

光学讲义

陈曦

August 15, 2020

目录

1	波动光学	iii
1.1	光的产生	iii
1.2	相干性	iv
1.3	干涉	iv
1.4	衍射	v
1.5	偏振	ix
2	光学成像	xi
2.1	几何光学基本方程	xi
2.2	几何光学成像	xiii
2.3	基于折射元件的几何光学成像	xiv
2.4	成像仪器	xvii
2.5	分辨率极限	xix
2.6	傅立叶光学	xix

1 波动光学

最简单的一种波是平面波。平面光波的一个电场分量记为

$$E(\vec{r}, t) = Ae^{i(\omega t - \vec{k} \cdot \vec{r} - \phi)} \quad (1.1)$$

上式的实部是真实的物理的场。 A 是实的量。真空中光强定义为

$$I(\vec{r}, t) = E^*(\vec{r}, t)E(\vec{r}, t) \quad (1.2)$$

对于平面波 $I = A^2$ 。

一个波的波前是指相位相同的点构成的面。波的传播方向垂直于波面。

我们在此将讨论的光波特指波长远大于原子尺度又远小于常用光学元件的尺寸（厘米量级）的电磁波，涵盖从红外到紫外的波段。用于处理光波的光学元件通常是基于折射和反射效应。

1.1 光的产生

一般地，光源分为热光源和激光两类。

热光源的辐射功率与发光体的热力学状态直接相关。光子产生于作为光源的一个宏观多体系统从其激发态到基态的跃迁。每个光子的产生都是一个近乎独立的事件，因此不同的光子之间缺乏关联。系统的能级结构决定了有哪些频率的光能出现在发光光谱中。系统处于各激发态的几率决定了不同频率的光的发光强度。对于热平衡态，这个几率由温度决定。例如黑体辐射产生于一个有连续能级机构的系统。气体中的分子有分立的能级结构。对于较稀薄的气体各分子近乎独立地辐射光子，因此光谱也呈现分立结构。由于能级的寿命、多普勒效应、分子间的碰撞，谱线会有展宽。发光二极管（LED）已经变得很普及。其辐射的光子来源于电子从半导体的导带向价带的跃迁。根据半导体材料的不同，光谱的范围也不同，通常有几十个nm的展宽。

激光由谐振腔和增益介质组成。谐振腔中只有某些频率的光可以产生共振，在腔体中建立起来电磁场的驻波。由于腔体两端的反射镜损耗很低，所以共振谱线的展宽很小，进而有很好的相干性。少部分能量可以通过反射镜到达腔外形成高度相干、高度准直的激光。为了使这种驻波能够形成并维持下去，腔中需要有增益介质。该介质由外界输入能量进行激发。如果介质激发态向基态跃迁的能量与驻波的相同，则处于激发态的介质通过所谓受激辐射过程与驻波产生共振。受激辐射保证了介质发射的光子与驻波保持相同的相位，进而向驻波中注入能量。如果介质的发光缺乏与驻波的相干性，驻波就会很快被衰减。发生在激光器中的过程有些类似于在一根两端固定的弦。在弦中建立起驻波后，如果在弦上的一些位置施加同样频率外力，并且每个外力的相位都与这个力施加位置处弦的相位相同，则会有能量注入到弦中。如果外力缺乏与弦的相干性，驻波很快会被破坏。

1.2 相干性

光由光子组成。由于激发态的有限寿命，发出的光子有一定频率展宽，这样每个光子可以看作是由原子发出一个波包。对于热光源，两个原子间没有关联，因此两个波包间没有相位关系。对于波包间没有确定相位关系的情形，测量空间上两点间光波的相位关系就会发现如果两点间的距离小于波包的长度时，这两点间的相位会有一些的关联，称为具有空间相干性。相干长度大致是波包的长度。在一点上进行相位测量时，如果两次测量的时间差小于相干长度除以光速，则这两个时间上的相位间会有一些关联，称为具有时间相干性。激光的两个光子间具有确定的相位关系，因此激光具有很长的相干长度。在这个意义下，激光是很好的相干光源。热光源通常相干长度很短。用一个小孔可增加相干性，因为经过小孔衍射后的光波来源于波前上的一个位置，具有相位的关联。单色仪也可以增加相干性。一个波包经过单色仪后会被滤掉大部分的频率成份，波包更接近一个平面波。尽管热光源的相干性可以无限制地提高，但光强会被损失。

相干性表示的是场在空间和时间上的关联。将两个光场进行叠加，电场是简单地相加，于是空间上某一点总的光强是

$$\begin{aligned} & (Ae^{i\phi_1(t)} + Be^{i\phi_2(t)}) \cdot (Ae^{i\phi_1(t)} + Be^{i\phi_2(t)})^* \\ & = A^2 + B^2 + 2ABe^{i(\phi_1(t) - \phi_2(t))} \end{aligned} \quad (1.3)$$

如果两个光场不相干，即没有固定的相位差，则第三项会以光的频率在振荡。一般地，探测器的响应时间远长于电场的振荡周期。所以在光探测器看来第三项只能探测到其平均值。对于无规的相位差 $\phi_1 - \phi_2$ ，这个平均值为0。所以，两束光非相干叠加，总的强度是两束光的强度之和。如果 $\phi_1 - \phi_2$ 有一个确定值，两束光进行相干叠加，第三项对总的光强有重要贡献，称为干涉项。

1.3 干涉

光作为电磁波，我们通常只考虑其电场的成份，原因是电磁波中电场与物质的相互作用往往比磁场强。电场是矢量，可以与物质相互作用，可以叠加，以波长为空间分布的特征长度。这些性质导致了光的一系列现象。例如，相干的两束光相遇时的干涉效应。产生干涉效应有两种常用的方法。方法一，在空间中的两个位置上做两个狭缝，狭缝间距在相干长度以内。这样从两个狭缝出射的光会有一些相干性，在屏幕上会形成稳定的干涉条纹。这种方法称为分波前干涉。方法二，用半反射镜将一束光分为两束。两束光各经过一定路径后相遇。如果两束光的传播时间之差短于相干时间，则随着光程差的变化叠加后的光会出现强度振荡。这种方法称为分振幅干涉。

1.4 衍射

在不考虑光的波动性时光经过一物体后会投下一个边界清晰的影区。但事实上光的波动性会使边界不再清晰。这个现象称为衍射。严格的对电场的描述可以通过解波动方程得到。在标量场近似下，如图1.1，可以得到P点的场强与透光位置上的场强的关系

$$U(\vec{r}) = \frac{1}{i\lambda} \int U(\vec{r}') \frac{\exp(ikR)}{R} d\vec{r}' \quad (1.4)$$

P点的场强可以看作是透光位置上各点作为点源形成的球面波 $\exp(ikR)/R$ 的相干叠加，称为惠更斯原理。这意味着透光的区域中光波要有比较好的相干性。

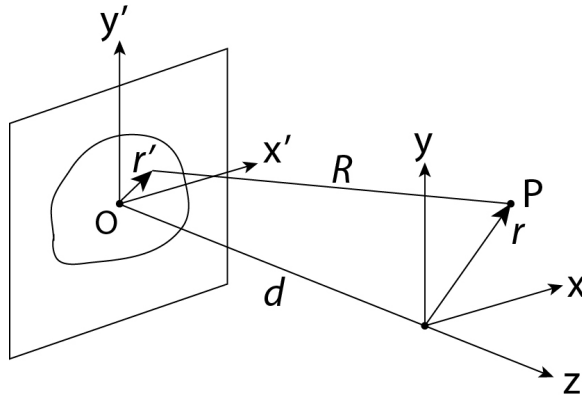


图 1.1: 惠更斯原理

$R^2 = (x - x')^2 + (y - y')^2 + d^2$ 。当 d 与 r 和 r' 相比足够大时，对 R 做近似

$$R \approx d + \frac{1}{2d}(x - x')^2 + \frac{1}{2d}(y - y')^2 \quad (1.5)$$

这里保留到了 x' 和 y' 的二次项。由于通常 $d \gg \lambda$ ，指数函数对 R 很敏感，相比之下分母上的 R 对高阶项的依赖性很弱，可以用 d 来近似。对于平面波入射，假设 $U(\vec{r}') = U_0$ 。于是

$$U(\vec{r}) = \frac{U_0 e^{ikd}}{i\lambda d} \int e^{\frac{ik}{2d}(x-x')^2} e^{\frac{ik}{2d}(y-y')^2} dx' dy' \quad (1.6)$$

考虑一个方形的窗口，边长为 $2a$ 。积分化为

$$U(x, y) = \frac{U_0 e^{ikd}}{i\lambda d} \int_{-a}^a e^{\frac{ik}{2d}(x-x')^2} dx' \int_{-a}^a e^{\frac{ik}{2d}(y-y')^2} dy' \quad (1.7)$$

定义Fresnel正弦和余弦积分分别为

$$S(z) = \int_0^z \sin(\pi t^2/2) dt \quad (1.8)$$

$$C(z) = \int_0^z \cos(\pi t^2/2) dt \quad (1.9)$$

并有 $S(-z) = -S(z)$ 和 $C(-z) = -C(z)$ 。进而

$$\int_0^z e^{i\pi t^2/2} dt = C(z) + iS(z) \quad (1.10)$$

Fresnel积分见图1.2。其级数展开分别为

$$S(z) = \sum_{n=0}^{\infty} \frac{(-1)^n (\pi/2)^{2n+1}}{(2n+1)!(4n+3)} z^{4n+3} \quad (1.11)$$

$$C(z) = \sum_{n=0}^{\infty} \frac{(-1)^n (\pi/2)^{2n}}{(2n)!(4n+1)} z^{4n+1} \quad (1.12)$$

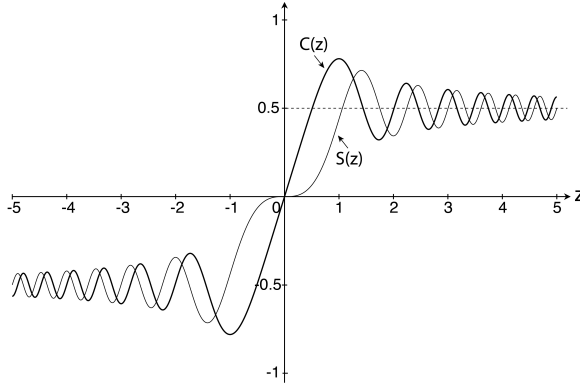


图 1.2: Fresnel积分

令 $s = a^2/\lambda d$, $\xi = \sqrt{2s}(x-x')/a$, 对 x' 的积分有

$$\begin{aligned} \int_{-a}^a e^{\frac{ik}{2d}(x-x')^2} dx' &= \int_{-a}^a e^{i\pi \frac{(x-x')^2}{2} \frac{2}{\lambda d}} dx' \\ &= \int_{-a}^a e^{i\pi(2s) \frac{(x-x')^2}{2a^2}} dx' \\ &= \frac{a}{\sqrt{2s}} \int_{-\sqrt{2s}(1+x/a)}^{\sqrt{2s}(1+x/a)} e^{i\pi u^2/2} d\xi \\ &= \frac{a}{\sqrt{2s}} (C(\sqrt{2s}(1+x/a)) + C(\sqrt{2s}(1-x/a))) \\ &\quad + i \frac{a}{\sqrt{2s}} (S(\sqrt{2s}(1+x/a)) + S(\sqrt{2s}(1-x/a))) \end{aligned} \quad (1.13)$$

令 $u(x) = \frac{1}{\sqrt{2}}(C(\sqrt{2s}(1+x/a)) + C(\sqrt{2s}(1-x/a))) + i \frac{1}{\sqrt{2}}(S(\sqrt{2s}(1+x/a)) + S(\sqrt{2s}(1-x/a)))$, 则

$$\int_{-a}^a e^{\frac{ik}{2d}(x-x')^2} dx' = \frac{a}{\sqrt{s}} u(x) = \sqrt{\lambda d} u(x) \quad (1.14)$$

于是

$$U(x, y) = -iU_0 e^{ikd} u(x)u(y) \tag{1.15}$$

而光强是

$$|U(x, y)|^2 = U_0^2 |u(x)|^2 |u(y)|^2 \tag{1.16}$$

例如 $a \sim 1 \text{ mm}$ 、 $\lambda \sim 1 \mu\text{m}$ 的情形， $d \sim 1 \text{ cm}$ 应该是足够远，这时 $s \sim 100$ 。图 1.3 显示的是当 d 逐渐变大 (s 逐渐变小) 时， $|u(x)|^2$ 的变化。

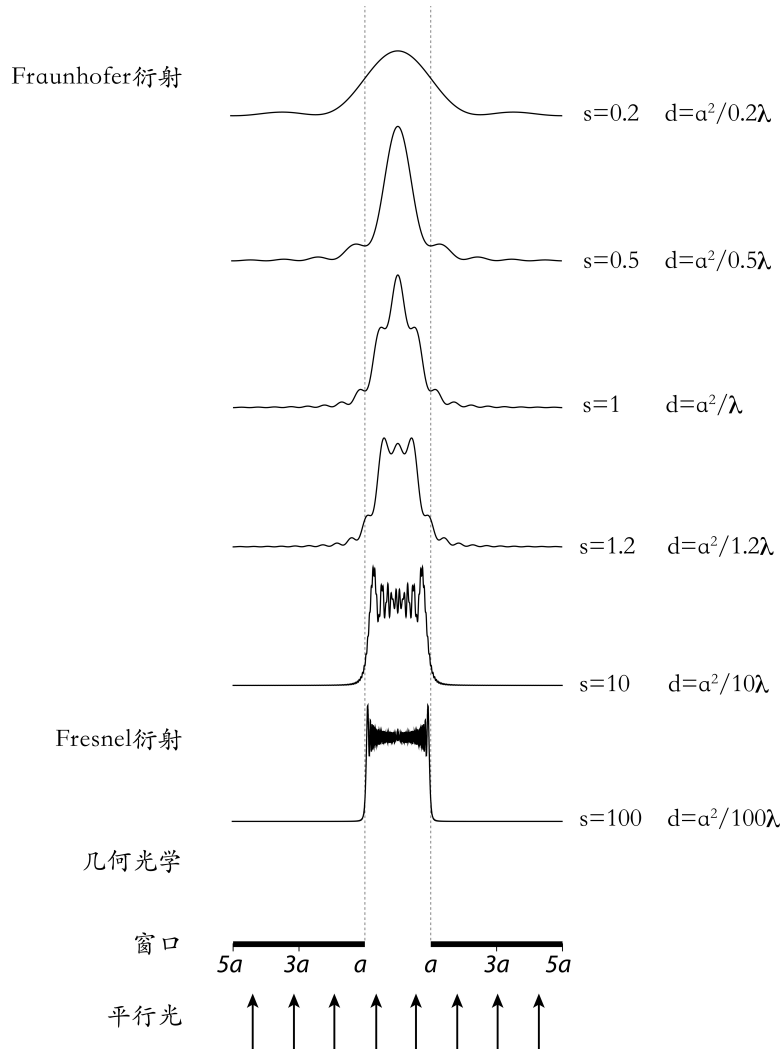


图 1.3: 从Fresnel衍射到Fraunhofer衍射

在窗口附近，投射光非常接近几何光学的描述。由于光的相干叠加，光强出现振荡。在数学上，振荡来源于图1.2中Fresnel积分的波动部分。远离窗口时，衍射图样发生明显变化。接近窗口的衍射称为Fresnel衍射（近场衍射），远离窗口的衍射称为Fraunhofer衍射（远场衍射）。转换发生在 $s \sim 1$ 的区域。在 s 很小 (d 很大) 时，场强计算中对 R 的近似只取到 x' 和 y' 一阶，即

$$R \approx d + \frac{1}{2d}(x^2 + y^2) - \frac{1}{d}(xx' + yy') \tag{1.17}$$

这时,

$$\begin{aligned}
 U(x, y) &= \frac{U_0 e^{ikd} e^{i(k/2d)(x^2+y^2)}}{i\lambda d} \int_{-a}^a e^{-\frac{ik}{d}xx'} dx' \int_{-a}^a e^{-\frac{ik}{d}yy'} dy' \\
 &= \frac{U_0 e^{ikd} e^{i(k/2d)(x^2+y^2)}}{i\lambda d} \left(\frac{\lambda d}{\pi x} \sin(2\pi sx/a) \right) \left(\frac{\lambda d}{\pi y} \sin(2\pi sy/a) \right) \quad (1.18)
 \end{aligned}$$

于是

$$|U(x, y)|^2 = 16U_0^2 s^2 \left(\frac{\sin(2\pi sx/a)}{2\pi sx/a} \right)^2 \left(\frac{\sin(2\pi sy/a)}{2\pi sy/a} \right)^2 \quad (1.19)$$

$(\sin \alpha/\alpha)^2$ 见图 1.4。形状与图 1.3 中最上面的曲线相同。

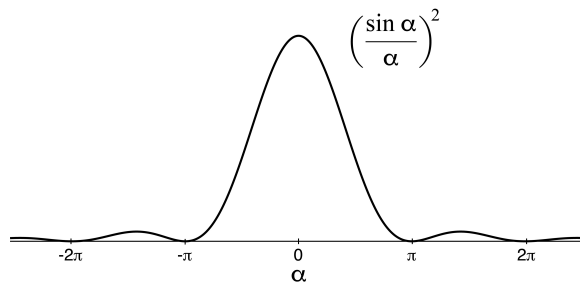


图 1.4: $(\sin \alpha/\alpha)^2$

从Fresnel衍射到Fraunhofer衍射的过渡来源于窗口范围内光波进行相干叠加时经历的相长与相消的交替次数随距离 d 的变化。这个过渡发生在 $s \sim 1$ 。当 d 比较小时, 透过的光基本上符合几何光学的规律, 形成一个与窗口几乎相同的光斑, 但光强会有振荡, 尤其是在边缘。Fresnel衍射发生在 $s = a^2/d\lambda \gg 1$, 即 $d \ll a^2/\lambda$ 时。如图 1.5, 当 $d \gg a$ 时, 窗口边缘到P点与中心O到P点的光程差是

$$\sqrt{a^2 + d^2} - d \approx a^2/2d \quad (1.20)$$

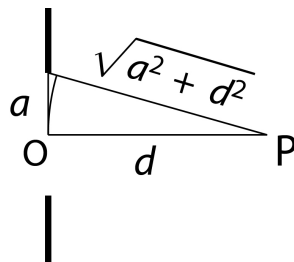


图 1.5: 光程差与 d 的关系

该光程差与半波长之比 $s = a^2/\lambda d$ 决定了P点的光强来源于多少次相长与相消的交替。在Fresnel衍射区, $s \gg 1$, P点以及偏离轴线的点的光强都来源于对多次振荡的求和并敏感地依赖于偏离轴线的距离。这时可观察到随位置快速振荡的光强。从 $s = 1$ 开始, 光程差小于半波长。在 $s \ll 1$ 时, 需要比较远的偏离才能看到光强明显的变化。这就是Fraunhofer区。

对于Fraunhofer衍射，由图1.4可知强度的第一个极小值点出现在 $2\pi sx/a = \pi$ ，即 $x = \lambda d/2a$ 。这时中心到极小值点与边缘到极小值点的光程差正好是半波长。通过计算还可以得到圆孔衍射的第一个极小值点出现在 $1.22\lambda d/D$ ， D 是圆孔的直径。

1.5 偏振

电场是一个矢量。以 \hat{k} 作为传播方向的单位矢量， \hat{i} 和 \hat{j} 是相互正交并与 \hat{k} 垂直的单位矢量。一个单色平面波的电场一般地可以写为

$$\vec{E}(\vec{r}, t) = (E_1 e^{i\phi_1} \hat{i} + E_2 e^{i\phi_2} \hat{j}) \exp(i\vec{k} \cdot \vec{r} - i\omega t) \quad (1.21)$$

这个平面波称为具有椭圆偏振。当 $\phi_1 - \phi_2 = n\pi$ ， $n = 0, 1, 2, \dots$ 时，称为线偏振。当 $E_1 = E_2$ ， $\phi_1 - \phi_2 = n\pi/2$ ， $n = \pm 1, \pm 3, \dots$ 时，称为圆偏振。

对偏振的测量可以得到光源的信息。另外，光穿过介质后通常偏振性质会发生变化。由此可得到介质的一些性质。

2 光学成像

光的主要用途有成像、光谱、操纵和信息传递。这里我们只讨论成像。光学（还包括电子学）往往不交代问题的本质，这与物理学的主流文化不同，我们希望尽量避免。

2.1 几何光学基本方程

通常用于成像的光的波长都短于1微米，远小于宏观物体的尺度。因此大多数情况下无需考虑光的波动性，称为几何光学。

我们将在 $\lambda \rightarrow 0$ 的条件下得到几何光学的基本方程。在此我们关心的是在空间某一点上光的传播方向，并用光线的概念来描述。光线是空间中的一条曲线，其每一点的切线与电磁波的传播方向重合。光场可以看作由一束光线组成。平面波是最简单波动行为。对于平面波，光线是空间中的一族直线，它垂直于波前（相位相同的点构成的面）并与波矢 \vec{k} 重合。光场在局域的小尺度范围上可以看作平面波。这样，一般地每一点的光的传播方向就由局域的波矢决定。电磁场的某一个分量用 f 表示。对于一个局域的单色平面波， $f = ae^{i\psi}$ 。相位 $\psi = \vec{k} \cdot \vec{r} - \omega t$ ， $k = n\omega/c$ 。折射率 n 是空间的函数。在传播过程中，光的频率 ω 不变，但波矢 \vec{k} 会因折射率的不同而改变，因而一般地是位置的函数。这样光线就有可能不再是一条直线，而是依赖于折射率的分布。我们来求光线满足的方程。

令 $\vec{r}(s)$ 为光线的轨迹， s 是沿曲线的路径长度。光线的切线在每一点都与波矢 \vec{k} 重合。 $d\vec{r}/ds$ 是曲线切线的单位矢量，于是有

$$\frac{d\vec{r}}{ds} = \frac{\vec{k}}{k} \quad (2.1)$$

或

$$k \frac{d\vec{r}}{ds} = \vec{k} \quad (2.2)$$

两边对 ds 求导数

$$\frac{d}{ds} \left(k \frac{d\vec{r}}{ds} \right) = \frac{d\vec{k}}{ds} \quad (2.3)$$

而

$$d\vec{k} = \frac{\partial \vec{k}}{\partial x} dx + \frac{\partial \vec{k}}{\partial y} dy + \frac{\partial \vec{k}}{\partial z} dz = d\vec{r} \cdot \nabla \vec{k} \quad (2.4)$$

于是

$$\frac{d}{ds} \left(k \frac{d\vec{r}}{ds} \right) = \frac{d\vec{r}}{ds} \cdot \nabla \vec{k} = \frac{\vec{k}}{k} \cdot \nabla \vec{k} = \frac{\nabla(k^2)}{2k} \quad (2.5)$$

将 $k = n\omega/c$ 带入, 得到

$$\frac{d}{ds} \left(n \frac{d\vec{r}}{ds} \right) = \nabla n \quad (2.6)$$

这是几何光学的基本方程。对于折射率为常数的介质, $d^2\vec{r}/ds^2 = 0$, 代表一条直线, 即均匀介质中光线为直线。

基本方程还可以用积分的形式表示。费马原理指出光从一点传播到另一点总是选择时间最短的路径。写成公式, 光的传播时间

$$T = \frac{1}{c} \int_a^b n ds \quad (2.7)$$

时间最短的路径意味着 T 的变分 $\delta T = 0$ 。 cT 称作光程。我们来证明由费马原理可以导出基本方程的微分形式。 a 和 b 两个端点固定, 对最短路径的一个小的偏离产生的 T 的变化是

$$\delta T = \int_a^b \delta n ds + \int_a^b n d\delta s \quad (2.8)$$

下面分别计算 δn 和 δs 。

$$\delta n = \nabla n \cdot \delta \vec{r} \quad (2.9)$$

而

$$(ds)^2 = d\vec{r} \cdot d\vec{r} \quad (2.10)$$

两边取变分

$$(ds)(d\delta s) = d\vec{r} \cdot d\delta \vec{r} \quad (2.11)$$

于是

$$d\delta s = \frac{d\vec{r}}{ds} \cdot d\delta \vec{r} \quad (2.12)$$

将 δn 和 $d\delta s$ 代入到积分中得到

$$\delta T = \int_a^b (\nabla n \cdot \delta \vec{r}) ds + \int_a^b n \frac{d\vec{r}}{ds} \cdot d\delta \vec{r} \quad (2.13)$$

另外,

$$d \left(n \frac{d\vec{r}}{ds} \cdot \delta \vec{r} \right) = d \left(n \frac{d\vec{r}}{ds} \right) \cdot \delta \vec{r} + n \frac{d\vec{r}}{ds} \cdot d\delta \vec{r} = d \left(n \frac{d\vec{r}}{ds} \right) \cdot \frac{\delta \vec{r}}{ds} ds + n \frac{d\vec{r}}{ds} \cdot d\delta \vec{r} \quad (2.14)$$

即

$$n \frac{d\vec{r}}{ds} \cdot d\delta \vec{r} = d \left(n \frac{d\vec{r}}{ds} \cdot \delta \vec{r} \right) - d \left(n \frac{d\vec{r}}{ds} \right) \cdot \frac{\delta \vec{r}}{ds} ds \quad (2.15)$$

代入到积分得到

$$\begin{aligned} \delta T &= \int_a^b (\nabla n \cdot \delta \vec{r}) ds + \int_a^b d \left(n \frac{d\vec{r}}{ds} \cdot \delta \vec{r} \right) - \int_a^b d \left(n \frac{d\vec{r}}{ds} \right) \cdot \frac{\delta \vec{r}}{ds} ds \\ &= \int_a^b \left(\nabla n \cdot \delta \vec{r} - d \left(n \frac{d\vec{r}}{ds} \right) \cdot \frac{\delta \vec{r}}{ds} \right) ds + d \left(n \frac{d\vec{r}}{ds} \cdot \delta \vec{r} \right) \Big|_a^b \end{aligned} \quad (2.16)$$

a和b两个端点固定，因此上式最后一项等于零，于是

$$\delta T = \int_a^b \left(\nabla n - \frac{d}{ds} \left(n \frac{d\vec{r}}{ds} \right) \right) \cdot \delta \vec{r} ds \quad (2.17)$$

对于时间最短路径， $\delta T = 0$ ，且对于任意的 $\delta \vec{r}$ 成立。这意味着

$$\nabla n - \frac{d}{ds} \left(n \frac{d\vec{r}}{ds} \right) = 0 \quad (2.18)$$

得到基本方程。

使用费马原理可以直接得到一些重要的结论，如均匀介质中光沿直线传播、反射定律、折射的Snell定律等。折射光学元件往往是一个光学成像系统的核心。例如，如图所示的凸透镜。沿PP'直线传播的光线经过的介质较厚，而沿任一路径PQQ'P'传播的光线虽然路程较长但经过的介质较薄。合适的设计可以使得二者具有相同的光程，从而实现P发出的沿任何角度的光线都会汇聚到P'点。P和P'称为透镜的焦点。

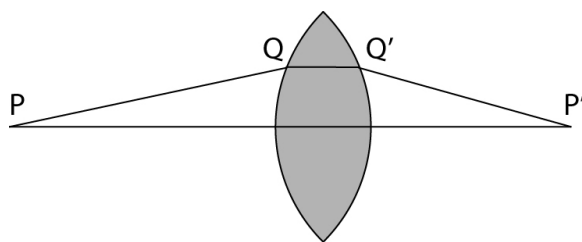


图 2.1: 透镜的焦点

2.2 几何光学成像

光照射到一个物体上后，物体上的每一个点都会散射光。散射光带有这一点上的信息。如果我们能够用散射光尽可能地将这个信息带到感光元件，如CCD、视网膜等，就可以将物体成像。但这里的一个障碍是物体上每个点发生的是漫散射，接近于一个点光源。如果将成像介质直接置于物体前，则感光元件上每一个点接收到的是物体上所有点散射出来的光的总合，因此需要很复杂的算法才能重建物体的图像。需要找到这样一种办法方便图像重建。最简单的一种方案是设法建立物体上的点与感光元件上的点之间的一一映射关系，将从物体上某一点散射出来的光带到感光元件上唯一的一个点且物体上不同的点对应感光元件上不同的点。一般地，我们还要求映射到感光元件上的各点间保持与物体上相应各点间相同的空间次序关系。这样就可以实现严格成像。这个方案可以有几种方法实现。

方法一是小孔成像。如图，小孔的存在限制了可以到达成像元件的光线，从而建立了物体上每一点与感光元件上每一点之间的一一映射关系。除了相与物成倒立关系外，相上点之间的空间次序关系与物相同，达到严格成像的目的。

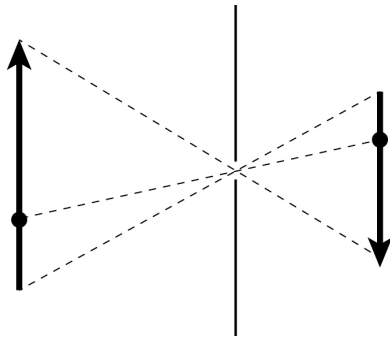


图 2.2: 小孔成像

另两种方法，即复眼和单眼，已经在自然界中通过生物进化得以实现。

复眼是昆虫的视觉器官。复眼由大量小眼组成。每个小眼是一个独立的感光元件，并与一根视神经相连。复眼中经常含有色素细胞来吸收斜射的光线，使得只有直射的光才能到达感光细胞。这意味着每个小眼都有准直的功能，从而建立了物体与感光元件间的点点对应关系，并且空间次序关系不变，实现严格成像。

单眼的主要部件之一是一个透镜。透镜的功能是进行信息变换。物体和感光元件分别位于透镜的两侧。透镜的折射效应改变光的传播方向，将物体上各点的信息进行组合而变换到另一侧的感光元件上。一般地，落在感光元件上一点的光强是

$$I(P') = \int \tau(P', P) I(P) dP \quad (2.19)$$

$I(P)$ 是物体上 P 点的光强。变换函数的形式依赖于透镜。选择合适的透镜，则可以使变换函数取这样一个形式以至于 $\tau(P', P)$ 对于处于特定空间范围（称为相平面）内的一个点 P' 只对物上某一个 P 点附近才有非零的贡献，并保证相上的点 P' 与物上的点 P 在空间次序关系上相同，实现严格成像。如果放松严格成像的要求，理论上讲已知了函数 τ 也可以通过对象进行计算反演出物来。这里值得注意的是上式使用的是光的非相干叠加，即光强的相加。

2.3 基于折射元件的几何光学成像

严格成像意味着相似变换。两个二维图形间满足相似关系意味着如下的线性变换： $x \rightarrow \alpha Ax + t$ 。其中 x 是一个矢量， A 是一个实现转动的正交变换矩阵（包括反演）， α 是放大系数， t 是平移。一个简化的实现方法是：物和像各存在一个原点，在极坐标下一个物点的坐标是 (ρ, θ) ，该物点映射到一个唯一的像点上，该像点的坐标是 (ρ', θ') ，满足 θ 和 θ' 只相差一个常数以及 $\rho' = \alpha\rho$ 。这需要证明一下几点：一、物上任何一点发出的所有光线经过成像系统后能汇聚于一点。这两个点称为共轭点。二、对于物处于同一个垂直于轴的平面上的点，它们的像也处于一个垂直于轴的平面上。这样的一对平面称为共轭面。三、保证 θ 和 θ' 只相差一个常数。四、保证 $\rho' = \alpha\rho$ 。这些条件保证了清

晰无畸变的成像。我们将发现存在一种具有轴对称性并且无手性的理想成像系统可以实现这样的变换。由于轴对称性，在物平面上从原点到物点的矢量与映射到像空间后形成的矢量之间的夹角为常数。由于没有手性，这个常数只能是0或 π 。这个一般的讨论就可以保证第三条，并且能将这个常数确定下来。下面讨论其他几条。

我们将证明由若干球面透镜排列在一条公共轴线上可以非常地接近一个理想系统。这条轴线通过所有球面透镜的球心，称为光轴。

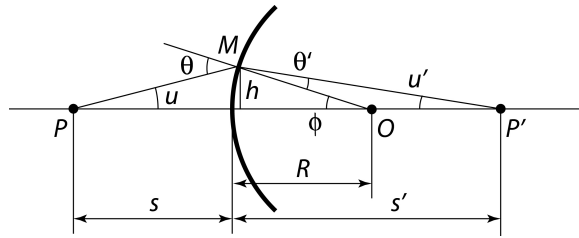


图 2.3: 傍轴近似

球面相比其他曲面更便于精确加工，因此大多数的折射光学元件都是球面透镜。我们先证明位于光轴上的一点存在一个同样位于光轴上的共轭点。如图 2.3，球面光学元件的折射率为 n ，位于大气中。如果我们只感兴趣光轴近旁的光线，则 u 、 u' 、 ϕ 、 θ 、 θ' 很小，可以做傍轴近似，对角度的三角函数只取最低阶近似。折射定律给出

$$\theta = n\theta' \quad (2.20)$$

而

$$\theta = \phi + u, \theta' = \phi - u' \quad (2.21)$$

由此得到

$$u + nu' = (n - 1)\phi \quad (2.22)$$

傍轴近似还给出

$$u = h/s, u' = h/s', \phi = h/R \quad (2.23)$$

最后得到

$$\frac{1}{s} + \frac{n}{s'} = \frac{n - 1}{R} \quad (2.24)$$

这表明 P' 点的位置与 M 点无关， P 点发出的光都汇聚到 P' 点。 P 与 P' 互为共轭。第一个条件得以满足。

再来讨论轴外物点的傍轴成像。如图所示，将直线 PP' 绕球心转过一个小的角度到达 QQ' 。由球对称性可知 Q 与 Q' 共轭。由于 Q 非常接近轴，可以认为 Q 与 P 处于一个垂直于光轴的平面上， Q' 与 P' 也有同样的情形。穿过 P 和 P' 并垂直于光轴的两个面互为共轭。第二个条件满足。

下面再看第三个条件。如图 2.4，在傍轴条件下

$$\frac{y'}{y} = \frac{\overline{OP'}}{\overline{OP}} = \frac{s' - R}{s + R} \quad (2.25)$$

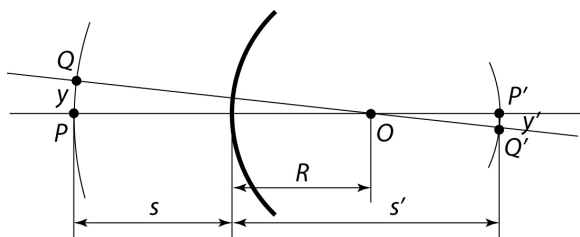


图 2.4: 轴外物点成像

即共轭的两个点到光轴的距离之比是常数。

综上所述，在傍轴条件下，共轴球面透镜组成的系统能够比较好地实现严格成像。

常用的透镜有两个球面。对于薄透镜，我们不加证明地给出

$$\frac{1}{s} + \frac{1}{s'} = \frac{1}{f} \quad (2.26)$$

其中

$$f = \frac{1}{(n-1) \left(\frac{1}{R_1} + \frac{1}{R_2} \right)} \quad (2.27)$$

称为焦距。 R_1 和 R_2 分别是两个球面的半径。

透镜系统的实际成像对理想成像存在偏离，称为相差。相差分为两类：纯几何来源的单色相差和折射率对波长的依赖性导致的色差。

傍轴光学是一阶近似，即取了 $\sin \theta = \theta$ 和 $\cos \theta = 1$ 。更高一阶近似需要取 $\sin \theta = \theta - \theta^3/6$ 和 $\cos \theta = 1 - \theta^2/2$ 。在此近似下，相差分为五类：球差、彗差、像散、像场弯曲、畸变。一般地，各种相差的大小依赖于物点到光轴的距离 y 。其中，球差与 y 无关，彗差正比于 y ，像散和像场弯曲正比于 y^2 、畸变正比于 y^3 。在光轴上的物点，球差是唯一的单色相差。如图2.5，从光轴上一点发出的光线，如果角度不同，经过透镜后汇聚的点也会有所不同，称为球差。

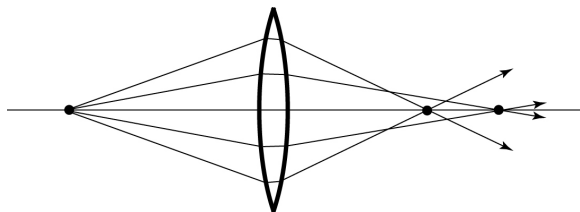


图 2.5: 球差

2.4 成像仪器

一个典型的成像系统包括一个透镜和成像介质。理想情况下的严格成像，物点与像点间有一一对应关系。但是像都需要由一个介质来采集，如胶片、视网膜、CCD等等。任何一个介质都由分立的感光单元组成，如胶片上的颗粒、视网膜上的感光细胞、CCD上的像素。感光单元都有一定大小，而不是一个理想化的点。于是每个感光单元都对应物上的一个区域，而不是一个点。这个区域中的物点发出的光会落到同一个感光单元上，因此不可分辨。这就导致了空间分辨率的问题。如果两个物点经过光学系统后它们的像点分别落在了两个不同的感光单元上，我们就称这两个物点可以分辨。将物划分成多个区域。代替点对点关系的是物上的一个区域与一个感光单元一一对应。可能的最小感光单元是单个原子，这决定了可能的最高空间分辨率。在设计上，仪器利用折射效应将两个物点的光线带到在垂直光轴的平面上尽可能分得远的两个像点上。另外，一个感光单元有其能够产生响应的光强阈值。所以仪器还需要将从一个物点发出的光线尽可能多地带到像点。

成像仪器的最基本构型是一个透镜和一个感光介质，如CCD。成像的核心问题是分辨率。如图2.6。

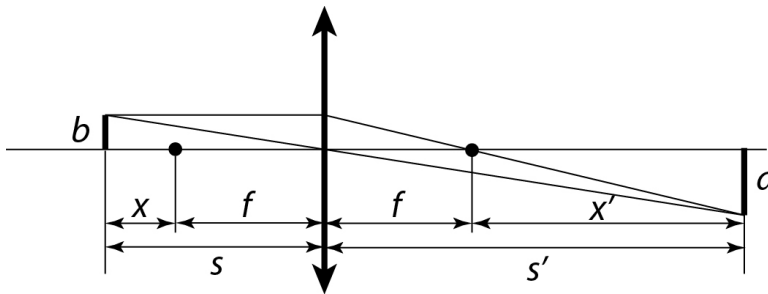


图 2.6: 照相机

图中 a 是相邻两个感光单元之间的距离， b 是可以分辨的最小空间尺寸。

$$b = a \frac{s}{s'} \quad (2.28)$$

代入 $s' = sf/(s - f)$ 得到

$$b = a \left(\frac{s}{f} - 1 \right) = a \frac{x}{f} = a \frac{f}{x'} \quad (2.29)$$

容易看到物体越接近焦点，能够分辨的尺寸就越小，分辨率越高。一个最简单的设计是将 s' 固定下来，这样最小能够分辨的尺寸也就固定下来。使用时只需要调节物距得到清晰的像即可。虽然不用同时调节物距和相距以方便使用，但分辨率也被固定下来。为了使分辨率可调，可以再加一个透镜，如图2.7。这个结构称为显微镜。接近被观察物体的透镜称为物镜，另一个透镜称为目镜。物镜到目镜、目镜到感光介质的距离都固定。这样每一个物镜与目镜的组合就对应一个确定的分辨率。通常一台显微镜上会有几个物镜用于

切换。调节显微镜使得物体处于物镜焦点外侧合适的位置就可以得到清晰的像。通常物镜成的像很接近目镜的焦点，即 $x'_2 \approx \Delta$ 。有如下关系：

$$b = a \frac{f_1 f_2}{x'_1 x'_2} = a \frac{f_1 f_2}{x'_1 \Delta} \quad (2.30)$$

改变物镜的焦距就可以得到不同的分辨率。

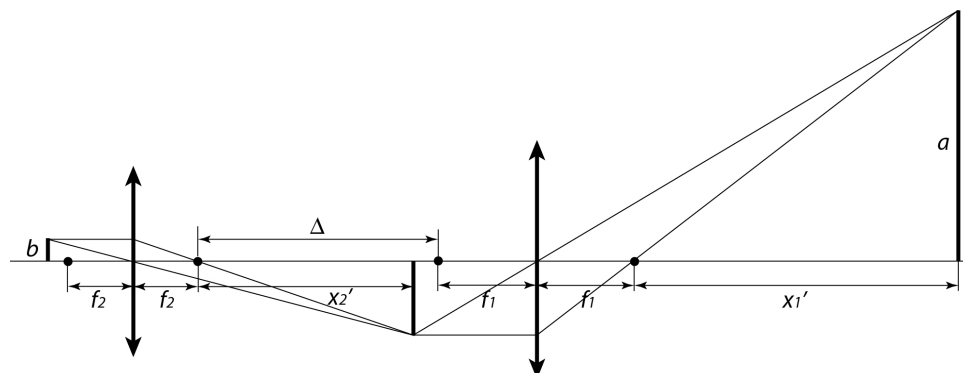


图 2.7: 显微镜

另一种情况是物距不易调节，即 s 几乎是固定的，例如星体。这时可以用类似的方法，即物镜，来对分辨率进行调节。一般地，物镜与目镜之间的距离可以改变。这个结构称为望远镜。如图 2.8。由于物距很大，物体成的像非常接近焦点并比物小得多。这时再用一个目镜（也可用一个显微镜）对其成像。通常让目镜的焦点置于非常接近物镜焦点的位置。对于望远镜，人们通常关心的是角分辨率，而不是绝对大小。这是因为物距可以有任意的大小，不同物距对应的空间分辨率不同，而角分辨率可以一样，因此更方便表征仪器特性。远处两个物点对于观测者形成一个张角 α 。我们想知道使两个物点能被分辨的最小张角是多少。张角为 α 的两个物点，经过物镜成像后两个像点的距离是 αf_2 。如这两个像点经目镜成像后正好落在相邻的连个感光单元上，则有

$$\alpha f_2 = a \frac{f_1}{x'_1} \quad (2.31)$$

即

$$\alpha = \frac{a f_1}{x'_1 f_2} \quad (2.32)$$

记 $\alpha' = a/x'_1$ ，则有

$$\alpha = \alpha' \frac{f_1}{f_2} \quad (2.33)$$

由于物镜成的像距离目镜焦点很近，所以 α' 几乎就是两个相邻感光单元对于目镜中心张成的角。一般地，使用大的物镜焦距和小的目镜焦距会提高角分辨率。

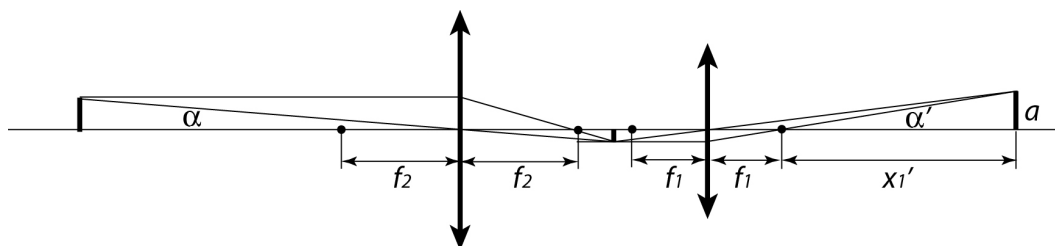


图 2.8: 望远镜

2.5 分辨率极限

光是电磁波。由于衍射，一个点的像不再是个点，而是一个光斑。光斑的大小依赖于光学系统的孔径。衍射现象限制了成像的分辨率。由此而来的分辨率为 $1.22\lambda f/D$ 。

2.6 傅立叶光学

在讨论光学成像时提过最理想的成像构型是像点和物点一一对应。到一个像点的光都来自同一个物点，而不同的像点的光来自不同的物点。但一般地，一个像点上的光会来自不同的物点。如果用非相干光成像，来自不同物点的光的强度相加，没有干涉效应。如果用相干光成像，需要由电场振幅进行叠加。我们考虑这样一种情形：电场按照傅立叶变换的权重进行叠加，结果成的像就是物的傅立叶变换。如图2.9。考虑焦平面上的一点 P' 。这一点的电场振幅 A 来源于一组平行光的叠加。

$$A(P') \sim \int A(P) e^{iky \sin \theta} dP \quad (2.34)$$

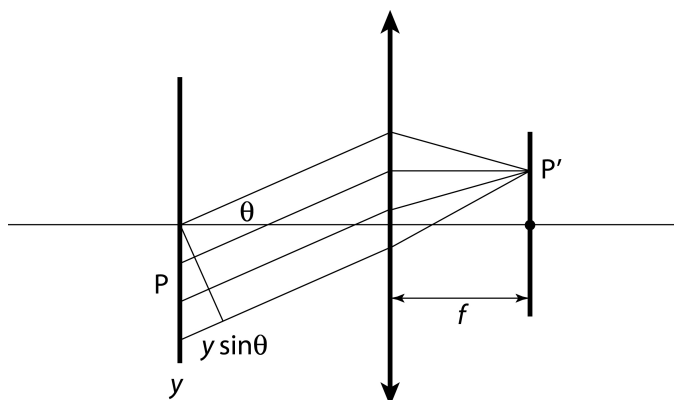


图 2.9: 傅立叶光学

于是 P' 点的光实质上是对物的傅立叶变换。