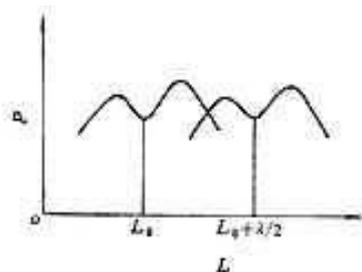
2. 扫描调制和极值检测

前述扫描调制型象分析器的主要工作条件是首先要使被测目标象相对于作为测量基准的测光窗口人为地进行周期性的往复扫描振动，产生交变的光载波；其次，扫描运动振动中心的偏置量作为被检测的位移信

息使载波信号受到调制，形成各种形式的调制载波。

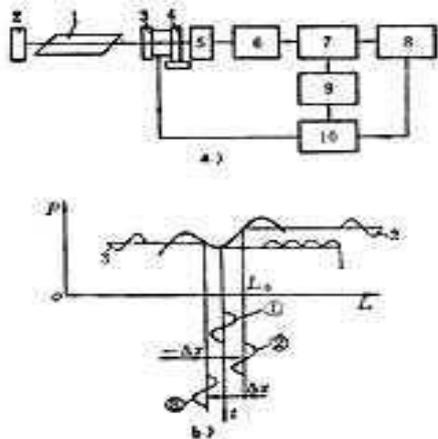
扫描调制是时间信号调制技术在位置测量中的移植和应用，具有许多优良的性质，例如交变的载波信号有利于信号的传递和处理，能改善信噪比、提高测量灵敏度和可靠性；目标象和测量基准间的相对扫描作用还能扩大测量范围等。由于这些理由，扫描调制测量在很多种几何量的形位测量和空间目标的扫描搜索中得到了成功的应用，具有普遍的意义。

在一个光学装置中，若其输出光通量值与其结构参数或几何形位间存在依赖关系，并事先能确定它们的定位特性，则通过检测输出光通量的数值就能确定这些参数的变化。若进一步利用极值特性的极值特点，通过扫描调制能判断装置对该极值的偏移，或通过反馈系统维持该装置处于极值状态上。这种方法称作光电装置的极值检测和极值控制。



激光器输出光功率与腔长的关系曲线

利用激光器谐振腔扫描调制来稳定激光频率的方法是光学极值控制的典型应用。氦氖激光器的输出光功率 P 与谐振腔长 L_0 之间的关系如图所示。当腔长增加 $\lambda/2$ 的整数倍时，输出功率周期性地重复极大极小值。对应每一个凹陷位置，激光器具有同样的振荡频率。因此只要控制腔长的数值使之维持在凹陷的底部即可以得到稳定的激光频率。



控制谐振腔的激光频率稳定系统

a) 频率系统示意 b) 波形图

1-激光器增益管 2-固定反射镜 3-活动反射镜

4-压电陶瓷 5-光电检测器 6-选频放大器

7-相敏检波器 8-直流放大器 9-参考电源 10-电压驱动器

上图 a 给出了控制谐振腔的稳频系统示意图。图中激光器由增益管 1，谐振腔固定反射镜 2 和活动反射镜 3 组成。反射镜 3 被固定在压电陶瓷 4 上，在交流电压驱动器 10 作用下，反射镜 3 沿光轴方向振动，周期性地改变腔长来实现扫描。扫描频率由参考电源 9 决定。若腔长扫描的平均位置恰为谐振频率 f_0 所对应的腔长 L_0 (如图 b 中的曲线①)，则输出光功率具有两倍于扫描频率的交变分量。此信号经过光电检测器 5 和选频放大器 6 后在相敏检波器 7 中和参考电压相比较。此时相敏检波器的输出为零，所以直流放大器 8 中没有直流电压控制驱动器 10 工作，因此腔长保持不变。不论由于何种外界原因使腔长改变时，腔长的扫描调制作用将在新的腔长位置上进行 (如图中曲线②或③)。此时载波信号具有和参考电压相同的频率分量，

因此相敏检波器的输出信号随偏离频率中心的偏差成比例增加，极性取决于偏离方向。极性和数值不同的误差电压经直流放大后通过驱动器给压电陶瓷以相应的偏移，使活动反射镜向维持腔长不变的方向移动。这样便实现了腔长的稳定，也就是激光频率的稳定。

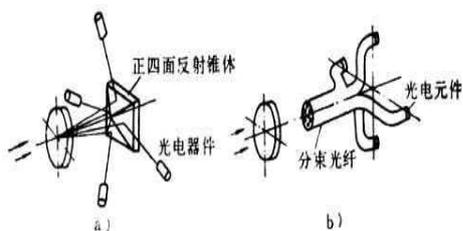
扫描调制的原理还应用于多种类型的扫描干涉仪和光学系统的自动调焦等许多光学技术中，已成为一种行之有效的光电技术方法。

亮度中心检测法

亮度中心检测的主要作法是将来自被测目标的光辐射通量相对系统的测量基准分解到不同的坐标象限上，根据它们在各象限上能量分布的比例可计算出目标的亮度中心位置，这种确定目标空间位置的方法称作象限分解法。

适用于该方法的目标可以是远处的宏观物体如星体、飞行物等，在经过成像系统的轻度离焦后可看作是弥散圆，更多的应用是用主动照明产生的标准规则图形。亮度中心检测法在空间目标的定向跟踪以及光学装置的准直对中和集成电路工艺设备的对准技术中得到应用。

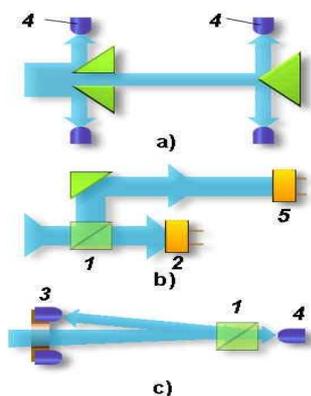
将目标象按坐标象限分解可以采用光学象分解和利用象限检测器的方法。这里简单说明光学象分解的基本方法。



实现四象限分解的光学零件

a) 正四面锥体 b) 四分束光纤

光学象分解是在光学系统中附加各种分光元件使入射光束分别向确定的不同方向传播，再在各自光路终端安装上具有单一光敏面的光电元件。典型的象分解方法包括：透射反射式、光纤分束式和全息分光式等。在平面坐标内实现四象限分解的光学零件表示在上图中。除了确定目标光点的平面坐标外，有些场合还要求确定光束的投射方向。下图给出了能同时实现沿平面坐标和绕 x、y 轴转动的多自由度分解的光学方法。



光学式象分解器

a) 中心孔式 b) 分光式 c) 反射式

1-半透半反镜 2、5-四象限光电池 3、4-光电元件

9.4 光学图象的扫描

扫描的基本原理和分类

1. 图象扫描

简单的光学目标具有规则的几何图形和只有黑白（或明暗）二值化的光强分布，因此能够采用比较简单的方法实现目标信息的采集和测量。在更普遍的情况下，光学目标是复杂的二维或三维的图形或景物，有更为复杂的光强分布形式，被测量对象的辐射强度 I 不仅与坐标位置 x 、 y 有关，而且可能随时间和辐射波长改变，属于这类光学图象的包括字符、图表、照片、产业制品的实体及自然景物和地形地貌等。

为了对复杂目标进行传真、录放、检测、处理、显示和存储，需要进行图象数据的采集和再现。也就是在光信号采集的情况下将物体在空间域的光强分布变换成时域内的电信号，或者在图形再现的情况下将时序电信号变成空间光强分布。

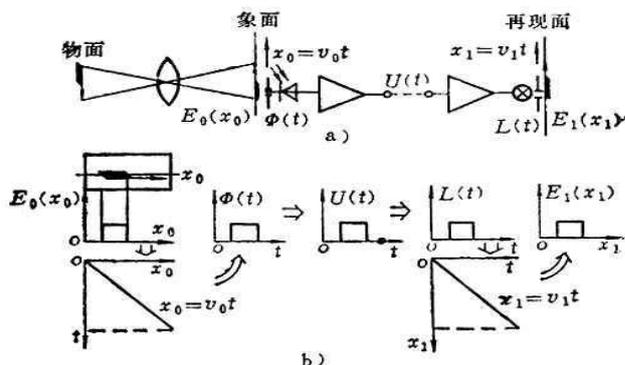
目前，在工程上能实现目标空时（或时空）变换和光电（或电光）变换的最常用技术是光电扫描技术。它具有在大视场范围内精确分辨图形细节的能力，在不同类型的图象仪器中几乎是不可缺少的。

在这种方法中只用了一个窄视场的光学系统和一个光电检测通道。当使该系统按一定时间顺序和轨迹串行逐点扫视目标物象空间的各点时，能获得瞬时值与被测目标的光学参数成比例的时序电信号。这个过程就称作扫描。

在相反的情况下，在扫描过程中用图象信号调制扫描发光点的瞬时发光值可以得到再现的光学图象。

这样，扫描方法能以单一窄视场的光电检测通道实现大范围的图象拾取和再现。它即具有宽广的观察测量范围，又具有较高的空间频率和灰度等级的分辨能力，这是它的重要特点。

2. 图象扫描的时空转换和光电转换性质



直线扫描原理示意图和信号变换

a) 原理示意图 b) 信号变换过程

我们用图给出的最简单的一维直线扫描原理图来定量说明图象扫描实现时空和光电双重变换的过程和关系，图中包括了图象采集和显示两个部分。设被测物体经光学系统在象面上得到的照度分布为

$$E_0 = E(x_0)$$

式中， x_0 为象面上的位置坐标。

在象面上设置带有取样窗口的光电接收器，组成的扫描单元沿象面与时间成正比地作直线运动。在扫描过程的运动方程为

$$x_0(t) = v_0 t$$

式中， t 是时间， v_0 是扫描运动速度。

上式表征了空间与时间的变换关系。 $1/v_0$ 是空时变换系数，它表示对应单位位移所需的扫描时间。取样窗口截取扫描点 x_0 处的象面光强，得到的光通量 ϕ 与坐标位置有关，但在扫描运动过程它是时间的函数，有

$$\phi(t) = SE[x_0(t)] = SE(v_0 t)$$

式中， S 为窗口截面积。

通过取样窗口的光通量经光电转换得到时序电信号，表示为

$$U(t) = K_{\phi} \phi(t) = K_{\phi} S E(v_0 t)$$

式中， K_{ϕ} 为光电转换系数。

这样，扫描单元实现了被测图象光强空间分布的空间-时间和光量-电量的变换。这个过程同时也是图象数据的采集过程。所得到的电信号与被测图象间有严格的对应关系，用上面的几个关系式表示。

为了实现图象的再现，将图象电信号经过传输线传到显示端。在这里，发光器件在电信号激励下产生随图象信号变化的发光亮度 $L(t)$ 。

$$L(t) = K_L U(t) = K_L K_{\phi} S E(v_0 t)$$

式中， K_L 是电光转换系数。当发光器件连同透光窗口作随时间成正比的扫描运动时，发光点的运动可表示为

$$x_1(t) = v_1 t$$

式中， v_1 是时空变换系数， x_1 是显示端的位置坐标，这时，在显示光屏上随着发光亮度 $L(t)$ 的变化在不同的位置上形成随位置变化的照度分布 $E_1(x_1)$ ，并有

$$E_1[x_1(t)] = K_E L(t) = K_E K_L K_{\phi} S E(v_0 t)$$

式中， K_E 为光度变换系数。

将上式中的变量 t 用位置坐标 x_1 代替可得

$$E_1(x_1) = K_E K_L K_{\phi} S E(x_1 \cdot v_0 / v_1) = K E(v x_1)$$

式中， $K = K_E K_L K_{\phi} S$ 是整个扫描变换系统对被测图形的灰度变换因子， $v = v_0 / v_1 = x_0 / x_1$ 是扫描系统的坐标变换因子。

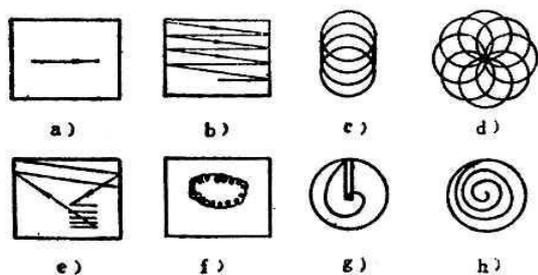
比较可知，采集端象面照度分布 $E_0(x_0)$ 和显示端光屏照度分布 $E_1(x_1)$ 之间具有相似的函数形式，二者之间仅相差灰度变换因子 K 和坐标变换因子 v 。

其中，灰度变换因子 K 取决于光电变换和电光变换系数等与图形灰色值有关的因素，而坐标变换因子 $v = x_0 / x_1$ 确定了扫描变换前后图形的尺寸比例。这些参数决定了扫描系统的工作品质。

为了保持扫描前后图形形状和光度值的比例关系，应该减少图形的几何失真和灰度失真，这就要求坐标变换因子和灰度变换因子是时间不变和空间不变的。此外，正确的扫描动作还要求图象采集端和显示端的扫描动作是严格同步的。

3. 扫描系统的分类和工作参数

1) 扫描系统的分类



几种扫描轨迹示意图

扫描系统有很多分类方法。按瞬时视场在空间的扫描轨迹可分为：直线扫描（图 a）、光栅扫描（图 b）、圆周扫描（图 c）、随机扫描（图 h、d、f）、螺线扫描（图 g）和分时割扫描（图 e）。按实现扫描的物理方法可分为：机械扫描、光学扫描、利用衍射光栅或全息光栅的扫描、利用电磁场控制的电子束偏转扫描、利用电光、磁光和声光效应的扫描、利用电荷耦合原理的移位电场扫描等。

2) 扫描装置的基本参数

扫描装置的基本参数包括扫描周期、扫描线密度和空间分辨率、扫描变换的直线性和动态范围、信道频带宽度、扫描数据率以及扫描精度等。这些参数描述了扫描装置对图象分解和合成时的质量标准，其中最重要的指标是扫描分辨率和扫描精度。前者表示装置分辨图象细节采集图形微细结构信息的能力，用单位长

度内可能分辨的最高黑白线对数表示，主要由扫描线的孔阑和信道通频带决定。

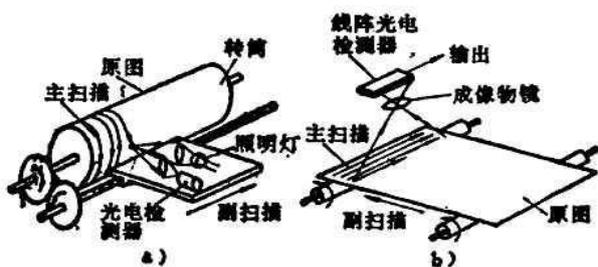
扫描精度表示时序电信号和空间光分布间光、电强度和时、空位置的正确对应关系。扫描轨迹的几何形状畸变和扫描运动速度的不均匀性会造成扫描非线性畸变，这将引起扫描的几何失真。扫描精度不论是对图象的不失真传送还是对图象测量都是一个重要的指标。

典型的扫描方法

1. 图形扫描

扫描技术的一个重要的应用方面是用作光学图片的分解与合成，称作图象扫描。它在现代的电视传真、复制、计算机图象输入设备、印刷分色制版、雕板等有关图象的生成、处理等应用中有很大的实用性。

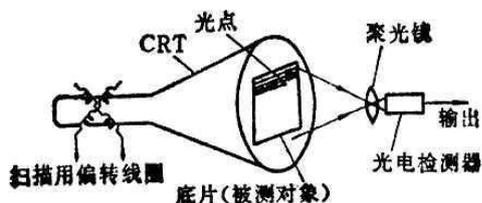
图象扫描的形式大致可分为：机械扫描、电子束扫描和光束扫描。下图表示了二种机械式扫描装置，它在传真及印刷制版中的应用由来已久。



转筒式和平板式机械扫描结构示意图

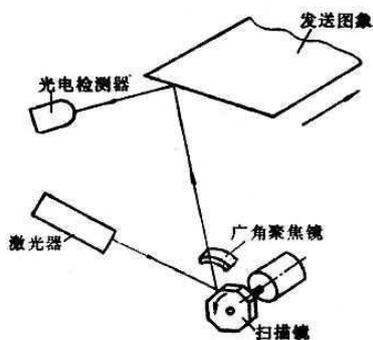
a) 转筒式 b) 平板式

在电真空器件中，电子束在交变电场或磁场作用下会改变它的运动轨迹，实现扫描作用。这在各种类型的电子摄像和显象器件中应用得很普遍，长期以来被作为一种主要的扫描手段。一种利用扫描电子束管的飞点扫描法，表示在下图中。它们用于高质量的传真和电视广播，主要的扫描器件飞点管实质上是 CRT 阴极射线显象管，在图中，飞点管荧光屏上由电子束形成的光点在偏转磁场作用下给出光栅型轨迹。被测半透明软片位于光屏上，扫描光点透过软片其强度被图片内容调制，经聚光镜由光电检测器接收变成图象信号。



飞点扫描装置示意图

使激光束扫描的装置称作光束扫描器或光学扫描器。目前，用得最多的是多面反射镜转鼓和检流计振镜。此外还有声光偏转器和全息光栅扫描器等。



机械转镜法扫描装置示意图

多面反射镜转鼓也称多面转镜，是一种最早应用的光学扫描器，上图给出了利用多面转镜的装置示意

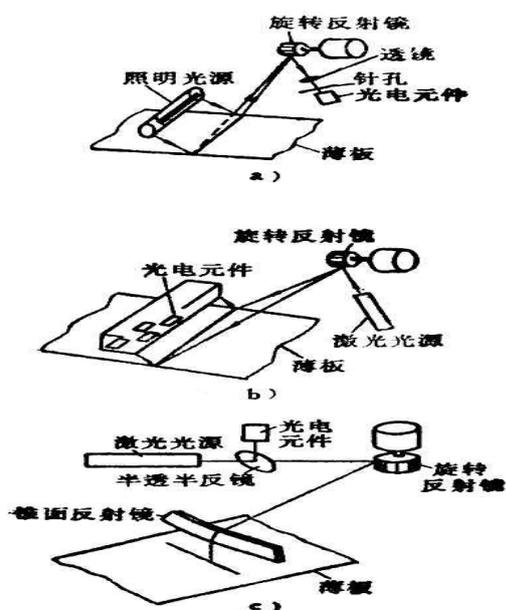
图。这种扫描的特点是扫描速度快（10000 行/s），扫描角度宽，精度很高，易获得高分辨率（40 线/mm）。缺点是结构复杂，成本高。

2. 实体扫描

在光学图象的检测中，除了各种平面图形外，更多的场合是要对实际的三维物体或景物进行扫描采集。例如各种工业制品的外观和尺寸检查、医疗检查以及景物和地形的测量等。

根据被测对象的状态不同，实体扫描大体可分为：对直线运动物体的扫描，对静止物体的扫描和对空间特定目标的搜索扫描等类型。

1) 直线运动物体的扫描



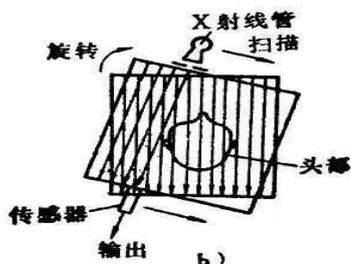
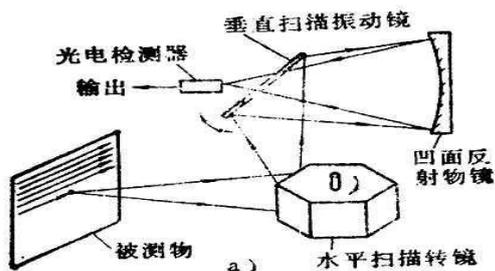
三种直线扫描检测的方式

a) 飞象扫描 b) 飞点扫描 c) 混合扫描

适用于生产车间传输带上的工件和条带状制品的表面状态或伤痕的自动检查。扫描测量装置应该完成一个方向的快速直线扫描，同时检测出经被测物体透射、反射的光强度，另一个运动方向由工件本身的缓慢传送实现。

根据照明和接收方式的不同这种扫描可以分为：光学飞象扫描方式（图 a）、光学飞点扫描方式（图 b）和光学混合扫描方式（图 c）。后者的照明和检测共用一个光路，因此结构紧凑，扫描精度高，是一种有应用前景的工业扫描装置。

2) 静止物体的扫描



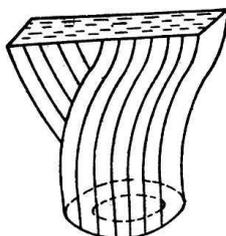
二种实体扫描装置示意图

a) 转镜-振镜扫描 b) 转动-直线扫描

对于单个静止物体，扫描检测应该具有二维或多维的扫描能力。

图 a 给出的振动转动反射镜的方法就具有这种功能，它是热成像照相机的基本形式。被测物体的温度辐射经水平扫描转镜和垂直扫描振动镜被连续进行光栅扫描，相当于二维的飞象方式。凹面反射物镜将反射光强通过振镜的中心孔会聚到光电检测器上，完成后续的信号处理。

图 b 中给出了用作 X 射线断层摄影的多维扫描方法示意图，能以很近的间隔从不同角度摄取人体各个断面的 X 射线透过强度分布。扫描装置相对人体转动，并在不同转动位置进行直线扫描，是典型的多维扫描的实例。



光纤圆直变换器示意图

在各种扫描装置中采用由光学纤维组成的直圆变换器或圆直变换器能比较简单地变换扫描轨迹。所谓光纤直圆变换器或圆直变换器是使图象的扫描轨迹由直行变换为环形或由环形变换为直行的光纤装置，为此，把光纤的一端集束成矩形截面，另一端则排列成环形，这样便可以将矩形的图象分布转换成环形的分布。它常和阵列式摄像器件组合使用，结构简单，功能多样，是一种有发展前途的新型扫描器件。

9.5 几何参量的光电检测

前面介绍了简单光学目标的调制变换和复杂光学目标的扫描采集。现在将进一步讨论几何参量的光电检测方法，这些方法主要应用于空间参量（几何量）或可变换为空间参量的其它物理量，如机械量、运动参量、物质成份等的光电检测中。

下表给出了基于几何光学原理检测几何参量的常用光电方法。

几何参量的光电检测方法分类表

下面将着重介绍光电测长和轴向测距二种基本技术。

光电测长

在光电检测的应用系列中一个重要的方面是测量物体的外形尺寸和相互间的距离。这种测量包括二个基本的环节，其一是通过瞄准确定被测物体的轮廓边缘，其二是计测边缘标志间的距离或长度。这两个环节可以是分开进行，也可以是同时完成的。

实时光电测长法大致可以分为：

1) 光度测量法

根据被测物体的遮光、反光、离焦、象偏移等引起的光量变化，采用单个光电元件测量物体尺寸的方法。这种方法装置比较简单，大范围的测量精度一般不超过 0.5%~5%，其中最重要的是瞄准测量法，具有较高的精度。

2) 成象测量法

利用物镜对物体成象，在象面上设置光敏阵列或摄像器件，在观察图象的同时确定轮廓的边缘和边缘间的距离。电视测量方法就是其中一例。这类方法的测量精度可高达 0.1%~2%。当物体沿景深方向移动时会造成放大倍率的变化，这会降低测量精度。因此这种方法需要解决自动调焦的问题。

3) 扫描测量法

利用扫描细光束周期性地照明被测物体，在物体边缘上形成强对比度的随时间周期性变化的光分布，通过测量通光遮光的时间差可以进行测长。例如各种形式的激光扫描测长装置采用的都是这种方法。这类装置结构相对复杂，有较高的测量精度，可达到 0.01%~0.1% 的相对精度，并且对于物面位置没有严格要求。

1. 瞄准测量方式

在这种测长方式中瞄准和测量的功能是用不同装置分别完成并同步进行的，也就是首先利用各种光电瞄准装置确定被测物体的轮廓，然后利用各种类型长度、角度标准器计量出被测边缘间的间隔尺寸。

瞄准是指将测量仪器的基准线和被测物体上的轮廓线准确重合对中的操作。它是影响几何量测量仪器测量精度的重要因素。光电瞄准的基本过程是首先将被测标志相对仪器基准的位置偏差转换为光电信号，然后根据电信号偏差值确定目标的对准程度、偏离大小和方向。

按照被测标志的不同形式，光电瞄准方法可以分为下列三种类型：

1) 线纹瞄准

这是一种最简单的光学目标定位，主要方法已在前面介绍。由于目标形状简单，采用一般的电子电路即可实现，并且可以得到 0.01~0.1 μm 的较高瞄准精度。

2) 轮廓瞄准

在这种方法中，瞄准系统象面上目标象的光强分布呈半边黑半边白的阶跃函数形式。不对称的轮廓形状增加了自动瞄准的难度，但其应用却比线纹瞄准更加广泛。例如在小孔瞄准，螺纹面形、轴承、刀具、齿轮等的形状测量，物位测量，热膨胀和电致伸缩微位移测量以及 IC 工艺过程的图形对准等都需要用到轮廓瞄准。最简单的轮廓瞄准可利用有矩形窗口的光电元件沿轮廓法线扫描来实现。

3) 图形瞄准

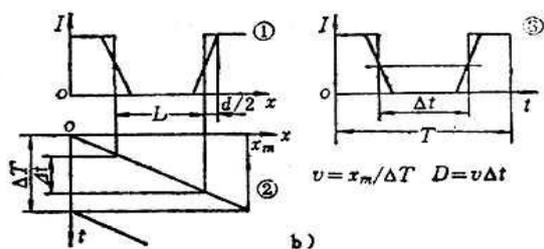
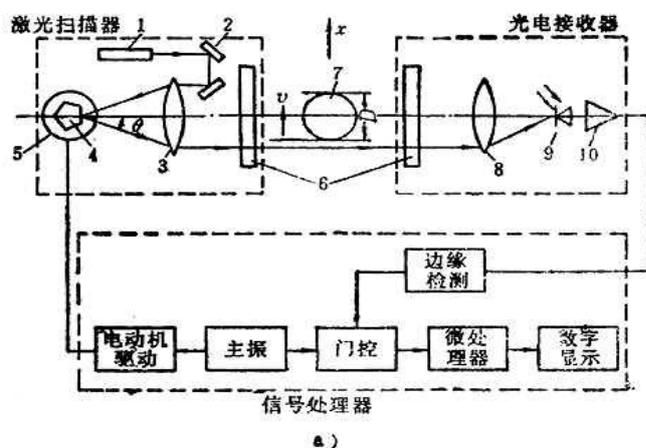
这是一种适用于复杂图形和实体的瞄准方法。它根据图形的局部特征成统计特点，利用数字化后的图象数据与计算机内存的标准数据比较，根据误差信号判断瞄准状态，是一种基于图象识别的原理发展起来的有广泛应用前景的瞄准方法。

可以作为测长基准的标准器有：刻线尺、计量光栅、光电编码器、磁尺、感应同步器、干涉仪等。在光波干涉比较仪中使用的动态光电显微镜和干涉测长装置的组合便是该种方式的典型应用。

2. 扫描测长方式

扫描是图象工程中进行图形采集、传送、显示、再现的基本方法，它的主要应用是图象的分解和合成。

但是，利用精确而稳定的同步扫描可以得到与被测物体光强空间分布成比例的电量时序变化。在这种光电信号中，时序间隔与空间坐标间的严格定量关系是图形几何量测量的重要依据，基于这种原理逐渐发展出一种图象测量的新技术。



激光扫描测径仪工作原理

a) 结构原理图 b) 信号波形

1-激光器 2-反射镜 3-f θ 透镜 4-多面反射镜 5-同步电动机

6-滤光片 7-被测物 8-聚光镜 9-光电检测器 10-放大器

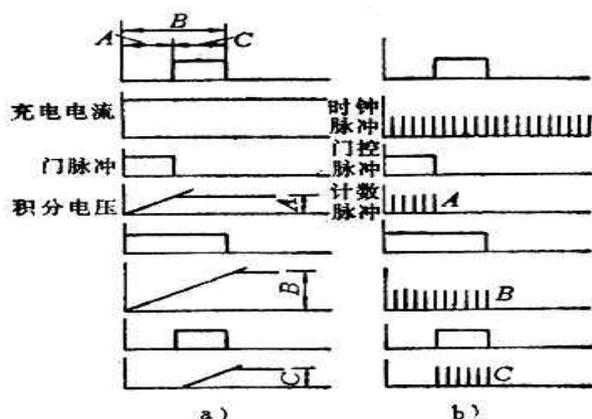
作为扫描测长的典型实例，下面讨论激光扫描测径仪的工作原理。上图给出了利用激光束扫描不透明物体来测量圆柱物体直径的简单示意图。扫描测量系统由三个部分组成（图 a）：将激光聚成细束平行投射并对被测物体进行扫描的激光扫描器，放置在被测物体后接收扫描激光束的光电接收器和处理接收信号的信号处理器。将直径为 D 的被测圆柱体放置在扫描器和光电接收器之间，若用细束平行光照明，则在一维截面上通过被测物体的光强分布如图 b 中所示，随着扫描器产生的细光束以速度 v 横向平行扫描物体时，通过物体的光信号被聚光镜收集并由光电检测器转换为电信号。这个光电信号呈脉冲系列形式，其形成过程如图 b②和③的波形图所示。其中②为激光束的扫描运动图，③为扫描物体后的光电时序信号图。由于扫描光点直径 d 的影响，脉冲前后沿有一过渡区域，在近似的情况下半峰值点所对应的脉冲宽度 Δt 代表了物体的直径，根据扫描变换公式有

$$D = v \Delta t$$

由式可知，由于其中激光束扫描运动速度 v 是恒定已知，只要精确地测量到 Δt ，便可以确定被测物体的直径 D 。

激光束的扫描可以用多面反射转镜或振动音叉等方法实现。为保证测量的准确性，要求在整个测量时间内，在扫描空间的各个位置上，扫描速度 v 的数值和方向应保持恒定不变。这需要采取同步检测和 $f \theta$ 镜等技术措施。此外，准确地检测到物体的边缘轮廓以及测量图象信号的脉冲宽度，作为共性技术也是很重要的。

成象系统的光学传递函数和扫描线孔阑效应使检测系统的频率特性变坏，造成图象信号脉冲前后沿的过渡区，理想的矩形图象信号变成了钟形的波形。因此，为了准确地确定物体的轮廓，需要进行边缘检测。常用的方法包括阈值比较法、微分阈值法和二次微分法。



二值化信号脉宽的测量方法

a) 恒流积分 b) 时钟脉冲计数

为了测量二值化图象信号的脉冲宽度，常采用恒流积分法和时钟脉冲计数法。上图表示了二种脉宽测量方法的工作原理。在图 a 的恒流积分法中，采用门脉冲来控制恒流源对积分电容的充电时间。门脉冲以扫描同步脉冲为时间起点，以被测图形二值化信号的升降沿为终点。充电电容上的积分电压值即可作为图形边缘位置的度量。用类似的模拟积分法可以得到表征图形宽度的积分电压。在图 b 的时钟脉冲计数法中，代替恒流充电电流，采用了时钟脉冲源，在图形二值化信号周期内进行时钟脉冲计数。累计的脉冲值即能以一定的比例确定被测物体的边缘位置和宽度。

轴向测距

在以前的讨论中所涉及的是一维或二维图形的尺寸测量。在更普遍情况下需要解决纵向深度的测量也就是测量沿检测系统光轴方向的位移或距离，通常也称作光电距离检测。这种方法在复杂工件的三维形状检测、仿形测量和光学装置的自动调焦以及机器人视觉和激光雷达等方面应用得很普遍。纵向距离测量的方法大致可分作下列基本形式：

1) 光波测距法

包括相位测距和脉冲测距。基本原理如前所述。这种方法适合于远距和超远距的测量，例如军用测距仪测距范围为 100~10000m，精度 $\pm 5m$ ，人造卫星测距仪测距范围几百公里到几万公里，相对精度达 10^{-6} 以上。

2) 干涉测距法

包括单频或双频干涉仪测距。基本原理将在后面讨论。这种方法适合在实验室环境下使用，有较高的环境要求。有效测距范围在 100m 以内，有 10^{-8} 以上的测距精度。

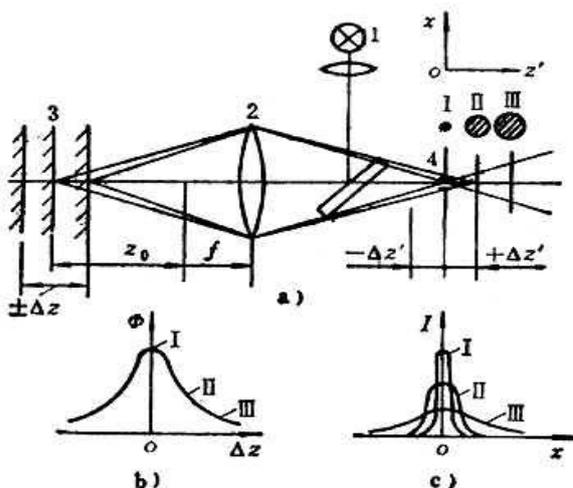
3) 离焦检测法

这是一种利用光学成像系统的有限焦深范围，当物体轴向偏移时引起的离焦现象进行距离检测的方法，这种方法测距范围较短在 10m 以内到几微米的范围。测距分辨率可达到测距范围的 1/100~5/1000。

这里介绍二种利用离焦检测的测距方法。

1. 象点轴上偏移光焦点法

在成像光学系统中，物象之间有严格的几何关系，被测物体离开理想物面时会引起象面上光照度分布的变化，这种现象可用来测量轴向位移。光焦点法利用点光源对被测物体表面照明，使用成像物镜对该光点成像聚焦。当物体沿光轴方向位移时，象面焦点扩散形成弥散圆。只有准确处于物面位置时，良好的成像状态才能保证象面上有集中的光密度分布。这种以聚焦光斑光密度分布的集中程度来判断物体轴向位移的方法称作光焦点法。



光焦点法测距原理图

a) 原理示意 b) 轴向定位特性 c) 象面光强分布

1-点光源 2-成象物镜 3-被测物 4-针孔光阑

上图表示了光焦点法的工作原理图。图 a 中，点光源 1 通过成象物镜 2 在被测物 3 表面成点象，该象点作为新的发光物点，折回成象物镜的光路中，在象面上成清楚象，象面的光照度分布呈衍射斑的形式（表示在图 c 中的曲线 I）。当被测物表面相对理想物面前后偏移 $\pm \Delta z$ 时，象点相对理想象面同方向地前后移动 $\pm \Delta z'$ 。 Δz 和 $\Delta z'$ 之间按几何光学规律有关系

$$\Delta z' = \beta^2 \Delta z$$

式中， β 为成象物镜的横向放大率，设物镜焦距为 f ，物距为 z_0 ，则 β 值可表示为

$$\beta = f/z_0$$

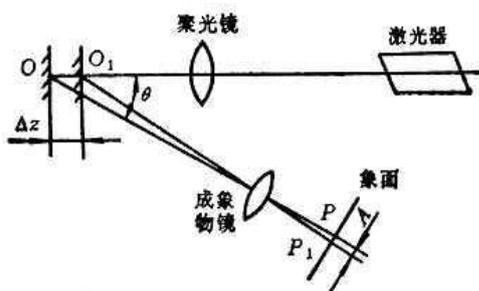
象点的 $\Delta z'$ 偏移，使象点照度分布扩散，如图 a 和 c 中的 II 和 III。

现在，在初始象面位置上设置针孔光阑 4，使其直径小于光斑直径。这时，前后移动光阑位置 $\Delta z'$ ，通过针孔的光通量 Φ 将随 $\Delta z'$ 而改变。若将光斑和针孔光阑看作是一象分析器，则它们的轴向定位特性表示于图 b 中。它表明了通过针孔光阑的光通量 Φ 和物点轴向偏移量 Δz 之间的关系。轴向定位特性相对于初始物面位置 z_0 呈有极值的对称分布。曲线范围由物镜像深决定，超过焦深后能量密度急剧下降。

利用所获得的轴向定位特性，根据前面关于象分析器的说明，可以组成各种形式的象面离焦检测系统，例如扫描调制检测、双通道差分检测等。读者可借此自行设计一些方案。

2. 象点轴外偏移的象偏移法

象偏移法又称光切法，它是一种利用三角测量方式的轴向位移测量方法。当将光束照射到被测物体时，用成象物镜从另外的角度对物体上的光点位置成象，通过三角测量关系可以计算出物面的轴向偏移大小。这种方法在数毫米到数米的距离范围可实现高精度的测量。在工业领域内的离面位移检测中常常用到。



象偏移法测距原理图

上图给出象偏移法测量轴向位移的原理示意图。照明激光源经聚光镜成象于被测物面上，反射亮点由成象物镜接收，在象面上成点象。设被测表面轴向偏移 Δz 时，照明光点由 O 移到 O_1 ，相应象点由 P 偏移到

P_1 ，则象点横向偏移量 A 可表示为

$$A = |PP_1| = \Delta z \beta \sin \theta$$

式中， β 为横向放大率。

上式可改写为

$$\Delta z = A / (\beta \sin \theta) \Big|_{\theta \rightarrow 0} = A / \beta \theta$$

上式表明，在已知 β 和 θ 的情况下只要测量出象点的横向偏移量 A ，即可计算出待测的物面轴向偏移 Δz 。

象点横向偏移量 A 的检测可以采用 CCD 摄像器件或者半导体位置检测器 (PSD)。前者的空间分辨率取决于固定分割的光敏象素尺寸，比 PSD 稍为优越。它的温度漂移较小，工作稳定。但是 CCD 装置需要扫描驱动电路，背景光的调制作用会引入干扰信号。此外，它的测量响应速度由扫描速度决定，高速应用受到限制。

采用 PSD 的检测方法，信号是模拟输出，分辨率受入射光功率的影响。只有光功率达到一定值时才能有和 CCD 装置相接近的分辨率。PSD 的响应速度高，可进行光点位置的连续检出，调制光也可检测，因此，容易和背景光相分离，此外，它的信号处理电路简单。至于照明光源可以选用白炽灯、He-Ne 激光器、半导体激光器和发光二极管等。

第十章 相干光电信号变换

各种干涉现象都是以光波为基础的信号变换过程。作为实际的载体，光波载荷了被测信息，它的特征参量与被处理的长度、距离、角度、面形、微位移、运动方向和速度、传输介质物理属性等信息存在着严格的内在联系，表现出随时间和空间变化的外观特性。

利用光电方法对光波的各种干涉现象进行检测和处理，最终解算出被测几何参量和物理参量的技术统称作光电干涉测量技术。随着现代光学和光电子技术的发展，光电干涉技术以其潜在的生命力在信息科学中崭露头角，取得了长足的进展。本章将介绍几种物理变换的光电方法，着重讨论相干光信息的调制和检测技术。

本章内容包括：[光学干涉和相干光信息分类](#)、[非相干探测和相干探测](#)、[相干光的相位调制和检测](#)、[相干光的频率调制和检测](#)。

10.1 光学干涉和相干光信息分类

1. 光学干涉和干涉测量

利用相干光作信息传输或检测，需要将被测信息载荷到光载波上，使光载波的特征参量随被测信息变

化。

但是由于光波波动频率很高，到目前为止各类光探测器尚不能直接感知光波本身的振幅、相位、频率、偏振的变化。所以在大多数情况下要利用光的干涉现象将这些特征参量转化为光强度的变化或转化为光探测器能敏感的较低频率的光电载波信号。因此各类型的光干涉现象是利用光波传输信息的基础。

所谓光干涉是指可能相干的二束或多束光波相重叠，它们的合成光波随其间的相位关系表现出不同的光强度空间分布或时序变化的现象。干涉测量的基本作用在于把光波的相位关系或频率状态以及它们随时间的变化以光强度空间分布或光强度随时间变化的形式检测出来，这个作用有时称作相幅变换。

以双光束干涉为例，设二相干波面波振幅 $U_1(x, y)$ 和 $U_2(x, y)$ 分别为

$$U_1(x, y) = a_1 \exp\{-j[\omega_1 t + \phi_1(x, y)]\}$$

$$U_2(x, y) = a_2 \exp\{-j[\omega_2 t + \phi_2(x, y)]\}$$

式中， a_1 、 a_2 为光波的振幅， ω_1 、 ω_2 为角频率， ϕ_1 、 ϕ_2 为初始相位。

当它们合成时，所形成干涉条纹的强度分布 $I(x, y)$ 可表示成

$$I(x, y) = a_1^2 + a_2^2 + 2a_1 a_2 \cos[\Delta \omega t + \phi(x, y)]$$

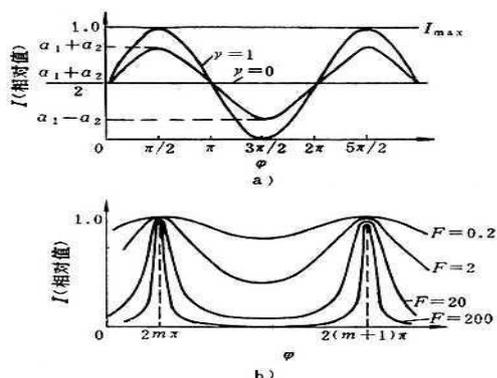
$$= A(x, y) \{1 + \gamma(x, y) \cos[\Delta \omega t + \phi(x, y)]\}$$

式中， $A(x, y) = a_1^2 + a_2^2$ 是条纹的光强直流分量， $\gamma(x, y) = 2a_1 a_2 / (a_1^2 + a_2^2)$ 是条纹的对比度， $\Delta \omega = \omega_1 - \omega_2$ 是光频差， $\phi(x, y) = \phi_1(x, y) - \phi_2(x, y)$ 是相位差。

当二光束光频相同即单频光相干时，有 $\omega_1 = \omega_2$ 或 $\Delta \omega = 0$ ，此时干涉条纹不随时间改变，呈稳定的空间分布，上式变成

$$I(x, y) = A(x, y) \{1 + \gamma(x, y) \cos[\phi(x, y)]\}$$

这表明单一频率双光束干涉时，随着相位差的变化，干涉条纹强度的分布表现为有偏置的正弦分布如下图 a 所示。以此为基础的干涉测量形成了干涉条纹检测技术。图 b 为多光束干涉时的条纹光强分布。



单频光束干涉条纹的强度分布

a) 双光束干涉 b) 多光束干涉

当两光束频率不同时，由于光频率约 10^{15} Hz，而目前光电检测器件的频率响应不超过 10^{10} Hz，因此对频率相差较大的二束相干光，将观察不到干涉的交流分量，但频差较小的二束双频光相干，检测器可检测到干涉条纹以 $\Delta \omega$ 的角频率在波动，形成了光学拍频信号。这就是外差干涉现象。

采用电信号处理器可以直接测量光拍信号的各种参量，从而能以极高的灵敏度测量出相干光束本身的特征参量，形成了新型的外差检测技术。

条纹干涉强度取决于相干光的相位差，而后者又取决于光传输介质的折射率 n 对光的传播距离 ds 的线积分，即

$$\phi = (2\pi / \lambda_0) \int_0^L n ds$$

式中， λ_0 为真空中光波波长， L 为光程。对均匀介质，上式可简化为

$$\phi = 2\pi nL / \lambda_0$$

对上式中的变量 L 和 n 作全微分可得到相位变化量 $\Delta \phi$

$$\Delta \phi = (2\pi / \lambda_0) (L \Delta n + n \Delta L)$$

这表明光传播介质折射率和光程长度的变化都将导致相干光相位的变化，从而引起干涉强度的改变。

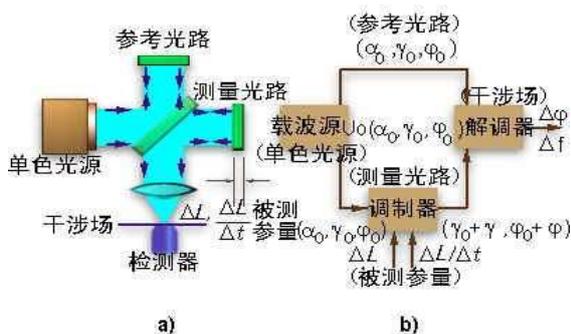
这一性质被用于改变光载波的特征参量，以形成各种光学信息。实际上有许多参量能造成光程差的改变，例如几何距离、位移、速度、温度引起的热膨胀等会引起传播距离的改变；介质的成份、密度、影响密度的温度、压力以及电场、磁场、介质的电光和磁光效应、应力引起的光弹效应等都能引起折射率的变化。进一步，从物体表面反射光波的波面分布可以确定物体的形状。

因此，从这个意义上讲光学干涉技术是一种能检测非光学参量的传感技术。能形成干涉现象的装置是干涉仪，它的主要机能是将光束分割成两个不同的支路，由外屏引入光程差后再重新合成借以观测干涉现象。

2. 干涉测量中的调制和解调

从信息处理的角度来看，干涉测量实质上是被测信息对光频载波的调制和解调的过程。各种类型的干涉仪或干涉装置是光频载波的调制器和解调器。所谓光调制是将一个携带信息的信号叠加到光载波上。能完成这一调制作用的装置称作光调制器。

光调制器能使光载波的特征参量如振幅、相位、频率、偏振、波谱分布等随被测信号的变化而改变。相对应的光调制技术分为光振幅调制 (AM)、相位调制 (PM)、频率调制 (FM)、偏振调制 (POM) 和光波谱调制 (SM)。解调是调制的相反过程，它能从被调制的光载波中以与被测量成比例的光强信号或电信号形式检测出被测参量。解调器可以是光学的、电子的或光电混合的。



干涉仪是光学调制器和解调器的组合

a) 原理示意图 b) 等效方框图

下面用最常见的迈克尔逊干涉仪来说明干涉仪是光学调制器和解调器的组合。上图给出了它的结构配置 (图 a) 和等效方框图 (图 b)。就信息传递的实质而言，实际的干涉仪结构和工作过程可以用下列方式描述：干涉仪中的单色光源是相干光载波的信号发生器，它产生振幅为 a_0 ，频率为 ν_0 ，初相位为 ϕ_0 的光载波信号，用 $U_0(a_0, \nu_0, \phi_0)$ 表示，载波信号分两路引入干涉仪。

在参考光路中光载波作为基准保持其原有的参量。在测量光路中， $U_0(a_0, \nu_0, \phi_0)$ 受到被测信号的调制。

如果被测信号是位移 $\delta(x)$ ，则引起光频载波的相位变化 $\Delta \phi$ ，称作相位调制，形成 $U_s(a_0, \nu_0, \phi_0 \pm \Delta \phi)$ 的调相信号。

若被测信号是运动速度 $v(t)$ ，则引起光频载波的频率偏移 $\Delta \nu$ ，称作频率调制，产生 $U_s(a_0, \nu_0 \pm \Delta \nu, \phi_0)$ 的调频信号。

这里测量光路连同其中的变送器起到信号调制器的作用。已调制的光频波在干涉面上和来自参考光路的参考光波重新合成，形成具有稳定干涉图样 (在测位移情况下) 或确定光拍频率 (在测速情况下) 的输出信号。这个信号解调了光频载波，以干涉条纹的相位分布或光拍的时序变化表征出被测量的特征，因此这被看作是光学解调的过程。

从光频载波的调制和解调的角度建立干涉测量的模型有助于我们从变化繁多的干涉现象和实用装置中提取出信息传输的线索，理解物理变换过程的本质。

3. 相干光信息分类

光学干涉形成了各种各样的光学现象，这种以光波的时空相干性为基础，在被测量（传输或处理）信息的调制下，光载波所表现出的各种形式的光强度空间分布或时间变化统称为相干光信息。它的形成和检测过程实质上是光载波受信息调制和已调制光被解调再现为原信息的过程。

相干光信息可以按多种方式分类。

1) 按物理效应分类为：双光束和多光束干涉条纹和干涉图、衍射图和莫尔拓扑图、散射图、全息照相和全息干涉及计算全息图、散斑照相和散斑干涉图、傅里叶变换、光学外差或零差效应。

2) 按被测目标的表面性质分为：带协作镜面靶和粗糙表面被测目标。

3) 按相干光的频率分为：单一频率双光束干涉、单一频率多光束干涉、双频率双光束干涉，多频率多光束等。

4) 按相干场的时空范围分为：相干场限制在局部空间随时间变化的干涉、在二维平面内发生的空间干涉（干涉图）、在二维平面内发生的随时间变化的干涉。

5) 按调制解调的类型分为：相位调制（PM）、频率调制（FM）、振幅调制（AM）、偏振调制（POM）、光波谱调制（SM）。

下表给出了相干光信号的分类，分别列出了它们的载波性质、调制方式、干涉图样、光电检测方法以及典型应用。从表中可对许多干涉现象及其间的联系有一个较全面的了解，并能揭示当前干涉测量的发展趋势。表中所列的多数应用将在本章中以相位调制和频率调制为线索逐项加以说明。

相干光学信息的类型

时空类型	光载波	调制方式	光学图样	检测方式	典型应用
一维时间调制	单频光	PM	干涉条纹	条纹计数测相	迈克尔逊干涉仪
		FM	干涉条纹扫描	条纹频率测量	傅里叶光谱仪
	双频光	FM, PM	外差型光拍频	外差测频	光通讯
		FM	零差型光拍频	外差测频	多普勒速度计
		FM	互差型光拍频	条纹测频	萨纳克转速计
		PM	外差型光拍频	拍频测相	波前检测、双频干涉仪
DFM	外差型光拍频	拍频测频	外差分光测量、调频波测距		
二维空间调制	单频光	SPM	散斑图	扫描光度测量	散斑干涉仪
		SPM	全息图	数字图样处理	全息干涉仪
		SPM+PM	干涉条纹扫描	条纹扫描或锁相跟踪	扫描干涉仪、锁相干涉仪
	双频光	SFM	平面外差拍频图	扫描光拍检测	平面外差干涉仪
		SDFM	平面外差拍频图	扫描光拍检测	调频波干涉仪

注：PM——时间相位调制；FM——时间频率调制；DFM——直接频率调制；SPM——空间相位调制；SFM——空间频率调制；SDFM——空间直接频率调制。

10.2 非相干探测和相干探测

为了用光电方法检测和解调出光载波所载荷的信息，通常采用两种探测方式，即非相干探测和相干探测。

非相干探测指的是当将信号光波直接投射到光电检测器件的光敏面时，检测器只响应入射于其上的平

均光功率，无论是对相干光或非相干光，检测系统只能检测出由光强度调制所形成的信号，这种方式也称直接探测。与此相对应的，相干探测能响应相干光波的波动性质，检测输出的电信号能间接表征光波的振幅、频率、相位或与之有关的信息。它的主要方式是外差探测。

1. 非相干探测的原理和特性

设被测光信号的波振幅 $U_s(t)$ 为

$$U_s(t) = a_s \sin(\omega_s t + \phi_s)$$

式中， a_s 、 ω_s 和 ϕ_s 分别为信号光波的振幅、角频率和相位。

它产生的入射功率 P_s 为

$$P_s = a_s^2 [\sin^2(\omega_s t + \phi_s)]_{\text{平均}} = a_s^2 / 2$$

光电检测器件只响应平均功率，其输出信号 I_{ds} 为

$$I_{ds} = P_s \cdot \eta q / h \nu = K P_s = K a_s^2 / 2$$

式中， $K = \eta q / h \nu$ 为光电灵敏度。

上式表明在直接探测情况下，光电检测器件的响应与平均入射功率成正比或与波振幅平方成正比。它不响应载波光的频率和相位变化，只适用于振幅或强度调制的光载波的解调。直接探测是光电检测的主要方式，在“非相干光光电信号变换方法”这一章的讨论中采用的都是这种方式。

非相干探测的探测能力可用探测功率信噪比 SNR_p 和最小可测入射功率 P_{min} 表示。为计算 SNR_p ，设对应入射光功率 P_s ，在 1Ω 光电器件负载电阻上的电功率为 I_{ds}^2 ，则由上式有

$$I_{ds}^2 = (P_s \cdot \eta q / h \nu)^2$$

输出光电信号的噪声 I_N^2 包括被测辐射功率 P_s 和背景辐射功率 P_b 引起的散粒噪声、器件暗电流 I_d 的散粒噪声和探测器内阻 R 造成的热噪声，可表示为

$$I_N^2 = (2 \eta q^2 / h \nu) (P_s + P_b) + 2q I_d \Delta f + 4KT \Delta f / R$$

由上两式可得直接探测功率信噪比 SNR_p 为

$$SNR_p = (q \eta P_s / h \nu)^2 / [2 \eta q^2 (P_s + P_b + h \nu I_d / \eta q) \Delta f / h \nu + 4KT \Delta f / R]$$

在忽略热噪声和背景辐射引起的噪声时，有

$$SNR_p = \eta P_s^2 / 2h \nu (P_s + h \nu I_d / \eta q) \Delta f$$

在只考虑受光电器件暗电流散粒噪声限制的直流探测中，信噪比为

$$SNR_p = (q / 2I_d \Delta f) (\eta / h \nu)^2 P_s^2$$

利用上式，设 $SNR_p = 1$ ，可计算直接探测情况下的最小可测入射功率 P_{min} 为

$$P_{\text{min}} = (2h \nu / \eta) (\Delta f I_d / q)^{1/2}$$

在理想情况下，器件暗电流 I_d 可忽略，只考虑信号光功率引起的散粒噪声，则有

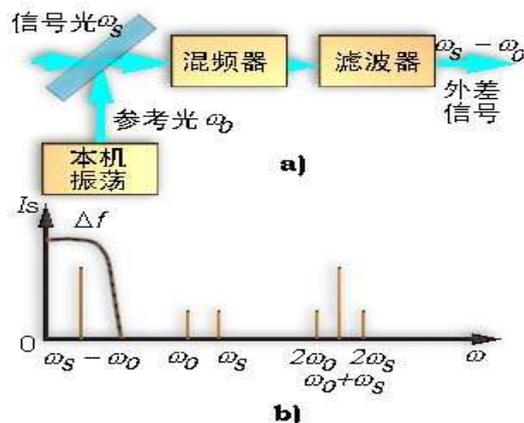
$$SNR_s = \eta P_s / 2h \nu \Delta f$$

称作直接探测的极限信噪比或量子极限。

2. 外差探测的原理和特性

在无线电电子学中广泛使用外差接收技术。被信息调制的高频载波在接收端与有一定频差的本机振荡信号相混频，得到频率为二者之差的中频信号，该信号保持了调制信号的特征，通过检测中频信号能最终解调出被传送的信息。这种方法引伸到光频干涉中发展成各种形式的光学外差干涉技术。

光学外差探测是将包含有被测信息的相干光调制波和作为基准的本机振荡光波，在满足波前匹配条件下在光电探测器上进行光学混频。探测器的输出是频率为二光波光频差的拍频信号。该信号包含有调制信号的振幅、频率和相位特征。与非相干探测相比，这种方式也称作相干探测。下图给出了光学外差探测的原理示意图。



光学外差探测原理示意图

a) 原理图 b) 频谱分布

1) 光学外差探测原理

在图 a 中，设入射混频的信号光波复振幅 U_s 和参考光波复振幅 U_0 的形式分别为

$$U_s(t) = a_s \sin(\omega_s t + \phi_s)$$

$$U_0(t) = a_0 \sin(\omega_0 t + \phi_0)$$

式中， $\omega_s = 2\pi\nu_s$ 和 $\omega_0 = 2\pi\nu_0$ 是二光波的角频率， ν_s 和 ν_0 是对应的光波频率。

采用平方律探测器进行光混频后，探测器的输出 I_{hs} 为

$$\begin{aligned} I_{hs} &= K |U_s(t) + U_0(t)|^2 \\ &= K \{U_s^2(t) + U_0^2(t) + 2U_0(t)U_s(t)\} \\ &= (K/2) \{a_s^2 + a_0^2 - a_s^2 \cos(2\omega_s t + 2\phi_s) \\ &\quad - a_0^2 \cos(2\omega_0 t + 2\phi_0) \\ &\quad - 2a_0 a_s \cos[(\omega_s + \omega_0)t + (\phi_s + \phi_0)] \\ &\quad + 2a_0 a_s \cos[(\omega_s - \omega_0)t + (\phi_s - \phi_0)]\} \end{aligned}$$

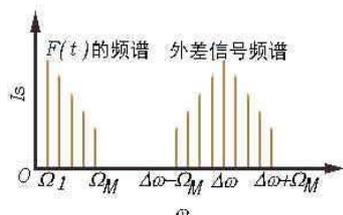
式中， K 为探测器的光电灵敏度。

由式可见混频后的光电信号包含直流分量、二倍参考光频和二倍信号光频分量以及参考光和信号光的和频和差频分量。它们的频谱分布表示在图 b 中。其中的倍频项与和频项不能被光电器件接收，只有当 ν_0 和 ν_s 足够接近使频差 $\Delta\nu = \nu_s - \nu_0$ 处于探测器的通频带 Δf 范围内才能被响应。此时探测器的输出信号变成

$$I_{hs} = K a_s a_0 \cos(2\pi \Delta\nu t + \Delta\phi)$$

式中， $\Delta\phi = \phi_s - \phi_0$ 为双频光波的相位差。

上式即为光学外差信号表达式。



调幅信号及其外差信号的频谱变换

在外差干涉系统中，参考光束（又称本机振荡光束或简称本振光）是二相干光的光频率和相位的比较基准。信号光可以由本振光分束后经调制形成，也可以采用独立的相干光源保持与本振光波的频率跟踪和相位同步。前者多用于干涉测量，后者用于相干通讯。不论哪种方式，由上式可知在保持本振光的 a_0 、 ν_0 、

ϕ_0 不变的前提下, 外差信号的振幅 $Ka_s a_0$ 、频率 $\Delta \nu = \nu_s - \nu_0$ 和相位 $\Delta \phi = \phi_s - \phi_0$ 可以表征信号光波的特征参量 a_s 、 ν_s 和 ϕ_s , 也就是说外差信号能以时序电信号的形式反映相干场上各点处信号光波的波动性质。即使是信号光的参量受被测信息调制, 外差信号也能无畸变地精确复制这些调制信号。这一点可以用简单的调幅信号加以说明。设信号光振幅 a_s 受频诸如上图中的调制信号 $F(t)$ 的调幅, 则上式中的 $a_s(t)$ 为

$$a_s(t) = A_0 [1 + F(t)] = A_0 \left[1 + \sum_{n=1}^M m_n \cos(\Omega_n t + \phi_n) \right]$$

式中, A_0 是调制信号的振幅; m_n 、 Ω_n 和 ϕ_n 分别是调制信号各频谱分量的调制度、角频率和相位。将上式代入原式中, 可得外差信号为

$$\begin{aligned} I_{hs} &= Ka_0 A_0 \left[1 + \sum_{n=1}^M m_n \cos(\Omega_n t + \phi_n) \right] \cos(\Delta \omega t + \Delta \phi) \\ &= Ka_0 A_0 \cos(\Delta \omega t + \Delta \phi) \\ &\quad + Ka_0 A_0 \sum_{n=1}^M (m_n/2) (\Delta \omega + \Omega_n) t + (\Delta \phi + \phi_n) \\ &\quad + Ka_0 A_0 \sum_{n=1}^M (m_n/2) \cos[(\Delta \omega - \Omega_n) t + (\Delta \phi - \phi_n)] \end{aligned}$$

它的频谱分布表示在上图中的对应位置上。由图及上式可见信号光波振幅上所载荷的调制信号双道带地转换到外差信号上去。对于其它种调制方式也有类似的结果。这是直接探测所不可能达到的。

在特殊的情况下, 若使本振光频率和信号光频率相同, 则原式变成

$$I_{hs} = Ka_s a_0 \cos \Delta \phi$$

这就是零差探测的信号表达式。式中 a_s 项也可以是调制信号。例如在调幅波情况下, 可得零差信号为

$$\begin{aligned} I_{hs} &= Ka_0 A_0 \cos \Delta \phi + Ka_0 A_0 \left[\sum_{n=1}^M (m_n/2) \cos(\Omega_n t + \phi_n + \Delta \phi) \right. \\ &\quad \left. + \sum_{n=1}^M (m_n/2) \cos(\Omega_n t + \phi_n - \Delta \phi) \right] \end{aligned}$$

简化计算, 令 $\Delta \phi = 0$, 则

$$I_{hs} = Ka_0 A_0 \left[1 + \sum_{n=1}^M m_n \cos(\Omega_n t + \phi_n) \right]$$

这表明零差探测能无畸变地获得调制信号的原形, 只是包含了本振光振幅的影响。此外, 在信号光不作调制时, 零差信号只反映相干光振幅和相位的变化而不能反映频率的变化, 这就是单一频率双光束干涉相位调制形成稳定干涉条纹的工作状态。

2) 光学外差探测的特性

与直接探测相比, 外差探测有许多突出的特点:

a) 探测能力强。外差探测不仅能检测出振幅和强度调制的光波而且可以检测相位和频率调制的光波, 因此是一种测试光的波动特性的有力方法。

b) 探测灵敏度高。外差探测中光电探测器输出电流幅值 I_{hsm} 为

$$I_{hsm} = Ka_0 a_s = 2K \sqrt{P_0 P_s}$$

式中, P_0 和 P_s 分别是信号和本振光的功率。

上式表明外差探测的输出信号取决于入射信号光和本振光功率的乘积, 本振光功率对外差信号有决定作用。提高本振光功率 (例如比信号光高 $10^7 \sim 10^8$ 倍) 可以得到较强的外差信号。与直接探测相比, 它们的信号转换功率比 G 为

$$G = I_{hs}^2 / I_{ds}^2 = 4K^2 P_s P_0 / K^2 P_s^2 = 4P_0 / P_s$$

通常 G 可以高达 $10^7 \sim 10^8$ 。

c) 信噪比高。外差信号电流均方功率 $\overline{I_{ks}^2}$ 为

$$\overline{I_{ks}^2} = 2(\eta q/h\nu)^2 P_s P_0$$

对于受限于散粒噪声的检测器, 考虑到关系 $P_0 \gg P_s$, 噪声的均方功率 $\overline{I_N^2}$ 为

$$\overline{I_N^2} = 2q \Delta f (\eta q/h\nu) P_0$$

所以外差探测信噪比 SNR_h 为

$$SNR_h = \overline{I_{ks}^2} / \overline{I_N^2} = \eta P_0 / h\nu \Delta f$$

最小可测入射功率 P_{hmin} 可计算为

$$P_{hmin} = \Delta f h\nu / \eta$$

与直接探测情况相比较, 有

$$P_{dmin}/P_{hmin} = 2(I_d / \Delta f q)^{1/2}$$

通常, $P_{hmin} \ll P_{dmin}$, 例如 $I_d = 100\text{pA}$, $\Delta f = 100\text{Hz}$ 时, 有 $P_{hmin} < 0.2 \times 10^{-3} P_{dmin}$, 这表明外差探测可以检测到更小的入射功率, 特别是在检测器件暗电流 I_d 较大和信道通频带 Δf 较窄的情况下外差探测具有更大的优越性。

d) 有空间滤波能力。为了形成外差信号, 对信号光和本振光要求较高的波前匹配条件。例如空间对准、偏振相同等。如果背景辐射杂散光入射方向杂乱, 偏振状态不确定, 不能满足这些条件, 则不能形成有效的外差信号。所以外差探测有较强的空间滤波能力。

e) 有光谱滤波能力。只要二相干光波频率是稳定的, 当检测通道的通频带刚好覆盖有用外差信号的频谱范围, 则在此通带外的杂光即使形成拍信号也将被滤掉。因此外差探测系统即使不加光谱滤光片其效果也比加滤光片的直接探测好得多。

f) 检测的稳定性和可靠性高。外差信号通常是交变的射频或中频信号, 并且多采用频率和相位调制, 即使被测参量为零, 载波信号仍保持稳定的幅度。对这种交流的测量系统, 系统直流分量的漂移和光信号幅度的涨落不直接影响检测性能, 能稳定可靠地工作。

3) 实现外差探测的匹配条件

尽管外差探测具有许多优点, 但为获得最佳的外差信号, 要求相干光束满足一定的条件。

a) 参于混频的信号光和本振光应是理想相干的单频单模光, 要有稳定的振荡频率和相位。通常两光束取自同一激光器, 通过频率偏移器取得本振光或经被测信息调制得到信号光。在外差通讯时, 光发射和接收不在同一地点的情况下, 要求本振光对信号光载波保持频率跟踪。

b) 在光混频器上信号光与本振光要求偏振方向一致。

c) 信号光与本振光要求空间波前匹配, 即要求空间调准(准直、共轴), 波面吻合垂直入射于光混频表面。二光束入射角偏斜 θ 应满足关系

$$\theta < \lambda_0 / \pi l$$

式中, λ_0 是本振光波长, l 是光电探测器光敏面尺寸。

d) 光电探测器应有灵敏度均匀的光敏表面, 足够的高频响应和稳定的量子效率。外差探测通常采用光电倍增管、PIN管和雪崩光电管等。

4) 外差信号的检测

外差信号的频率范围取决于本振光和信号光的频差。通常为 10^4 到 10^{10}Hz 数量级。主要检测光拍信号的方法包括: 用电子示波器或中、高频频谱分析仪测试信号波形或频谱分量; 经选通放大器放大外差信号后用时域微分鉴频、斜率鉴频或相位鉴频等常用电子装置测定外差拍频; 对于射频外差信号, 采用射频混频的电信号外差接收, 转换为较低频信号后进行频率测量。更多的检测方法要根据光调制的方式和外差信号的频谱范围自行设计。

10.3 相干光的相位调制和检测

单频光相位调制和条纹检测

在使用窄光束单频光波相位调制的干涉测量中，干涉条纹的形成和检测是在光束重叠的较小空间范围内进行的，通常采用单元光电器件检测局部位置上的干涉条纹波数或相位随时间的变化。

1. 单频光的相位调制

在单一频率相干光路中，被测量使相干光波的相位发生变化，同时通过干涉作用把波相位的变化变换为振幅的变化，这个过程称单频光波的相位调制或称相幅变换。由前面的公式可知，能引起相位变化的参量是光路长 L 和介质折射率 n 。因此相位调制通常是利用不同形式的干涉仪，借助机械的、光学的、光电子学等变换器件将被测量的变化转换为光路长 L 和折射率 n 的变化。前者用来检测几何和机械运动参量，后者用于分析物质的理化特性。

为了定量描述被测参量对相位调制的影响，采用规一化相位响应表示在单位长度的光路内由被测参量引起的相位变化。

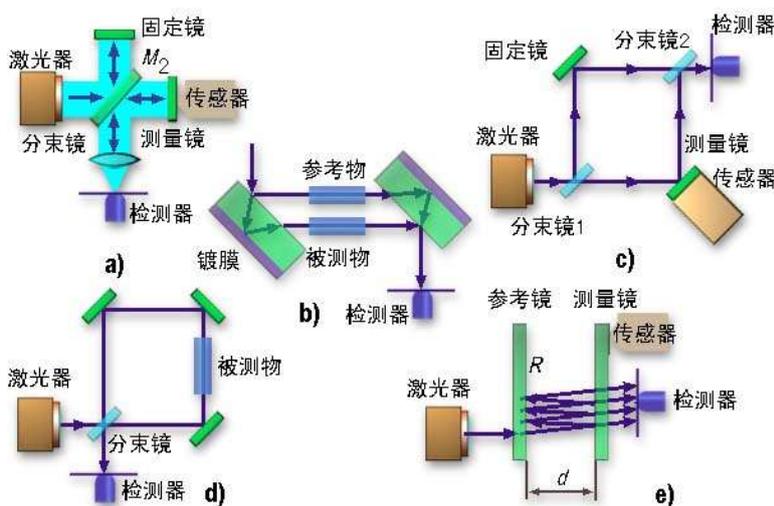
$$(1/L) (d\phi / dF) = (2\pi / \lambda_0) [dn/dF + (n/L) (dL/dF)]$$

式中， $(1/L) (d\phi / dF)$ 为规一化相位响应， L 为干涉光路长度， F 为被测参量。

等式右端两项分别表示折射率变化和光路长度变化引起的相位响应。上式可用来衡量相位调制的各种类型光学干涉仪和光纤干涉仪的工作特性。

1) 光学干涉仪相位调制

通常作为相位调制用的光学干涉仪有迈克尔逊干涉仪、吉曼干涉仪、马赫-泽德干涉仪、萨纳克干涉仪和法布里-珀罗干涉仪等。下图给出了它们的原理示意图。



典型的光学干涉仪原理示意图

除了法布里-珀罗干涉仪外，前述干涉仪皆属双光束干涉。干涉强度分布满足公式。图 a 的迈克尔逊干涉仪其特点是结构简单，条纹对比度好，信噪比高。测量镜 M_2 与被测物连接可以感知位移、变形等参量。由于 M_2 的位移量 Δx 引起测量光路 $2\Delta x$ 的变化，即 $\lambda/2$ 的位移引起干涉条纹一个周期的变化，所以条纹的计数和被测位移的计算关系简单。它的测量灵敏度达 10^{-13} m 的数量级。其缺点是输出光束能经分束镜返回

激光器，这将使激光器工作不稳定，这可以通过设置偏振器来防止。图 b 是吉曼干涉仪。同样厚度的二块玻璃板背面镀以反射膜，利用两玻璃表面的反射形成光束的分束和再合成。由于两光路的光程差很小，即使相干性较差的光源也可进行精密测量。它主要用来测定透光物质（例如气体）的折射率，可进行标准试样和被测试样的比较测量。若试样长度为 L，条纹测量精度为 $\lambda/50$ ，则折射率误差在 $\delta n = \lambda/50L$ 之内。图 c 是马赫-泽德干涉仪，由二片分束镜和二片反射镜组成。输出分束镜有两束干涉光输出，可用于布置多路接收器，它的返回散射光较少，有利于降低激光的不稳定噪声。被测位移的引入通过可移动反射镜进行，位移范围不能超过相干光束的截面。图 d 的萨纳克干涉仪其入射分束镜和出射合成镜合并为一。入射分束镜把入射光束分成二个传输方向相反的顺时针光路和逆时针光路，闭合回路的形状可以是任意的如矩形或三角形。反射镜的法向位移对两个反向光路长度的改变量相等，因此不引起干涉强度的变化。它用来感知如转角、转速以及磁场强度等非互易效应。图 e 是法布里-珀罗干涉仪。它的两块反射镜的反射率 R 高达 95% 以上。入射光束经不同次数的反射后分别透射，以平行光形式输出，由光电探测器接收，因此是多光束干涉。相干光强的变化表示为

$$I = I_0 / \{1 + [4R^2 / (1-R)^2] \sin^2(\phi/2)\}$$

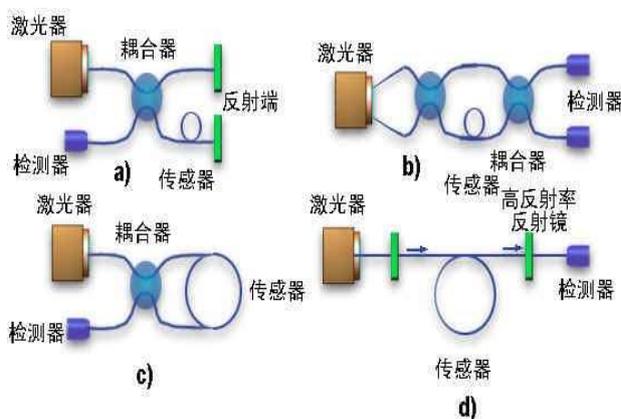
及

$$\phi = 4\pi nd / \lambda_0$$

式中， $F = 4R^2 / (1-R)^2$ 称作精细度系数，可表征干涉条纹的锐度（如图 b）；是相邻反射光束间的相位差； I_0 是平均光强度；d 为平面镜间隔； λ_0 为真空波长；n 为反射镜间介质折射率。

这两个式子表明透射干涉光强不是正弦分布，当 ϕ 为 2π 整数倍时光强为最大值， ϕ 为 π 的奇数倍时光强为最小值。最大最小值之比为 $(1+R)/(1-R)$ ，因此有很高的条纹对比度，是一种灵敏的传感器。由后一式可见，被测变量的调制作用可以通过改变 d 和 n 实现，后者可用更换平面镜之间的气体等方式进行。此外也可实现高分辨率波长变化的测量，其分辨率可达 $2 \times 10^{-5} \text{nm}$ 。光学干涉仪的共同特点是相干光在空气中传播，环境温度的改变会引起空气折射率的扰动；大气湍流和声波干扰也会导致光程的变化，降低了工作可靠性和测量精度。

2) 光纤干涉仪相位调制



光纤干涉仪

另一类相位调制可采用单模光纤作封闭光路组成的各种类型的光纤干涉仪。上图给出了它们的原理示意图。其中图 a 为迈克尔逊干涉仪，图 b 为马赫-泽德干涉仪，图 c 为萨纳克干涉仪，图 d 为法布里-珀罗干涉仪。在这些结构中，光纤光程代替了空气光程，定向光束耦合器取代了分束器，它能按一定比例将光束由一束光纤耦合到另一束光纤中，以实现光束分割和合成。由于光路的闭合避免了空气的扰动，并且不受结构空间的限制可以组成千米数量级长度的干涉仪，因此有利于提高测量灵敏度。光纤本身作为被测参量的敏感元件直接置于被测环境中，通过不同的物理效应感知被测参量的变化，形成光纤中光波相位、频率、振幅、偏振态的改变。与光学干涉仪相比，这种调制作用是通过光纤的内在性能达到的。

a) 机械应力效应相位调制

光纤受纵向机械力作用时，纵向应变引起光纤长度的变化，这种效应导致光纤中光相位的改变。对于单模光纤，设光纤的敏感长度为 L ，光纤芯和包层内平均折射率为 n ，光纤的光弹系数为 P_{12} ，纵向长度变化为 ΔL ，相位调制 $\Delta \phi$ 可表示为

$$\Delta \phi = (2 \pi L / \lambda_0) n^3 (1 - \nu^2 P_{12} / 2) (\Delta L / L)$$

纵向应力作用下的泊松效应引起光纤芯的直径变化。对于轴对称的光纤，相位变化与径向应变成比例，比例因子和光纤长度、折射率和光纤材料有关。此外当光纤受纵向应力作用时，由于光弹效应的影响引起折射率的变化。基于这些调相作用组成了许多力的传感器。

b) 温度胀缩效应相位调制

光纤处于被测温度场中，热胀冷缩引起长度变化，温度也能造成折射率的改变。由此引起的光波相位响应可由原式将被测参量 ΔF 用温度变化 ΔT 代替得到

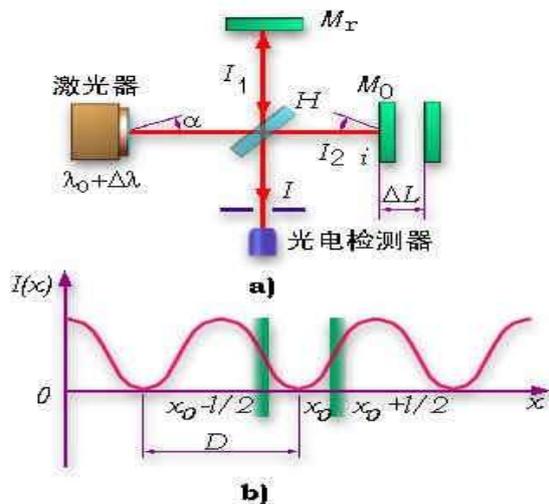
$$d \phi / dT = (2 \pi L / \lambda) [dn/dT + (n/L) (dL/dT)]$$

上式表示光波相位响应与光纤的敏感长度 L 成正比。选择足够长的光纤同时优化光纤材料和形状可使热胀系数达 $(4 \sim 13) \times 10^{-6} / ^\circ\text{C}$ ，折射率变化 dn/dT 在 $(-10 \sim +19) \times 10^{-6} / ^\circ\text{C}$ 范围内。用光纤检测温度变化可以获得 $10^4 / ^\circ\text{C}$ 信号脉冲的温度灵敏度。

2. 单频光干涉的条纹检测

在局部空间范围内检测干涉条纹的光强度或其随时间的变化称作干涉条纹检测。与此相对应在一定的空间范围内确定整个干涉图样的分布称作干涉图分析。基本的条纹检测法包括条纹光强检测法、条纹比较法和条纹跟踪法。

1) 条纹光强检测法



条纹光强的光电检测

a) 原理示意图 b) 波形图

这种方法是在干涉场中确定的位置上用光电元件直接检测干涉条纹的光强变化。上图 a 给出了一维干涉测长的实例。为了获得最佳的条纹光电信号，要求有最大的交变信号幅值和信噪比，这需要光学装置和光电检测器确保最佳工作条件，尽可能地提高两束光的相干度和光电转换的混频效率。

由前可知，单频相干光合成时的瞬时光强为

$$I(x, y, t) = a_1^2 + a_2^2 + 2a_1a_2 \cos[\phi(t)]$$

上式只有在检测时间 τ 内 $\cos[\phi(t)]$ 为恒定时才能得到确定的光强值，若 $\phi(t)$ 随时间变化，则合成光强是对 t 的积分

$$I(x, y, t) = a_1^2 + a_2^2 + 2a_1a_2 (1/\tau) \int_0^\tau \cos[\phi(t)] dt$$

用 $\phi(t)$ 的平均值 ϕ_0 等效表示这一积分值，即令

$$(1/\tau) \int_0^\tau \cos[\phi(t)]dt = \Gamma \cos \phi_0$$

式中，比例因子 Γ 称作两光束的相干度， $0 \leq \Gamma \leq 1$ 。当 $\Gamma = 1$ 时表示在 τ 时间内相位保持不变，相干度最大。当 $\Gamma = 0$ 时，表示 τ 时间内两光束不相干。

将上式代入原式中，得

$$I(x, y, t) = a_1^2 + a_2^2 + 2a_1a_2\Gamma \cos \phi_0$$

并有

$$\Gamma = \{(1/\tau) \int_0^\tau \cos[\phi(t)]dt\} / \cos \phi_0$$

上式表明，只有当 Γ 大时光强随相位的变化才明显，而当 $\Gamma = 0$ 时合成光强只有直流分量，与相位 ϕ 无关。因此相干度 Γ 是衡量干涉条纹光强对比的质量指标。

相干光源波长的非单色性 $\Delta \lambda$ 引起不同波长不同初相位的叠加，这会降低相干度。对常用干涉仪，可以计算在中心波长为 λ_0 、波长范围为 $\pm \Delta \lambda$ 时其相干度 Γ_λ 为 $\Delta \lambda$ 的 sinc 函数，有

$$\Gamma_\lambda = \text{sinc}(\Delta \lambda \cdot 2\pi \Delta L / \lambda_0^2)$$

上式表明，光程差 ΔL 愈小及单色性愈好 ($\Delta \lambda$ 愈小) 则 Γ_λ 愈大。当光程差 ΔL 等于单色光相干长度时干涉条纹消失。

相干光源的发散使不同光线产生不同的光程差，这引起相位 ϕ 的变化。对于平板干涉的情况，若光束发散角为 α ，入射光不垂直反射镜的偏角为 i 时，可以计算有附加光程差的相干度 Γ_α 为

$$\Gamma_\alpha = \text{sinc}(2\pi n \Delta L \alpha \sin i / \lambda_0)$$

这表明空间每条相干光线光程差不同会引起条纹信号交变分量的下降。

光电检测器把光信号转换为电信号，得到的是光敏面上的光强积分值。光电信号的质量不仅取决于干涉条纹的相干度，而且取决于接收器光阑和条纹宽度之间的比例关系。在图 8b 中，设接收光阑是 $h \times l$ 的矩形，由均匀照明光产生的平行直条纹的间距为 D ，空间坐标为 x ，则沿 x 向的条纹光强空间分布 $I(x)$ 为

$$I(x) = a_1^2 + a_2^2 + 2a_1a_2\Gamma \cos(x \cdot 2\pi / D)$$

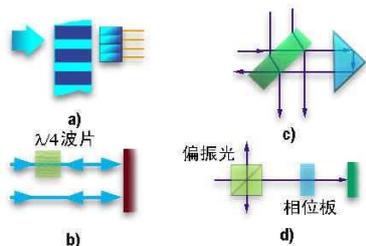
在不同位置 x 上光电检测器输出电信号 I_s 为

$$\begin{aligned} I_s &= K \int_{-h/2}^{+h/2} dy \int_{-l/2}^{+l/2} [a_1^2 + a_2^2 + 2a_1a_2\Gamma \cos(x \cdot 2\pi / D)] dx \\ &= Kh l \{ (a_1^2 + a_2^2) \\ &+ 2a_1a_2\Gamma [\sin(\pi l/D) / (\pi l/D)] \cos(x \cdot 2\pi / D) \} \\ &= Kh l [a_1^2 + a_2^2 + 2a_1a_2\Gamma \beta \cos(x \cdot 2\pi / D)] \end{aligned}$$

式中， $\beta = \text{sinc}(\pi l/D)$ 称作光电转换混频效率， $0 < \beta < 1$ 。

当 $l/D \rightarrow 0$ ， $\beta = 1$ 时光电信号交变分量幅度最大；当 $l = D$ ， $\beta = 0$ 时光电信号只有直流分量。由式可见混频效率 β 或光阑宽度与条纹宽度比 l/D 直接影响电信号的幅值。为了增大 β 值，在 D 值确定时应减少 l 值，但这样将降低有用光信号的采集。正确的作法是使干涉区域充分占据接收光阑，通过加大条纹宽度来增大 β 值。这一结果不论对采用均匀扩束照明还是采用单束激光（光束截面强度呈高斯分布）照明或者是采用圆孔形光阑的情况都是适用的。

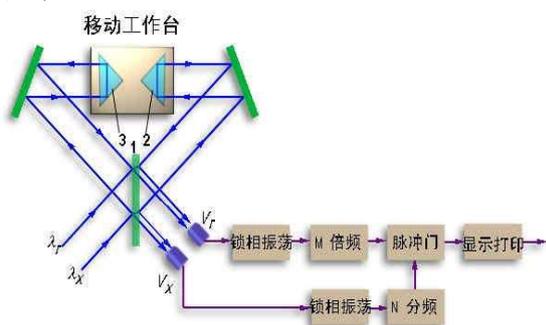
使用干涉仪作位移测量时为了消除振动的干扰和进行双向测量，条纹的检测通常采用可逆计数方式。为此要求检测装置提供彼此正交的两路条纹信号。此外正交信号的差分连接能消除直流光强的影响，并能进行倍频细分的计算。产生 $\pi/2$ 相位差信号的方法如下图所示，其中图 a 的条纹图形法是在干涉条纹分布的 $1/4$ 周期位置上分别放置点状检测器或采用四分光电池与整个条纹周期对应。相位板法（图 b）利用光学晶体作成的 $\lambda/4$ 波片对光波作相位延迟取得 $\pi/2$ 的光波相位差。图 c 的光路移相法利用组成干涉仪的角反射镜或多层镀膜半透镜在光反射时的相位变化形成 $\pi/2$ 的相位差。前述方法中正交两路信号是由干涉视场内不同区域获得的，所以易受到环境扰动的影响。偏光法不易受外界扰动，在干涉仪中采用偏光分束镜。利用光束的偏光作用形成 $\pi/2$ 的相位差（图 d）。



产生正交干涉信号的方法

2) 干涉条纹比较法

对于如图的干涉仪，若采用二束不同光频的相干光源，各自独立地组成干涉光路，使其中一束光频为已知，另一束是未知的，则对应共用测量反射镜的同一位移，两光束各自形成干涉条纹。经光电检测后形成二组独立的电信号。通过电信号频率的比较可以计算出未知光波的波长。这种对应同一位移，比较不同波长的两个光束干涉条纹的变化差异的方法称作干涉条纹比较法。从这种原理出发，设计出了许多精确测量波长的波长计。



条纹比较法波长计原理图

1-半透半反镜 2、3-圆锥角反射镜

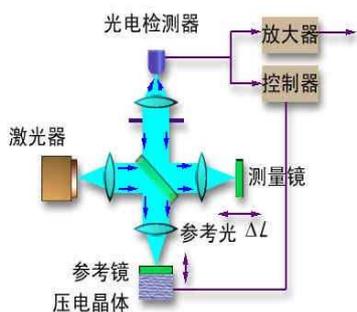
上图是波长测量精度为 10^{-7} 的条纹比较法波长计的简化原理图。已知波长 λ_r 的基准波和被测波长 λ_x 的光波由半透半反镜 1 分别投射到放置于移动工作台上的两个圆锥角反射镜 2 和 3 上。使二束光的入射位置分别处于弧矢和子午方向，保证它们在空间上彼此分开。每束光束的逆时针反射光和顺时针反射光在各自的光检测器 V_r 和 V_x 上形成干涉条纹。对应于工作台的同一位移，由于两束光的波长不同，产生的干涉条纹也有不同的变化周期，因而对应的光电信号显示出不同的频率。精确地测量出两信号的频率比值，根据基准波长的数值即能计算出被测波长值。在所介绍的装置中频率比的测定采用了锁相振荡计数的作法。两个锁相振荡器分别与 V_r 和 V_x 光电信号同步，产生与 λ_r 和 λ_x 的干涉条纹同频的整形脉冲信号。其中与 λ_r 对应的脉冲信号经 M 倍频器作频率倍频，而与 λ_x 对应的信号则作 N 倍分频。利用脉冲开关由 N 分频信号控制 M 倍频信号进行脉冲计数，最后由显示器输出。被测波长的计算按下式进行

$$\lambda = (\lambda_r/M) (c/N) (1 + \Delta n/n)$$

式中，c 为脉冲计数器的计数值， $\Delta n/n$ 是折射率的相对变化。

3) 干涉条纹跟踪法

这是一种平衡测量法。在干涉仪测量镜位置变化时，通过光电接收器实时地检测出干涉条纹的变化。同时利用控制系统使参考镜沿相应方向移动，以维持干涉条纹保持静止不动。这时，根据参考镜的位移驱动电压的大小可直接得到测量镜的位移。下图表示了利用这种原理测量微小位移的干涉测量装置。这种方法能避免干涉测量的非线性影响，并且不需要精确的相位测量装置。但是所用跟踪系统的固有惯性限制了测量的快速性，因此只能测量 10kHz 以下的位移变化。

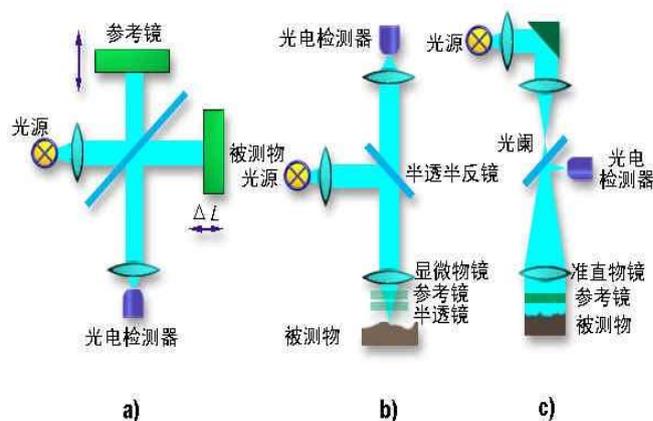


条纹跟踪法干涉系统示意图

单频光波面的相位调制和干涉图分析

当将相干光扩束成平面波或其它规则波面照射到被测物体时，干涉仪的相干场上会形成由干涉条纹组成的二维干涉图。它反映了被测物表面的微细面形分布或透射介质的折射率分布。这是对二维单频光波面的各点形成相位调制的结果。

1. 单频光波面的相位调制



可作波面相位调制的干涉仪

- a) 泰曼-格林干涉仪
- b) 米勒干涉仪
- c) 斐索干涉仪

常见的波面相位调制采用泰曼-格林干涉仪（上图 a）。设被测物波面 $U_0(x, y)$ 的相位为 $\phi_0(x, y)$ ；参考光是理想的平面波 $U_r(x, y)$ ，它的初始波相位 ϕ_r 空间不变；两光波的频率相同。此种情况下原式可简化为

$$U_0(x, y) = a_0(x, y) \exp\{-j\phi_0(x, y)\}$$

$$U_r(x, y) = a_r \exp\{-j\phi_r\}$$

干涉场上条纹的强度分布 $I(x, y)$ 为

$$I(x, y) = a_0^2(x, y) + a_r^2 + 2a_0(x, y) a_r \cos[\phi_r - \phi_0(x, y)]$$

由以前的式子，考虑到泰曼-格林干涉仪的光路折返，上式中的相位分布 $\Delta\phi(x, y) = \phi_r - \phi_0(x, y)$ 和折射率分布 $\Delta n(x, y)$ 、物面变形分布 $\Delta L(x, y)$ 之间有关系

$$\Delta\phi(x, y) = (4\pi / \lambda_0) [L\Delta n(x, y) + n\Delta L(x, y)]$$

通过干涉图分析，计算出 $\Delta\phi(x, y)$ 后，借助该式可获得被测量的 $\Delta n(x, y)$ 或 $\Delta L(x, y)$ 值。

能进行波面相位调制的干涉仪还有米勒干涉仪（图 b）和斐索干涉仪（图 c）。在米勒干涉仪中，光源经半透半反镜和长工作距显微物镜照明被测物面。反射波面经物镜成象在接收面上。在物镜和被测物间放置半透半反镜和参考镜，经参考镜和半透半反镜反射的波面作为参考光同时被成象在接收面上，与物光形成干涉条纹。由于所得到的干涉图是经显微镜放大的，所以适合于微细表面的干涉测量。在斐索干涉仪中激光束经聚光镜会聚在准直物镜的焦点上。由物镜出射的平行光垂直入射于半透半反参考镜上，其中的反射光作为参考光束经物镜会聚在干涉平面上。透射光经被测平面反射，再经物镜会聚与参考光束相遇，形成双光束干涉条纹。这些干涉仪所得到的平面干涉图适用于物体表面形状和粗糙度的测量以及非球面光学零件的

检验等。此外能获得形状和折射率分布的干涉图还有莫尔拓扑图、全息干涉图、散斑干涉图等。精确地测定干涉图形的条纹空间分布，计算各空间位置的相位值及其相对应的面形要进行干涉图的分析。

2. 干涉图的图样分析法

干涉图的分析判读通常包括：利用光电器件阵列、扫描装置或摄像装置采集干涉场的光度分布；计算确定干涉图各处的条纹间距和走向；确定各坐标点处的相位值；判断相位的极性符号；对于确定的干涉系统计算干涉图对应的被测面形或折射率分布。

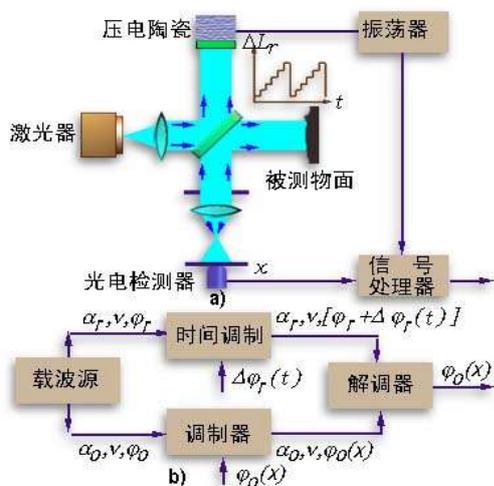
常用的干涉图分析是将干涉图看作是明暗相间的条纹图样，利用常规的数字图象处理技术进行。基本的步骤是首先通过低频滤波消除相干噪声和进行照明背景的阴影补偿，取得振幅均匀的条纹信号。其次通过条纹边缘或峰值的检测来确定条纹的中心分布和走向，称作条纹细线化。然后计算条纹的间距和各点的相位值，并完成被测数值的最终计算，给出数字的或图示的结果。

数字图象处理的方法需要较大型的数据处理计算机和相应的软件，且计算工作量较大。由于干涉条纹光强测量精度的限制，并且某点处的空间相位是从与周围条纹分布的比较中得到的，因此它的空间分辨能力和相位测量精度受到限制。在多数情况下的实际精度不超过 $\lambda/20$ 。更重要的是为了判断面形的凸凹，必须确定相位的极性符号。这需要更多的数据和复杂的判断，因而通常采用人机对话加以辅助。

二次相位调制和干涉图的动态检测

为了自动分析干涉图，近年来发展出各种能进行动态相位检测的干涉测量技术。其基本原理是使参考光的相位人为地随时间调制（称作二次调制），使干涉图上各点处的光学相位变换为相应点处时序电信号的相位。利用扫描或阵列检测分别测得各点的时序变化，就能以优于 $\lambda/100$ 的相位精度和 100 线对/mm 的空间分辨率测得干涉图的相位分布。这些干涉图测量法包括锁相干涉和条纹扫描干涉测量。它们为干涉测量开辟了实时、数字式和高分辨率的新途径。

1. 条纹扫描干涉法



条纹扫描干涉法二次相位调制

a) 原理示意图 b) 方框图

在如图 a 的曼-格林干涉仪中，若使用压电陶瓷等位移变送器使参考反射镜周期性地随时间成比例地移动 $\lambda/2$ ，在干涉面的各点上将形成同样周期的正弦型强度变化，在不同位置上时序正弦信号的初始相位与该点处被测波面的初始相位对应。用光电法比较各点处电信号的相位就可以计算出被测表面的形状分布，这是条纹扫描干涉法的基本原理。与前面提到的使被测变量直接进行相幅变换形成稳定的干涉条纹相比，这种首先人为地引入相位时间调制，然后进行被测变量相位调制的方法称作二次相位调制。这种用测量并比较时序信号相位来代替测量光强空间分布，解算出空间相位的作法，能有效地提高测且精度。它的调制信号流程表示在图 b 中。

在工程上为了便于数据采集，常使参考光路按阶梯波形（如图 a）变化，在参考镜所处的每一个阶梯位

置上用面阵图象传感器对干涉面上各点的光强值取样，对于每个点，用傅里叶和式累加各个阶梯上的测量结果可拟合出正弦变化曲线。由此可得到干涉面上各点的相对相位分布。

为使讨论简单，这里只分析一维的情况。由前面的式子可以看出，干涉面上任一点 x 的光强度可表示为 $I(x) = a_0^2(x) + a_r^2 + 2a_0(x)a_r \cos[\phi_r - \phi_0(x)]$

将上式看作是 ϕ_r 的余弦函数，它可表示成有直流分量和基波分量的傅里叶级数形式

$$I(x, \phi_r) = U_0(x) + U_1(x) \cos \phi_r + m_1(x) \sin \phi_r$$

令 ϕ_r 在 2π 周期内每次改变 $1/n$ 周期，共采样 p 个周期，即

$$\phi_{rj} = j(2\pi/n) \quad j=1, 2, 3, \dots, np$$

则与 ϕ_{rj} 对应的干涉强度可用上面的 $I(x, \phi_r)$ 式表示，此时式中 ϕ_r 用 ϕ_{rj} 代替。应用三角函数的正交关系， $I(x, \phi_{rj})$ 的各系数表示为

$$U_0(x) = (1/np) \sum_{j=1}^{np} I(x, \phi_{rj}) = a_0^2(x) + a_r^2$$

$$U_1(x) = (1/np) \sum_{j=1}^{np} I(x, \phi_{rj}) \cos \phi_{rj} = 2a_0(x)a_r \cos \phi_0(x)$$

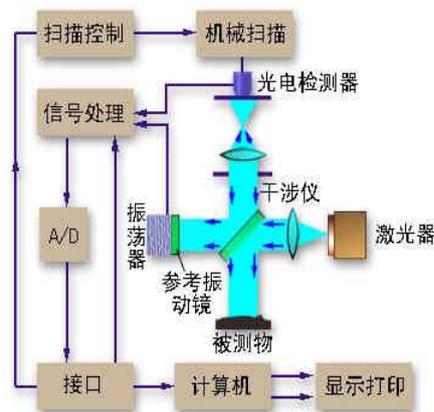
$$m_1(x) = (2/np) \sum_{j=1}^{np} I(x, \phi_{rj}) \sin \phi_{rj} = 2a_0(x)a_r \sin \phi_0(x)$$

上式是在最小均方意义上对干涉面上光强正弦变化的最佳拟合。由此，各点处的相位值 $\phi_0(x)$ 可用两个加权平均值之比给出，即

$$\phi_0(x) = \text{arctg}[m_1(x)/u_1(x)]$$

这是对 p 次周期测量数据的累加平均。利用该式对每个测量点分别测得 $n \times p$ 个数据即可得到被测面形的相位分布图。此外，该式的计算包含着相位的符号，根据相位的连续性可去除 2π 相位的不确定性，因而可以判断面形的凸凹。这种方法测量精度可达 $\lambda/100$ ，空间分辨率由象感器件决定，例如可取 32×32 或 100×100 光敏阵列等。

2. 正弦相位时间调制干涉法



正弦波相位调制干涉仪原理示意图

在如图的泰曼-格林干涉仪中，参考光路设置压电元件带动反射镜按正弦规律振动，使相干光束间的相位差作周期性变化，从而引入正弦波相位时间调制。设振动的振幅为 ΔL_m ($\Delta L_m < \lambda/2$)，角频率为 ω_c ，则相位的时间调制项 $\delta\phi(t)$ 可表示为

$$\delta\phi(t) = A \sin \omega_c t$$

式中， $A = \Delta L_m 4\pi / \lambda$ 为相位调制振幅。这时干涉场中光强表达式在一维情况下简化为

$$I(x, t) = a_0^2(x) + a_r^2 + 2a_0(x)a_r \cos[\phi_r - \phi_0(x) + A \sin \omega_c t]$$

利用三角函数的加法定理和贝塞尔级数展开式该式可写成

$$\begin{aligned} I(x, t) &= a_0^2(x) + a_r^2 \\ &\quad + 2a_0(x)a_r \cos[\phi_r - \phi_0(x)] \cos[A \sin \omega_c t] \\ &\quad - 2a_0(x)a_r \sin[\phi_r - \phi_0(x)] \sin[A \sin \omega_c t] \\ &= a_0^2(x) + a_r^2 \\ &\quad + 2a_0(x)a_r \cos[\phi_r - \phi_0(x)] [J_0(A) + 2J_2(A) \cos 2\omega_c t + \dots] \\ &\quad - 2a_0(x)a_r \sin[\phi_r - \phi_0(x)] [2J_1(A) \sin \omega_c t \\ &\quad + 2J_3(A) \sin 3\omega_c t + \dots] \end{aligned}$$

若振幅 A 很小, 则有下列近似

$$J_0(A) \approx 1 - A^2/4; \quad J_1(A) \approx A/2; \quad J_2(A) \approx A^2/8; \quad J_n(A) \approx 0 \quad (\text{当 } n \geq 3)$$

此时上式变成

$$\begin{aligned} I(x, t) &= a_0^2(x) + a_r^2 \\ &\quad + 2a_0(x)a_r \cos[\phi_r - \phi_0(x)] [1 - A^2/4 + (A^2/4) \cos 2\omega_c t] \\ &\quad - 2a_0(x)a_r \sin[\phi_r - \phi_0(x)] A \sin \omega_c t \end{aligned}$$

用光电接收器检测上述信号, 经过滤波和对调制频率选频放大后得到与相位调制频率相同的基频分量 i_s

$$i_s(x, t) = -2a_0(x)a_r \sin[\phi_r - \phi_0(x)] A \sin \omega_c t$$

上式表明通过正弦波相位时间调制, 检测的条纹信号是正弦时序信号, 它的频率是时间调制的频率, 振幅与检测点处波面的相位差成正弦关系。

采用相位锁定法可以测得被测相位差。实际上, 若在时间调制的同时给压电元件施加直流偏置位移 δL , 它将引起 $\delta \phi = (4\pi/\lambda)\delta L$ 的相位偏置。这时上式变成

$$i_s(x, t) = -2a_0(x)a_r \sin[\phi_r + \delta \phi - \phi_0(x)] A \sin \omega_c t$$

当满足条件

$$\phi_r + \delta \phi - \phi_0(x) = n\pi \quad n=0, \pm 1, \pm 2, \dots$$

时, 在光电检测器上干涉条纹的强度达到最大或最小。此时利用上式根据已知的 $\delta \phi$ 可以计算出

$$\phi_r - \phi_0(x) = n\pi - \delta \phi$$

一旦条纹位置错开, 光电信号的振幅将改变, 并且随错移方向的不同, 信号极性将相反, 所以可将光电信号的振幅作为误差信号, 采用类似前图中的条纹跟踪法控制振动反射镜的动作, 直至保持光电检测器上条纹强度处于最大或最小。这时使反射镜移动的直流偏置信号 $\delta \phi$ 的大小和极性就反映了该位置上被测相位的数值。

为了实现相位的灵敏锁定, 应使前式中的 n 值取偶数, 即使 $\phi_r + \delta \phi - \phi_0(x) = 2k\pi$, 因为此时 $i_s(x, t)$ 式的电流值有最大的变化斜率。但这种状态会使锁相跟踪误差信号为零。为此, 可将测量系统的通频带宽扩大到二倍调制频率。此时光电信号的交流分量可表示为

$$\begin{aligned} i_s(x, t) &= (1/2)a_0(x)a_r \cos[\phi_r + \delta \phi - \phi_0(x)] A^2 \cos 2\omega_c t \\ &\quad - 2a_0(x)a_r \sin[\phi_r + \delta \phi - \phi_0(x)] A \sin \omega_c t \end{aligned}$$

当满足 $[\phi_r + \delta \phi - \phi_0(x)] = 0$ 时, 尽管光电信号失掉了基波调制分量, 但却保留了二次谐波项, 从而保证了锁相系统的可靠工作。

采用锁相法正弦相位时间调制能以 $\lambda/100$ 的精度测定光路长度的变化或被测物面的形状。由于采用了补偿原理, 即使相干光束的强度发生变化, 对测量精度的影响也很小。通过计算机控制光电检测器扫描整个干涉图, 可以自动完成干涉场中相位分布的动态测量。

10.4 相干光的频率调制和检测

在前面介绍了光学外差探测的原理和特性，它被广泛应用于光电干涉测量中，称作外差检测或光学差频检测。为了形成差频检测必要的光频差，工程上采用了各种频率调制技术。根据光频差的获得和差频信号检测方式的不同，差频检测大致可分作三种类型：

(1) 参量调频法

被测参量直接对参考光波的频率进行调制形成与参考光有一定频差的信号光。检测差频信号的频率或相位可以测定被测参量值。这种方式当被测参量为零时光频差为零，故有时称作零差检测。

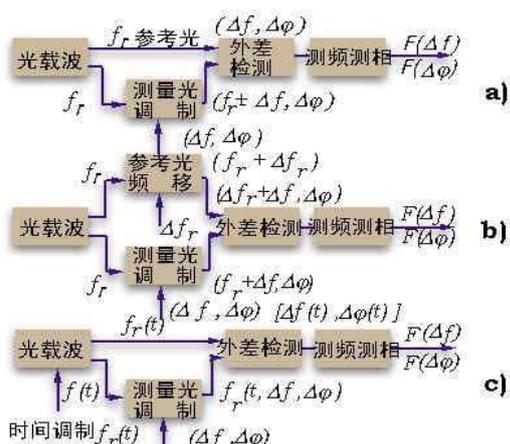
(2) 固定频移法

使用频移器件使参考光波相对信号光形成一固定的频率偏移，或利用双频光源形成有一定频差的两束相干光束，被测信号对其中一束光波进行调制（调频或调相），检测差频信号可以测定被测参量值。这种方法有时称作光学超外差。

(3) 直接调频法

利用可进行频率调制的激光器（如半导体激光器）产生随时间变化的调频参考光束。被测参量对其中一束光波作二次调制。检测外差信号可解调出被测参量值。

这三种方法的光学差频信号的形成及检测过程表示在下图所示的信号流程图中，从图中可看出它们的调制和检测的特征。



典型光学差频检测的信号流程图

a) 参量调频法 b) 固定频移法 c) 直接调频法

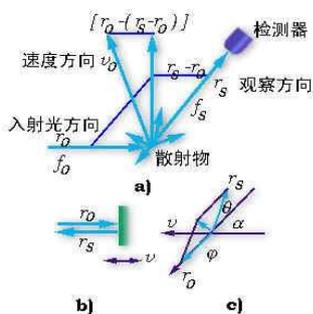
10.5 运动参量的频率调制和零差检测

1. 光学多普勒效应和运动差频检测

运动物体能改变入射于其上的波动性质（例如波动频率），这种现象称作多普勒效应，其中对入射光波频率的改变称作光学多普勒效应。对光学多普勒现象的分析表明：频率为 f_0 的单色光作用到以速度 v 运动的散射物体上，被物体散射的光波频率 f_s 会产生附加的频率偏移 Δf ，称作多普勒频移。 Δf 和散射方向有关，其数值表示为

$$\Delta f = f_s - f_0 = (1/\lambda) [\mathbf{v} \cdot (\mathbf{r}_s - \mathbf{r}_0)]$$

式中， \mathbf{v} 是物体运动速度矢量； $\mathbf{r}_s - \mathbf{r}_0$ 是散射接收方向 \mathbf{r}_s 和光束入射方向 \mathbf{r}_0 的单位矢量差，称作多普勒强度方向。该式表明：多普勒频移的大小等于散射物体的运动速度在多普勒强度方向上的分量和入射光波长的比值（如下图 a）。



光学多普勒效应

a) 多普勒频移 b) $\mathbf{r}_s = -\mathbf{r}_0$ c) 一般情况

在特殊情况下，当 $\mathbf{r}_s = -\mathbf{r}_0$ 时（如图 b），有 $(\mathbf{r}_s - \mathbf{r}_0) = -2\mathbf{r}_0$ ，代入上式有

$$\Delta f = -(2v/\lambda) \cos\theta = \pm 2|v|/\lambda$$

这正是迈克尔逊干涉仪用作速度测量时的情况。利用光学多普勒效应形成的运动频移可以测量物体的运动参数包括位移、速度和加速度等。典型的应用是激光多普勒速度计和流速计。

当激光束照射流动的散射粒子时，被运动粒子散射的激光受粒子流动速度的频率调制，得到如前式的运动频移。为检测光学差频数值，可利用光学外差法将散射前后或不同散射方向的激光在光电检测器上混频，以获得与运动速度有关的拍频信号。这种利用光学多普勒效应和光外差技术的测速装置称激光多普勒测速仪。

在一般情况下，若 \mathbf{v} 和 \mathbf{r}_0 的夹角为 α ， \mathbf{r}_0 和 \mathbf{r}_s 夹角为 θ （如图 c），前式变成

$$\Delta f = (2v/\lambda) \sin(\theta/2) \sin(\alpha + \theta/2)$$

这是多普勒测速的基本公式。当 \mathbf{r}_s 和 \mathbf{r}_0 相对 \mathbf{v} 对称布置并且满足 $\alpha + \theta/2 = 90^\circ$ 时，上式变成简单的形式

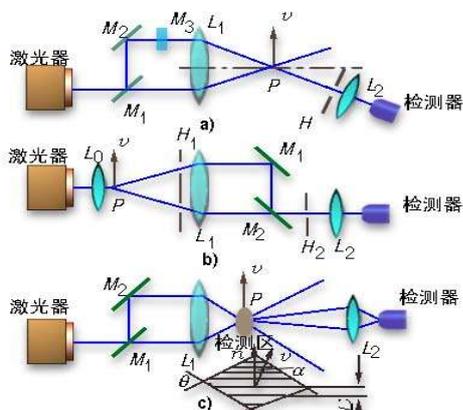
$$\Delta f = (2v/\lambda) \sin(\theta/2)$$

或

$$v = \Delta f \lambda / 2 \sin(\theta/2)$$

上式表示被测速度 v 和频差值 Δf 成正比。例如对于 $\lambda = 0.4880 \mu\text{m}$ 的氩激光，当 $\theta = 8.5^\circ$ ，被测流速 $v = 0.8$ 马赫数 $= 264\text{m/s}$ 时， $\Delta f = 77\text{MHz}$ 。

(1) 多普勒测速的频率调制方式



三种基本的多普勒速度计光路

a) 参考光束方式 b) 对称互差方式 c) 干涉条纹方式

多普勒测速仪中光频率的调制根据测量光路的不同大致可分作三种方式（如上图所示）。

1) 参考光束方式

在图 a 的示意图中，激光器发射的单色光经透镜分两路聚焦于流动颗粒 P 上。其中经反射镜 M₂ 的光束未经颗粒散射，作为参考光束未发生频移直接入射到光电检测器上。透过半反射镜 M₁ 的光束被颗粒散射，其中部分散射光载荷多普勒频移投射到探测器上。两束光经混频得到差频信号。为保证两束光强度值匹配，参考光常用中性滤光片 M₃ 调节。

2) 对称互差方式

在图 b 的光路中，入射激光经透镜聚焦于被测颗粒 P 上。两束对称的散射光载荷两个对称方向的多普勒频移分别经反射镜 M₁ 和 M₂ 由光探测器外差接收。差频信号是由两束受不同方向多普勒运动调频的光波形成的，故称作对称互差式。

3) 干涉条纹方式

图 c 的光路中激光束经分束镜后分成两束等光强平行光，由透镜会聚在测量场中。光束交叉的区域构成检测区，形成干涉条纹。条纹的形状是一组平行于入射光束角平分线的直线组，间距 D 是入射光束夹角 θ 的函数

$$D = \lambda / 2 \sin(\theta / 2)$$

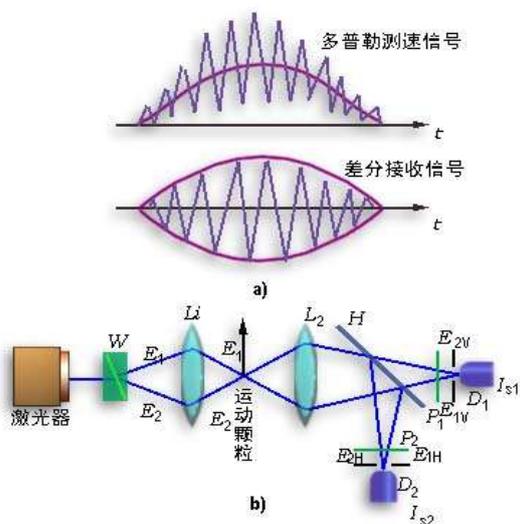
当运动颗粒与条纹法向 n 成 α 角通过检测区时，条纹图形将产生周期性变化，光检测器输出信号的频率 f_d 与被测运动速度 v 成正比。利用上式有

$$f_d = (v / D) \cos \alpha = (2v / \lambda) \sin(\theta / 2) \cos \alpha$$

以上述三种方式为基础，可组成许多实用的多普勒测速计。

(2) 调制信号的解调和处理

1) 多普勒调频信号的接收



差分检测的多普勒信号接收系统

a) 多普勒测速信号波形 b) 原理示意图

多普勒测速信号是一个随被测颗粒进入光束照明区而断续出现的夹有激光相干噪声的调幅调频波，有相当宽的频谱分布。它的幅度调制反映了照明激光束的径向光强高斯分布，频率调制的特征反映了被测颗粒速度的变化（如图 a 所示）。为了消除直流光强分量、高频噪声和钟形调幅包络线的影响，在只用单一接收检测器的情况下需要进行高频和低频滤波。在许多场合常采用差分接收。

图 b 给出了用差分检测的多普勒信号接收系统示意图。图中，线偏振激光经渥拉斯顿棱镜 W 分光成两束 $\pm 45^\circ$ 的偏振光，此二束光经物镜 L₁ 后交叉会聚构成检测区。被测散射颗粒的运动使检测区内两束偏振光 E₁ 和 E₂ 的角频率分别发生 $\Delta \omega_1$ 和 $\Delta \omega_2$ 的变化，同时杂乱地改变了它们的偏振方向。随后，出射光被半

反半透镜 H 分束后，分别通过检偏器 P₁ 和 P₂，检偏器的偏振面分别与发自光源的偏振光垂直和平行。两束调频光分别由探测器 D₁ 和 D₂ 接收。

设进入到 D₁ 的两束垂直偏振的光波 E_{1V} 和 E_{2V} 及进入到 D₂ 的两束平行偏振的光波 E_{1H} 和 E_{2H} 的复振幅分别为

$$E_{1V} = E_{1H} = a_1(t) \exp[-j(\omega_0 + \Delta\omega_1)t]$$

$$E_{2V} = E_{2H} = a_2(t) \exp[-j(\omega_0 + \Delta\omega_2)t]$$

式中， $\Delta\omega_1 = k_1 v$ 和 $\Delta\omega_2 = k_2 v$ 分别为两束入射光的多普勒角频移， k_1 和 k_2 为比例因子。

同一方向偏振光在光电检测器上混频得到外差输出信号，分别为

$$I_{s1} = K[a_1^2(t) + a_2^2(t) + 2a_1(t)a_2(t)\cos(\Delta\omega_d t)]$$

$$I_{s2} = K[a_1^2(t) + a_2^2(t) - 2a_1(t)a_2(t)\cos(\Delta\omega_d t)]$$

式中，K 是光电检测器的灵敏度， $\Delta\omega_d$ 是两束光多普勒运动角频移的差值，表示为

$$\Delta\omega_d = \Delta\omega_1 - \Delta\omega_2 = (k_1 - k_2)v$$

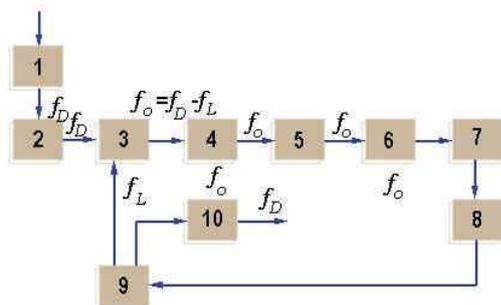
将二个输出信号 I_{s1} 和 I_{s2} 差分放大，并考虑到两束正交偏振光的极性，差分放大输出电流 ΔI_s 为

$$\Delta I_s = I_{s1} - I_{s2} = 4Ka_1(t)a_2(t)\cos(\Delta\omega_d t)$$

上式表明差分接收方式能消除信号直流分量的影响，并使交流信号振幅增大一倍。对应的信号波形表示在图 a 中。可以看出通过差分接收使复杂的信号波形得到简化，同时保持了原有的频率特征。

2) 频率跟踪器

根据前面的公式，一旦检测出多普勒差值 Δf ，便可计算到被测流速。通常利用下列方法进行信号的频率分析，即：多通道或扫描滤波型频谱分析仪，信号周期测量或波形计数，相关测量。最典型的方法是频率跟踪器。



利用频率跟踪器检测多普勒信号电子系统示意图

1-光检测器 2-宽带放大器 3-电子混频器 4-中频滤波器 5-限幅器

6-鉴频器 7-积分器 8-直流放大器 9-压控振荡器 10-显示单元

上图给出了利用频率跟踪器检测和处理多普勒信号的电子系统示意图。受频率调制的光波在检测器 1 中进行光混频后，拍频信号经前置宽带放大器 2 形成有一定强度的多普勒信号 f_D 。在混频器 3 中和来自压控振荡器 9 的本机振荡信号 f_L 混频，得到的外差信号 $f_D - f_L$ 经中频滤波器 4 选频，消除噪声并取其差频送到限幅器 5 以去掉幅度调制和消除放大倍率的波动影响。限幅器的方波输出进入鉴频器 6，它和中频滤波器均调谐在相同的频率 f_0 上。鉴频器的直流输出信号与混频输出 $f_D - f_L$ 和鉴频器中心频率 f_0 的差频成比例。此直流信号用积分器 7 加以滤波，通过直流放大器 8 反馈控制压控振荡器，形成频率跟踪闭环回路。当系统平衡时，有关系

$$f_D - f_L = f_0$$

或

$$f_D = f_L + f_0$$

上式表示当实现频率跟踪后压控振荡器的输出频率 f_L 被锁定在多普勒信号上，相差一个固定的 f_0 值，测定该频率值 f_L ，利用数字显示单元 10 即可计算和显示出被测流体的速度平均值或速度的变化。

频率跟踪器仅接收中心频率在多普勒信号上一定带宽内的噪声，故能较好地抑制背景噪声。此外对多普

勒频率的调制深度要求不高，在信噪比较低时也能工作，可在较大的范围内（如 15MHz）对变化的速度进行跟踪，便于实时显示。

2. 萨纳克效应和转动差频

当封闭的光路相对于惯性空间有一转动角速度 Ω 时，顺时针光路和逆时针光路之间将形成与转速成正比的光程差 ΔL ，其数值满足关系

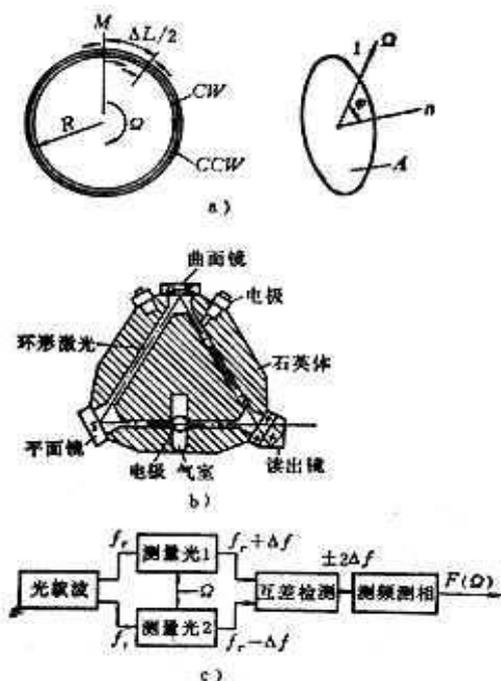
$$\Delta L = (4A/c) \Omega \cos \phi$$

式中， c 为光速， A 为封闭光路包围的面积， ϕ 为角速度矢量与面积 A 的法线间的夹角。

当光路平面垂直于 Ω 时，上式简化为

$$\Delta L = (4A/c) \Omega$$

这种闭合光路的反向光路光程差随转速改变的现象称作萨纳克效应。下图 a 给出这一效应的图解说明，可以看出当光路以 Ω 顺时针转动时，从光路上一点 M 发出的顺时针光束 CW 在绕光路一周重新回到 M 点时要多走一段光程，而反时针光束 CCW 却少走一段光程，于是形成了光程差。这种光程差的量值甚微，例如采用 $A=100\text{cm}^2$ 的环形光路对地球自转的速度为 $\Omega_E=7.3 \times 10^{-5}\text{rad/s}$ ，相应的 ΔL 仅为 10^{-12}cm 。只有利用环形干涉仪或环形激光器才有可能通过检测双向光路的微小频差得到这一角速度。



萨纳克效应的转动光程差和环形激光器

a) 萨纳克效应 b) 激光陀螺 c) 方框图

三个或三个以上反射镜组成的激光谐振腔使光路转折形成闭合环路。这种激光器称作环形激光器（图 b）。在环形激光器中，激光束的基频纵模频率 ν_{00q} 可表示为

$$\nu_{00q} = q \cdot c/L$$

式中， c 为光速， L 为腔长， q 为正整数。

此式表明激光谐振腔长 L 和光频 ν 之间有比例关系，即

$$\Delta \nu / \nu = \Delta L/L$$

式中， $\Delta \nu$ 是与光程差 ΔL 对应的光频差即 Δf ，可得

$$\Delta f = (4A/\lambda L) \Omega$$

此式即为环形激光器的测角速度公式。为了计算实际转角 θ ，可对光频差计数累加积分，其波数值 N 即为

$$N = \int_0^t \Delta f dt = (4A/\lambda L) \theta$$

这就是环形激光器的测角公式。

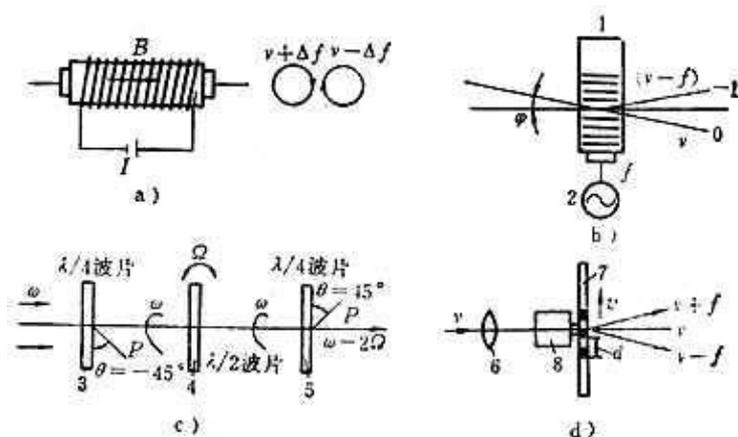
小型化的环形激光器及相应的光学差频检测装置组成了激光陀螺。它可以感知相对惯性空间的转动，在惯性导航中作为光学陀螺仪使用。此外，作为一种测角装置，它是一种以物理定律为基准的客观角度基准，有很高的测角分辨率，在 360° 角度范围内有 $0.05'' \sim 0.1''$ 的测量精度。

转动效应在相反方向光路中分别引起的光频率的增减是一种非互易效应。类似的现象也发生在电场和磁场对环形激光的相互作用。由于非互易效应引起的光学频差是双向的，因此，在这种干涉测量中，参考波同样受到调制而不再保持恒定，光频差是二束测量光波的混频形成的，它的信息调制过程表示在图 c 中，这种外差检测方式称作互差式。

10.6 固定频移的频率调制和外差检测

1. 固定频移的形成

固定频移可以用光学或光电子器件产生。



几种常用的器件偏频法

a) 塞曼效应 b) 声光效应 c) 旋转波片 d) 旋转光栅

1-声光器件 2-激励换能器 3、5- $\lambda/4$ 波片

4-半波片 6-透镜 7-光栅盘 8-电动机

(1) 塞曼效应激光频移

如图 a 所示，利用永久磁铁或螺线管在 He-Ne 激光器中形成轴向磁场，它使单模激光分裂成左右圆偏振的两个分量。二偏振光间存在频差，数值取决于外加磁场的强度和谐振腔品质因数。通常，几个高斯的磁场强度可得到 2MHz 到 300kHz 的频差。

(2) 声光效应激光频移

如图 b 所示，在声光器件 1 中以频率为 f 的超声波交变信号激励换能器 2，在透明介质内形成折射率的周期变化。当激光平行入射到介质内将产生零级和 ± 1 级衍射光，一级衍射光与零级衍射光频率相差 $\pm f$ ，可分别作为参考光和信号光。声光偏频所需的控制功率较低，频差可达 100MHz，变换效率为 80%。

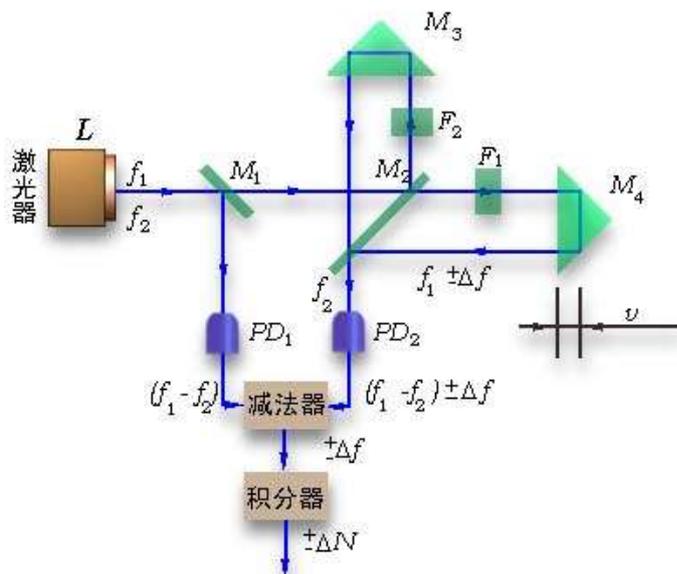
(3) 旋转波片激光频移

如图 c 所示，线偏振激光通过 $\lambda/4$ 波片 3 后输出圆偏振光。再通过以 Ω 转动角频率旋转的半波片 4 和固定的 $\lambda/4$ 波片 5，所输出的偏振光可得到 2Ω 的角频移。半波片转速由电动机控制，频差受限在 2~3kHz 以下，变换效率为 90% 以上。

(4) 旋转光栅激光频移

如图 d 所示，激光由透镜 6 聚焦在光栅盘 7 的刻线上，透射光被分作零级和 ±1 级衍射。光栅盘由电动机 8 带动旋转。若光栅上光点处的线速度为 v ，光栅刻线间距为 d ，则 ±1 级衍射光发生 $\pm f = v/d$ 的频移。频移稳定性与转速有关。频差可达 20MHz。变换效率一般为 20%。

2. 双频干涉和外差平面干涉法



双频激光干涉仪原理示意图

利用固定频移进行外差检测的典型应用是双频激光干涉仪。它的原理示意图表示在上图中。图中，双频激光装置 L 产生频率相差几兆赫兹的二种频率的激光 f_1 和 f_2 ，它们在基准光束分光镜 M_1 上分作二束。其中反射光中的 f_1 和 f_2 在光电检测器 PD_1 上混频得到二光频的差频信号作为参考信号。透射光受干涉反射镜 M_2 反射，经光学滤光器 F_2 得到频率为 f_2 的单频激光，它由参考用角反射镜 M_3 反射后成为干涉仪的参考光束。透过 M_2 的光束经光学滤光器 F_1 后得到频率为 f_1 的单频激光，经测量用角反射镜 M_4 的反射，附加了镜面运动引起的多普勒频移 Δf ，以 $f_1 \pm \Delta f$ 的光频在光电检测器 PD_2 中和参考光频 f_2 相混频，得到光学差频信号 $f_2 - (f_1 \pm \Delta f) = (f_2 - f_1) \pm \Delta f$ 。这相当于多普勒频移 Δf 对光学差频 $(f_2 - f_1)$ 的频率调制。由于光学差频 $(f_2 - f_1)$ 频率已进入到电信号处理的通频带内，因此，将 PD_1 和 PD_2 中检测到的两路外差信号经过电信号混频或作频率计数相减运算，即可以得到表征物体运动速度的光学差频信号 Δf ，并有

$$\Delta f = \pm (2/\lambda) v$$

若用积分器累加差频信号的相位变化或者对差频信号的波数 N 计数，可得

$$N = \int_0^t \Delta f dt = (2v/\lambda) dt = (2/\lambda) \int_0^t v dt$$

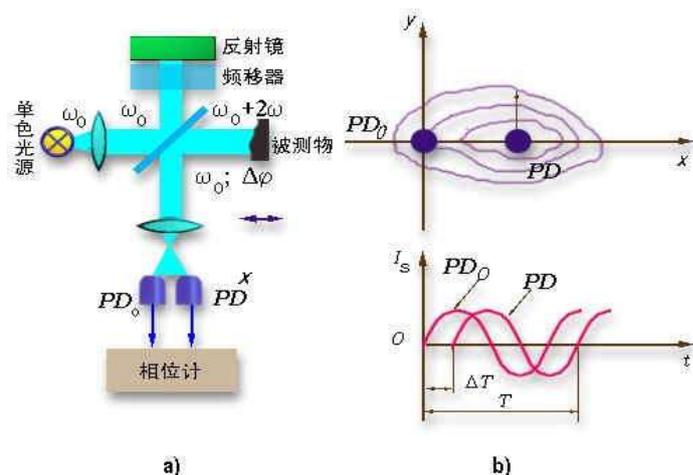
$$N = (2/\lambda) L$$

式中， $L = \int_0^t v dt$ 为运动物体的位移，于是有

$$L = (\lambda/2) N$$

这就是双频干涉测量装置的测量公式。该系统的优点在于整个系统中的信号是在固定频率偏差 $f_2 - f_1$ 的状态下工作的，这就克服了常规干涉仪中采用直流零频系统所固有的复杂通道耦合、长期工作漂移等不稳定因素，提高了测量精度和对环境条件的适应能力。通常，在频差为 10~50MHz，激光稳频度为 10^{-8} 时，能得到小于 $0.1 \mu m$ 的灵敏度和 $1 \mu m$ 的测长精度。

双频干涉仪的信息流程表示在前面的图 b 中。



平面外差干涉法原理示意图
a) 原理示意图 b) 波形图

利用固定频率偏移的方法使被测的反射波面各点处分别形成不同的差频信号，通过外差检测可以得到二维波面的相位分布图。上图表示利用这种原理的平面外差干涉法示意图，这是一种改进的泰曼-格林干涉仪。图中的干涉条纹是由两束光频率稍许不同的光束形成的。由激光器产生的单频光在参考光路中经过频移器，从原来的角频率 $\omega_0 = 2\pi f_0$ 偏移了 $2\omega = 2(2\pi \Delta f)$ 角频率。此时干涉图样上任意点 x 处的光波复振幅可表示为

$$U(x, t) = a_0(x) \exp\{j[(\omega_0 t + \phi_0(x))]\} + a_r \exp\{j[(\omega_0 + 2\omega)t + \phi_r]\}$$

式中， $a_0(x)$ 和 $\phi_0(x)$ 是测量光束各点处的光波振幅和相位； a_r 和 ϕ_r 是参考光束的振幅和相位，它们在相干平面上是均匀的。

在干涉平面上各点处的光强度为

$$I(x, t) = |U(x, t)|^2 = a_0^2(x) + a_r^2 + 2a_0(x)a_r \cos[2\omega t + \phi_r - \phi_0(x)]$$

式中， $2\omega = 4\pi \Delta f$ ， Δf 为光频差。

此式表明：干涉平面上各点处的合成光是以测量光和参考光的光频差 $2\Delta f$ 为变化频率，随时间按正弦规律变化的。其中 $\phi_r - \phi_0(x)$ 是二光束在没有频偏时稳定干涉条纹的空间相位，而在外差干涉仪中则转换为时序信号的相位差。这样，若在 x 点处放置平方律检测器就能将光的波动转换为交变的电信号。它们具有 $2\Delta f$ 的频率和 $\phi_r - \phi_0(x)$ 的初始相位。为了比较不同位置上的相位差，可选取 $x = x_0$ 处为相位基准点，利用两个光电检测器同时测量差频信号（如图 b 所示）。设被测信号相对参考信号的时间延迟为 ΔT ，信号周期为 T ，则相位差为

$$\phi_0(x) - \phi_0(x_0) = 2\pi \Delta T / T$$

使光电检测器扫描整个干涉场，测量出各点处的相位即得到被测波面的相位分布 $\phi_0(x)$ ，进而能换算出被测物体的表面形状。

平面外差干涉法具有很高的相位测量精度（ $\lambda/100 \sim \lambda/1000$ ）和空间分辨率（100 线对/mm）。特别重要的是外差干涉法在原理上不是取决于二相干光束的强度而只是利用它们的相位关系。因此，即使相干光强有时间和空间的变比也不会影响测量的结果。此外，频移装置造成的光频偏移的波动对二束光的影响是相同的，不致引起相对相位的变化，这就为高精度的测量提供了可靠的保证。外差干涉法在高质量光学元件的检查、干涉显微镜的测量，以及利用波面相位进行测量的能动光学系统中都获得了有成效的应用，它作为干涉测量技术的重大突破，是现代光电测量技术的重要方向。

10.7 直接光频调制和外差检测

1. 直接光频调制干涉测量原理

具有固定频移的外差检测要求有辅助的频移器或者能产生双频光的激光器以及附属的控制装置，这不仅增加了系统的复杂性，而且偏频的稳定性会直接影响测量精度的提高。

近年来，小型、坚固、低功耗的半导体激光器（LD）作为新型的相干光源得到了迅速的发展。它具有良好的工作特性，当注入电流改变时，激光器的振荡频率（或波长）能直接变化，因而可以实现光波频率随注入电流的直接调制，称作直接光频调制。利用这一性质已研制出小型、宽带、高速、高精度的波面形状、位移、距离和光频变化的测量装置。

(1) 半导体激光器的直接频率调制

在半导体激光器中，谐振腔应满足谐振条件

$$\lambda = (2/m) nL$$

式中， λ 是激光波长， n 为折射率， L 为谐振腔长， m 为整数。

当注入电流变化 Δi 时，引起谐振腔温度变化 ΔT 为

$$\Delta T = R \Delta i$$

式中， R 是谐振腔热阻。

ΔT 的变化引起腔长 L 和折射率 n 的改变，由前式知最终引起激光波长的改变 $\Delta \lambda$ 为

$$\Delta \lambda = (2/m) (L \frac{\partial n}{\partial T} + n \frac{\partial L}{\partial T}) \Delta T$$

其中， $L/T \ll n/T$ ，利用式 $\Delta T = R \Delta i$ 上式简化为

$$\Delta \lambda = (\lambda/n) \frac{\partial n}{\partial T} R \Delta i = \beta \Delta i$$

$$\beta = (\lambda/n) \frac{\partial n}{\partial T} R$$

此处 $\beta = \Delta \lambda / \Delta i$ 称作电流波长调制系数。对于 $\lambda = 780\text{nm}$ 的 GaAlAs 半导体激光器，有 $n = 3.5$ ， $\frac{\partial n}{\partial T} = 4 \times 10^{-4}/^\circ\text{C}$ ， $R \approx 0.07^\circ\text{C}/\text{mA}$ ，相应的 $\beta \approx 6 \times 10^{-3} \mu\text{m}/\text{A}$ 。

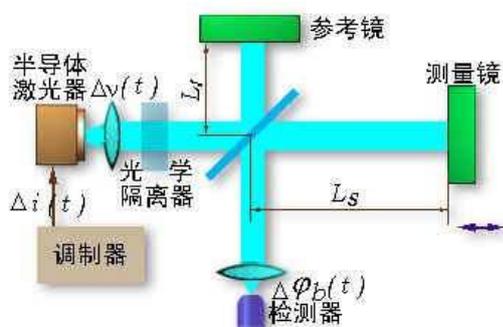
由波长和频率的关系可将上式变换为光频差表达式

$$\Delta \nu = -(c/\lambda^2) \Delta \lambda = -(c/\lambda^2) \beta \Delta i = K_m \Delta i$$

$$K_m = -(c/\lambda^2) \beta$$

式中， $K_m = \Delta \nu / \Delta i$ 称作电流频率调制系数。对于 $\lambda = 780\text{nm}$ ， $\Delta \nu$ 在 1MHz 以下的低频范围， $K_m = (1 \sim 7) \times 10^3 \text{GHz}/\text{A}$ 。连续的光频率变化被限制在一个跳模间隔内，最大变频范围为 $\Delta \nu_{\text{max}} = 50 \sim 70\text{GHz}$ 。

(2) 调频光干涉测量



迈克尔逊调频光干涉测量原理图

上图给出了迈克尔逊调频光干涉仪的基本组成和工作原理。由 LD 激光器产生的单模激光波长为 λ_0 ，频率 ν_0 。通过物镜准直后经光学隔离器引入干涉仪中。设参考光路长度为 L_r ，被测光路长度为 L_s ，光程差

为L, 则二光波相位差 ϕ_0 为

$$\phi_0 = 2\pi L / \lambda_0 = 2\pi v_0 L / c = 2m\pi + \phi$$

式中, $L = L_r - L_s$, ϕ 为半波长以下小位移时对应的相位角, 被测量 ΔF 引起的相位角变化 $\Delta \phi$ 为

$$\Delta \phi = (2\pi/c) (v_0 L \frac{\partial n}{\partial F} \Delta F + nL \frac{\partial L}{\partial T} \Delta T + n v_0 \frac{\partial L}{\partial F} \Delta F)$$

式中第二项表示了光频改变 Δv_0 对相位 $\Delta \phi$ 的影响。这样当保持 LD 的温度不变, 注入电流改变 $\Delta i(t)$, 光频变化 $\Delta v(t)$ 时, 引起的二相干光的附加相位偏移 $\Delta \phi_b(t)$ 为

$$\Delta \phi_b(t) = 2\pi n \Delta v(t) L / c$$

此时光电检测器的光电流为

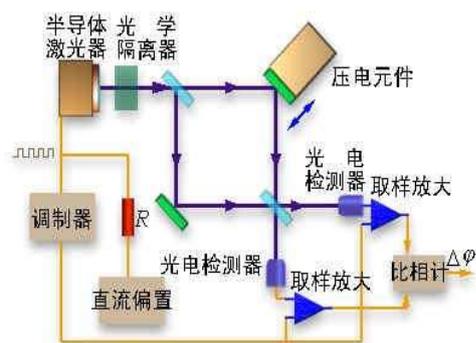
$$I_s(t) = I_0 + I_m(t) = I_0 \{1 + \gamma \cos[\phi + \Delta \phi + \Delta \phi_b(t)]\}$$

式中, I_0 为信号直流分量, γ 为交流分量幅值 I_m 与直流分量 I_0 比值 ($\gamma = I_m / I_0$)。由上两式可以看出, 直接调频法使合成光强以及相应的光电信号的相位随光频变化的规律进行调制, 与相幅变换相比可称作是频相变换。此时即使被测波面不随时间改变, 干涉信号也将随时间改变。这样, 只要测量出时间信号的相位值即可由 $\Delta \phi$ 式解出被测量来。

2. 直接光频率调制的信号解调法

为了利用 $I_s(t)$ 式从直接调频光电信号中解调出被测相位角 $\phi + \Delta \phi$, 有许多类型的解调方法。

(1) 双频切换干涉法



双频切换干涉法示意图

上图是采用马赫-泽德干涉仪的双频切换干涉法示意图。在波长为 857nm 的 GaAlAs 激光器中注入方波电流对激光器作时间调制, 使照明激光频率交替改变, 附加相位移 $\Delta \phi_b(t)$ 周期性地变为 0 或 $\pi/2$ 。这时, 由 $I_s(t)$ 式的交变分量中得到与 $\cos(\phi + \Delta \phi)$ 和 $\sin(\phi + \Delta \phi)$ 成比例的输出电流 I_{mc} 和 I_{ms} 。二个取样放大器分别取样出光电检测器的输出电流, 利用相位比较器取它们的比值, 可计算出被测相位 $\phi + \Delta \phi$ 为

$$\begin{aligned} \phi + \Delta \phi &= \arctg[\cos(\phi + \Delta \phi) / \sin(\phi + \Delta \phi)] \\ &= \arctg(I_{ms} / I_{mc}) \end{aligned}$$

该系统可检测反射镜的振动 (图中由压电晶体驱动) 和位移。对 1kHz 的振动, 振动相位的测量灵敏度为 $5 \times 10^{-5} \text{ rad/Hz}^{1/2}$ 。检测电路带宽 1Hz 时, 位移测量灵敏度为 7pm。

(2) 线性扫描调频干涉法

在泰曼-格林型干涉仪中, 使半导体激光器注入电流随时间成比例变化, 有

$$\Delta i = \alpha t$$

代入前面的式子得到相干光的调制相位 $\Delta \phi_L$

$$\Delta \phi_L(t) = (2\pi nL/c) K_m \alpha t = K_L t$$

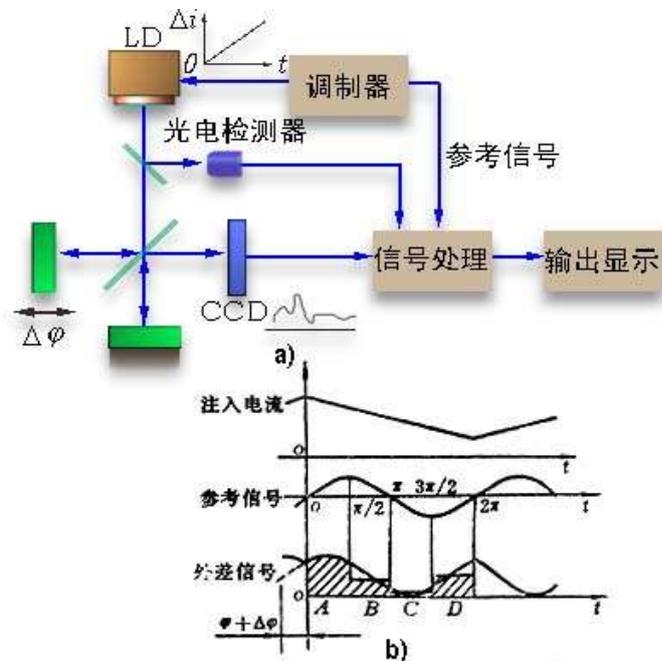
式中, $K_L = (2\pi nL/c) K_m \alpha$ 为比例常数。

将上式代入 $I_s(t)$ 式有

$$I_s(t) = I_0 [1 + \gamma \cos(\phi + \Delta \phi + K_L t)]$$

式中, $K_L=2\pi f_L$ 。

这表明输出光电流按正弦规律变化。调制信号的频率为 $f_L=(nL/c)K_m\alpha$ 。若检测到相干平面各点处电信号的初始相位即可确定波面的相位分布。



线性扫描调频干涉

a) 原理示意图 b) 波形图

上图给出利用线性扫描调频法的干涉仪示意图。采用三角形波的注入电流（见图 b），使干涉条纹本身进行周期性地扫描运动。伴随条纹的移动，面阵 CCD 摄像装置各个象素上的光电流输出也周期性地满足上式，波形如图 b 所示。控制注入电流的大小可使干涉条纹只在一个周期内移动。波形的折返点表示被测点上的初始相位 $\phi+\Delta\phi$ 。为了测定 $\phi+\Delta\phi$ 值，可以利用如图 b 所示的四段积分法（或称四斗式）。该方法是计算在 $\Delta\phi_L(t)=K_Lt$ 分别处于 $0\sim\pi/2$ 、 $\pi/2\sim\pi$ 、 $\pi\sim3\pi/2$ 、 $3\pi/2\sim2\pi$ 的四个区间内检测器输出电流的积分值，然后即可计算出被测相位 $\phi+\Delta\phi$ 值为

$$\phi+\Delta\phi=\arctg[(A-C)/(B-D)]$$

由于半导体激光器直接频率调制可达 10^9MHz 数量级，所以测量时间可以很短（50ms），有利于进行高速测量，可避免温度漂移和振动的影响。它的测量精度可达 $\lambda/50$ 。

2. 利用调频光的距离测量

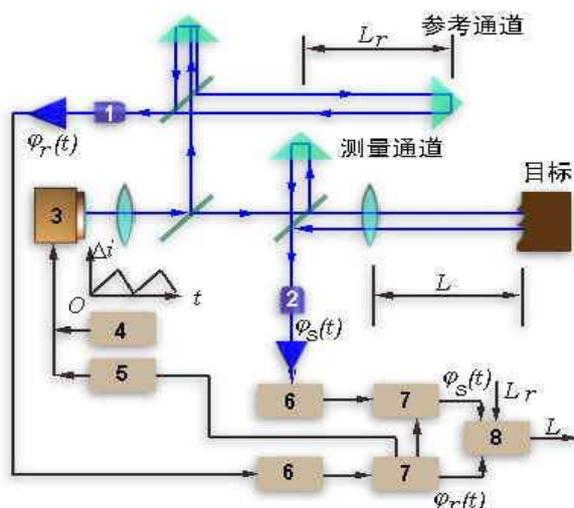
在迈克尔逊型双光束干涉仪中，相干光波的相位差 $\Delta\phi_0$ 和光程差 ΔL 之间有关系

$$\Delta\phi_0=2\pi\Delta L/\lambda_0=2\pi n v_0\Delta L/c$$

式中, $\Delta L=2(L_0-L_r)$, L_r 为参考光路长, L_0 为被测光路长或被测距离。

在前述讨论中所谓位移测量指的是当 ΔL 发生微小变化 δL 时, 测量出相位变化 $\delta\phi$, 利用公式 $\delta\phi=2\pi n v_0\delta L/c$ 计算 δL 的过程。而在这里的距离测量表示测最被测光路的绝对长度 L 。在许多利用调频光测距的方法中我们讨论二种典型的方法。

(1) 参考相位比较法



有参考通道的调频光测距仪

- 1-参考通道光电检测器 2-测量通道光电检测器 3-半导体激光器
 4-直流电源 5-调制器 6-带通滤波器 7-相位计 8-比例计

上图是具有参考通道的调频光测距干涉仪的原理图。注入波形为三角形的电流对单模半导体激光器调频，使之作为两个干涉仪的光源，并由测量用干涉通道得到准连续的正弦拍频信号。测得它的相位移为 $\Delta \phi_s(t)$ ，由前面的式子有

$$\Delta \phi_s(t) = 2 \pi n \Delta \nu L / c$$

$$L = \Delta \phi_s(t) c / 2 \pi n \Delta \nu$$

导入参考干涉通道的调频光形成参考用的正弦拍频信号，其初始相位差为 $\Delta \phi_r(t)$ 。设参考光程为 L_r ，则由 $\Delta \phi_r(t)$ 式可计算出调频光的频偏 $\Delta \nu$

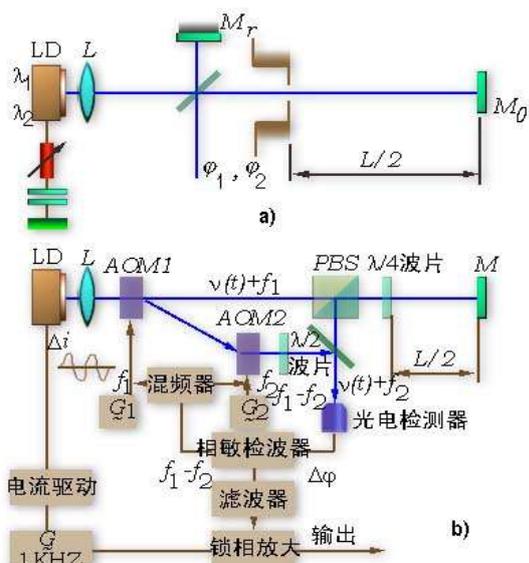
$$\Delta \nu = c \Delta \phi_r(t) / 2 \pi n L_r$$

将该式代入前式有

$$L = L_r \cdot \Delta \phi_s(t) / \Delta \phi_r(t)$$

由上式可见，同时测得二个干涉通道的相位移 $\Delta \phi_s(t)$ 和 $\Delta \phi_r(t)$ ，即可测得被测光程长或距离。这种方式由于采用比较测量，消除了频率调制的非线性和不稳定性的影响，测量精度很高。使用波长 780nm，功率为 15mW 的单模 GaAlAs 激光器，调制电流 $\Delta i = 16\text{mA}$ ，调制频率为 90Hz，光频最大偏移 $\Delta \nu = 40\text{GHz}$ ，在 0.2~1.7m 内相对测距精度为 9×10^{-6} （约 6~7 μm ）。

(2) 调频光多波长干涉法



调频光合成波长法测距

a) 多波长干涉测距 b) 结构示意图

对于长光路的距离测量，有时采用多波长干涉法，通过对不同波长相位差的比较来测定光路的绝对长度。在上图 a 的简单光路中，调整半导体激光器的注入电流使波长分别为 λ_1 和 $\lambda_2 = \lambda_1 + \Delta \lambda$ ，相干光的相位差 ϕ_1 和 ϕ_2 分别为

$$\phi_1 = 2 \pi nL / \lambda_1$$

$$\phi_2 = 2 \pi nL / (\lambda_1 + \Delta \lambda)$$

两相位差的差值 $\Delta \phi$ 在 $\lambda_1 \approx \lambda_2$ ，即 $\Delta \lambda$ 很小时有

$$\Delta \phi = \phi_2 - \phi_1 = 2 \pi nL [1 / (\lambda_1 + \Delta \lambda) - 1 / \lambda_1] \approx 2 \pi nL \Delta \lambda / \lambda^2$$

因此，当 λ 、 $\Delta \lambda$ 已知并测得 $\Delta \phi$ 后即可计算出 L 值，使用双波长时， $\Delta \lambda$ 的确定应确保 $\Delta \phi$ 在 2π 范围内。但在波长连续变化的情况下即使 $\Delta \phi$ 超过 2π 也能得到准确结果。

图 b 为利用多波长干涉原理的测距系统图。波长为 789nm，光功率为 20mW 的半导体激光器 LD 由声光调制器 AOM1 分束移频构成外差干涉，其中一束经偏振分光镜 PBS 由测距靶镜 M 反射，最后由光电检测器接收。另一光束经声光调制器 AOM2 再次移频，由半透半反镜 HM 导入光电检测器中。配置适当的偏光板可得到 450kHz 的外差信号。以 $f_m = 1\text{kHz}$ 正弦波电流调制 LD 激光波长时，根据 $\Delta \phi_b(t)$ 式，外差信号的相位也周期性地按正弦规律变化。采用相敏检波器检测被测相位，参考信号由驱动 AOM 的二个正弦信号经电信号外差混频得到。相敏检波器的正弦输出信号表示外差信号相位的正弦变化规律，由前面的式子有

$$\Delta \phi = (2 \pi nL / \lambda^2) \beta \alpha \sin 2 \pi f_m t$$

式中， $\alpha \sin 2 \pi f_m t = \Delta i$ 为激光器调制电流。

上式表明该相位变化的频率与注入电流调制频率相同，振幅与被测光程差（距离）成正比。利用锁相放大检测振幅值即可以计算出被测距离。这种方法的测距范围在 n 个毫米范围内，相位测量灵敏度约 10^{-5}rad ，测距灵敏度为 $0.1 \mu\text{m}$ 。