Lecture 34

Higher-order Gaussian Beams

We obtained the Gaussian beam solution as an exact solution to the paraxial wave equation by starting with a spherical wave, which we know is an exact solution to the full Helmholtz equation, and applying the paraxial (Fresnel) approximation to it. However, this turns out to give only the lowest-order solution of a whole family of possible solutions.

A general approach to solve the paraxial wave eqn. is to start with solutions of the form

$$\varphi(x, y, z) = A(z) e^{-ik\left(x^2+y^2\right)/2q(z)},$$

Plug into the paraxial wave eqn. and obtain differential equation for A(z) and q(z).

This approach is carried out in detail in Siegman sec. 16.4 and Haus appendix S. We don't have the time to carry out the solution, but it is important to note the final results for future reference. The two most common forms of the general solution are:

(i) Hermite-Gaussians, most useful when rectangular coordinates are the most relevant , and

(ii) Laguerre-Gaussians, most useful in systems with cylindrical symmetry

Hermite-Gaussian solution:

The general form of the solution ψ is then

ψ ~ (Hermit polynomials)$\times$(Gaussian).

$\psi_{nm} \sim \left[H_n(x)H_m(y)\right] \times$(Gaussian).

It is important to note that the complex radius q(z) still satisfies

$$\frac{1}{q(z)} = \frac{1}{R(z)} - i\frac{\lambda}{\pi w^2(z)}$$

The full expression is given in Siegman eqn. 16-54 or 17-40(see handout).

(note that our $\psi(x, y, z)$ is basically Siegman's $\tilde{\mu}(x, y, z)$)

Here it will suffice to note that the low-order Hermite polynomials are

$\mathrm{H}_0 = 1$ (this yields our lowest-order Gaussian beam)

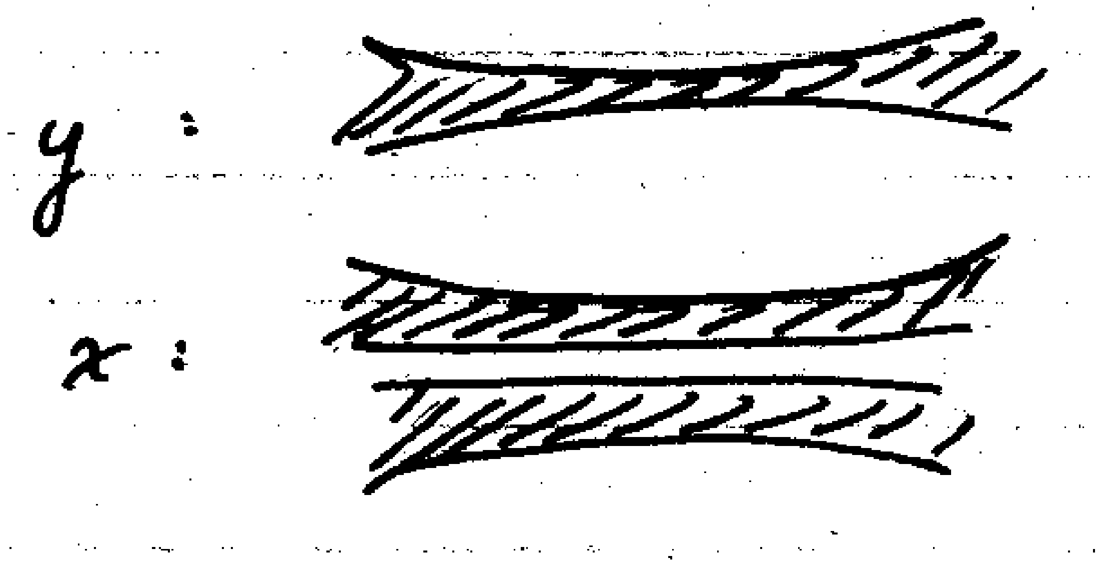
$\mathrm{H}_1(x) = 2x$ (odd)

$$\mathrm{H}_2(x) = 4x^2 - 2(\text{even})$$
$$\mathrm{H}_3(x) = 8x^3 - 12x$$
$$\vdots$$
$$\mathrm{H}_{n+1}(x) = 2x\mathrm{H}_n(x) - 2n\mathrm{H}_{n-1}(x)$$

For plots of these, see Siegman fig.17.19

Note that the arguments of the Hermite polynomials in the ψ are scaled to $w(z)$, which means that the <u>shape</u> of the beam does <u>not</u> change with propagation, only its size does.

e.g. ψ_1 looks like a zero-order Gaussian in y , and a 2^{nd} order Hermitz-Gaussian in x , for all z :



Higher-Order Hermite-Gaussian Mode Functions

The free-space Hermite-gaussian TEM_{nm} solutions derived in the preceding chapter can be written, in either the x or y transverse dimensions, and with the plane-wave e^{-jkz} phase shift factor included for completeness, in the normalized form

$$\begin{aligned} \tilde{u}_n(x,z) = \left(\frac{2}{\pi}\right)^{1/4} \left(\frac{1}{2^n n! w_0}\right)^{1/2} \left(\frac{\tilde{q}_0}{\tilde{q}(z)}\right)^{1/2} \left[\frac{\tilde{q}_0}{\tilde{q}_0^*}\frac{\tilde{q}^*(z)}{\tilde{q}(z)}\right]^{n/2} \\ \times H_n\left(\frac{\sqrt{2}x}{w(z)}\right) \exp\left[-jkz - j\frac{kx^2}{2\tilde{q}(z)}\right], \end{aligned} \tag{40}$$

where the H_n's are the Hermite polynomials of order n, and the parameters $\tilde{q}(z)$, $w(z)$ and $\psi(z)$ are exactly the same as for the lowest-order gaussian mode as given in Equation 17.5. These same functions can be written in alternative form, emphasing the spot size $w(z)$ and Guoy phase shift $\psi(z)$, in the form

$$\begin{aligned} \tilde{u}_n(x,z) = \left(\frac{2}{\pi}\right)^{1/4} \left(\frac{\exp[j(2n+1)\psi(z)]}{2^n n! w(z)}\right)^{1/2} \\ \times H_n\left(\frac{\sqrt{2}x}{w(z)}\right) \exp\left[-jkz - j\frac{kx^2}{2R(z)} - \frac{x^2}{w^2(z)}\right], \end{aligned} \tag{41}$$

where $\psi(z)$ is still given by $\psi(z) = \tan^{-1}(z/z_R)$.

Note the important point that the higher-order modes, because of their more rapid transverse variation, have a net Guoy phase shift of $(n + 1/2)\psi(z)$ in traveling from the waist to any other plane z, as compared to only $\psi(z)$ for the lowest-order mode. This differential phase shift between Hermite-gaussian modes of different orders is of fundamental importance in explaining, for example, why higher-order transverse modes in a stable laser cavity will have different oscillation frequencies; or how the Hermite-gaussian components that add up to make a uniform rectangular or strip beam in one transverse dimension at an input plane located in the near field (at a beam waist) can add up to give a $(\sin x)/x$ transverse variation for the same beam in the far field.

Hermite-Gaussian Mode Patterns

Figure 17.18 illustrates the transverse amplitude variations for the first six even and odd Hermite-gaussian modes. Note that the first few (unnormalized) Hermite polynomials are given by

$$\begin{aligned} H_0 &= 1 & H_1(x) &= 2x \\ H_2(x) &= 4x^2 - 2 & H_3(x) &= 8x^3 - 12x. \end{aligned} \tag{42}$$

These polynomials obey the recursion relation

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x) \tag{43}$$

which can provide a useful way of calculating the higher-order polynomials in numerical computations.

The Hermite-gaussian beam functions alternate between even and odd symmetry with alternating index n. The n-th order function has n nulls and $n+1$

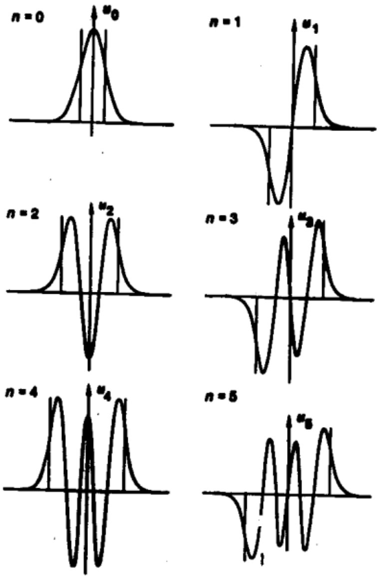


FIGURE 17.18
Amplitude profiles for low-order Hermite-gaussian modes.

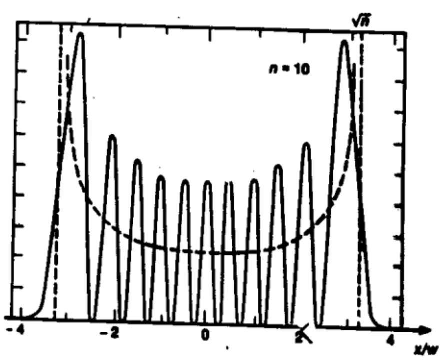


FIGURE 17.19
Intensity profile for the Hermite-gaussian mode pattern with $n = 10$.

peaks. These same Hermite-gaussian functions are also the quantum mechanical eigenfunctions for the linear quantum harmonic oscillator. Figure 17.19 illustrates the intensity variation, or the wave amplitude squared, for the $n = 10$ eigenmode, showing how the wave distribution approaches the classical proba-

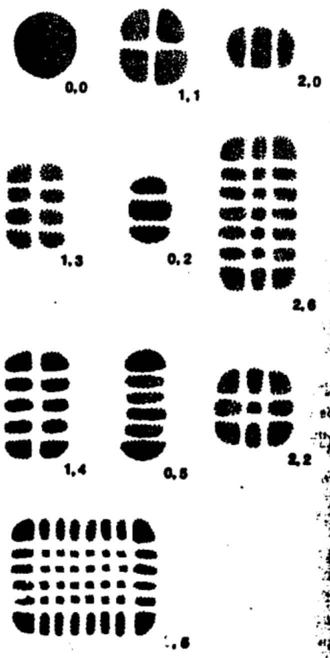


FIGURE 17.20
Transverse mode patterns for Hermite-gaussian modes of various orders.

bility density for a linear harmonic oscillator. It can also be seen that for larger values of n the outermost peaks become noticeably more in e than the inner peaks.

The complete set of Hermite-gaussian transve e mod . a beam in two transverse dimensions can then be written as $\tilde{u}_{nm}(x, y, z) = \tilde{u}_n(x, z) \times \tilde{u}_m(y, z)$, where in the most general situation a different $\tilde{q}(z)$ param ter, and even a different waist location, may apply to the x and the y variati . Figure 17.20 shows how the intensity patterns of various higher-order mod appear if the output beam from a laser oscillating in one of these higher-order modes is projected onto a screen. Note that the H rmite-gaussian functions are erwi e. scaled to th spot size w through the arguments x/w and y/w. He *he intensity pat ern of any given TEM_{nm} mode changes size but not sha pe s it propagates forward in z*—a given TEM_{nm} mode looks exactly the same, e ept for scaling at every point along the z axis.

The higher-order Laguerre-gaussian mode patterns also described in Section 16.4 (cf. Equation 16.64) are characterized by azimutha and radial symmetry, rather than by the rectangular symmetry of the Hermite-gaussian modes, as illustrated in Figure 17.21. As explained earlier, most real lasers r to oscillate in modes of rectangular rather than cylindrical symmetry, althoug with very

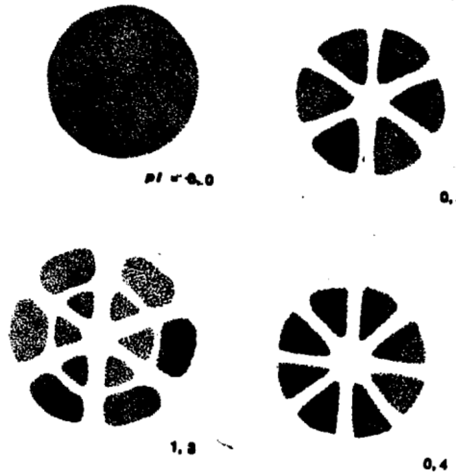


FIGURE 17.21
Transverse mode patterns for Laguerre-gaussian modes of various orders.

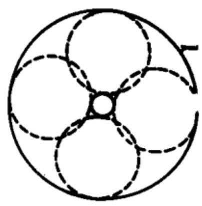


FIGURE 17.22
The "donut" mode is a linear superposition of 10 and 01 Hermite-gaussian modes.

careful adjustment, certain internal-mirror lasers can be made to oscillate in the cylindrical Hermite-gaussian modes.

The "Donut Mode"

In many laser experiments with stable laser resonators, the experimental procedure is to stop down an adjustable circular aperture inside the laser cavity until higher-order mode oscillation is completely suppressed and the laser oscillates only in the desired TEM_{00} mode. For aperture diameters slightly larger than this value, lasers are often observed to produce an output beam in the form of a circularly symmetric ring with a dark spot on axis, as illustrated in Figure 17.22.

This mode, often referred to as the "donut mode," cannot be an $m = 0$ mode, since an $m = 0$ Laguerre-gaussian mode does not have a null on axis. It might be interpreted as a higher-order $\tilde{u}_{pm}(r, \theta)$ Laguerre-gaussian mode with $p = 1$ and an azimuthal variation like $e^{jm\theta}$ with $m \geq p$. In most practical lasers, however,

Gouy effect: the n-m^{th} Hermite-Gaussian φ_{nm} contributes an excess phase

$$(n+m+1)\phi(z)$$

Where $\phi = \tan^{-1}\dfrac{z}{z_R}$ just as before.

Thus higher order Hermite-Gaussian has a larger excess phase.

(One of the most important physical consequences of this is that the transverse modes of lasers oscillate at slightly different frequencies; thus multi-transverse-mode lasers exhibit "mode-beating.")

Theorem: The Hermite-Gaussian solutions form a complete set of orthonormal functions.

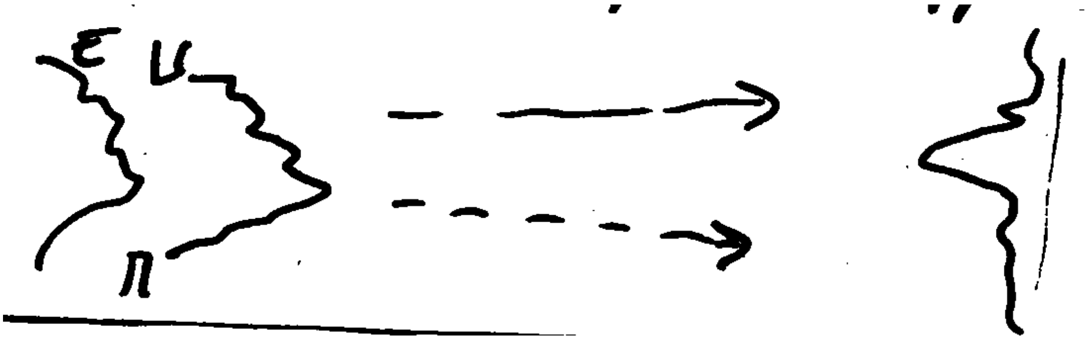
This means that any electric field at some plane z can be written as

$$E(x, y, z) = \varepsilon_0 \sum_n \sum_m c_{nm} \psi_{nm} e^{-ikz}$$

$$c_{nm} = \iint dxdy \frac{E(x, y, z)}{\varepsilon_0} \varphi_{nm}(x, y, z)$$

$$\iint \varphi_{nm} \varphi_{n'm'} = \delta_{nn'} \delta_{mm'}$$

This could be used to solve arbitrary diffraction problems (within the paraxial approximation, of course).



The complicated diffraction pattern on the observation screen is, in this picture, just the consequence of all the different phase shifts (n+m+1) $\phi(z)$, and the amplitudes c_{nm}.

In practice, this approach is useful only in limited situations, where only a few modes must be kept.

Laguarre-Gaussian modes

- Write paraxial wave eqn. in cylindrical coordinates (r, θ, z)

$\rightarrow$ solution

$\varphi \sim$ (Laguarre polynomial) $\times$ (Gaussian)

-see Siegman fig.17.21 for representative plots

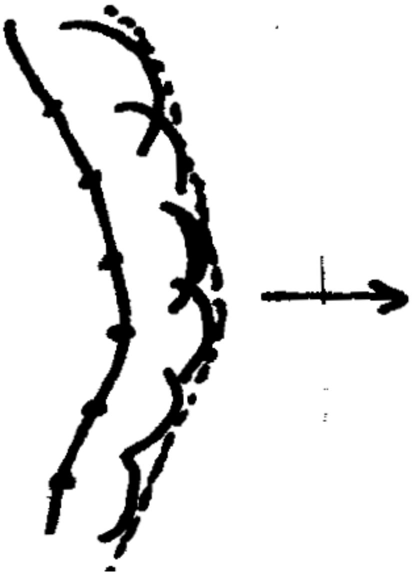
(note that this set of solutions also forms a complete orthonormal set)

Diffraction: Integral Approach

We shall shortly see how an integral solution to the Helmholtz equation can be obtained, but before doing so, it is worthwhile to consider diffraction from a different point of view, namely that of Huygen's Principle.

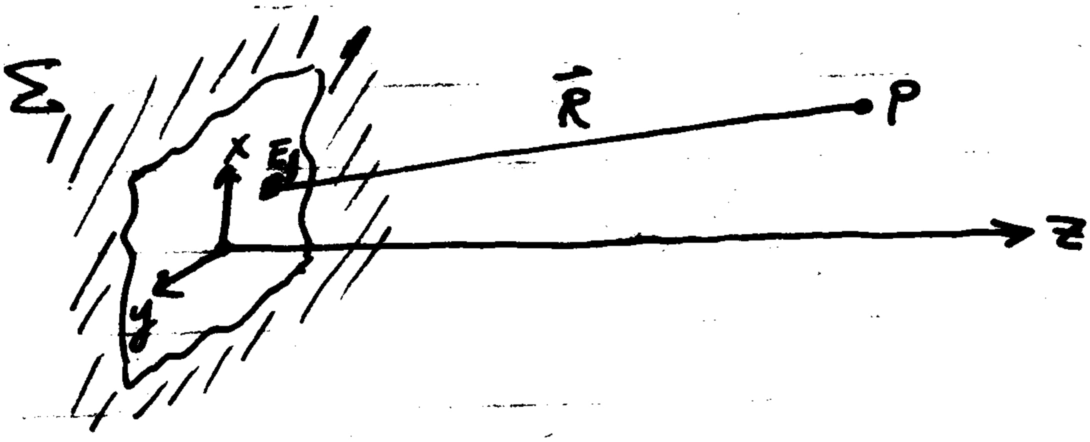
Huygen's Principle

Every point on a wavefront can be considered to be a source of a spherical wave (Huygen's wavelet). The envelope of the wavelets gives the wavefront at a later time.



(See Guenther P.326 for a simple derivation of the laws of reflection and refraction from this principle.)

Huygen's Principle is usually applied to diffraction problems in the following simple way. Suppose we know the (scalar) optical field E(x, y) over some aperture in a plane $\sum$ (see diagram).The basic idea is to consider each point in the aperture to be a source of Huggen's wavelets. The total field at an observation point P is just the sum (integral) of all the wavelets!



Roughly speaking, the i^{th} element in the aperture contributes a field $E_i\left(\vec{r}_i\right)$ to the total field in $\sum$, and radiates a Huggens spherical wave

$$E_i\left(\vec{r}_i\right) = \frac{e^{-i\vec{k}\cdot\vec{R}}}{R}$$

Now, again we crudely say that the total field at P is the sum of all the elements j in $\sum$. Of course we take the continuum limit so the sum goes over into an integral

$$E_P \sim \iint_{\Sigma} E_i\left(x, y\right)\frac{e^{-i\vec{k}\cdot\vec{R}}}{R}dxdy$$

In fact, this expression would be correct if we just put a factor (i/λ) in front of the integral.

We shall see below in a "rigorous" treatment of the problem how this factor arises. Here it will suffice to note that

1. The "I" comes essentially from the Gury phase shift
2. The "λ" factor has no "a priori" interpretation – but might be thought it as: source strength ~field per wavelength

Thus $\boxed{E_P = \frac{i}{\lambda}\iint_\Sigma E(x,y)\frac{e^{-i\vec{k}\cdot\vec{R}}}{R}dxdy}$

This is known as the Huygens-Fresnel integral. The physical idea behind it is quite simple and important to remember (!), but our argument is at best only a plausibility argument. Our next job is to provide some proof that indeed this does represent a valid solution to the problem.

Fresnel – Kirchhoff Theory

The goal is to find a solution to the Helmholtz eqn.

$$\left[\nabla^2 + k^2\right]\psi(\vec{r}) = 0$$

Where the (scalar) electric field is

$$E(\vec{r},\mathrm{t}) = E_0\psi(\vec{r})e^{i\omega t}$$

(i.e. we consider a pure time-harmonic wave).

¤Note slight change in notation / meaning of ψ - it's the scalar field amplitude not a slowly varying envelope

Of course, since Helmholtz's eqn. is linear, we can find solutions for waves more complicated than simple harmonic (i.e. for wave groups) by linear superposition of the pure-harmonic solutions.

Huygen's Principle says we should be able to find the field at any point P_0 given some initially known wave. The mathematical statement of this idea can be given starting from Green's theorem, which is obtained from Gauss's theorem familiar from vector calculus in the following way.

Gauss's theorem.

$$\iint_S \vec{F}\cdot\hat{n}ds = \iiint_V \nabla\cdot\vec{F}dV$$

Where $\vec{F}$ = vector function (nonsingular in V)

S = surface enclosing volume V

Let $\vec{F} = \varphi\nabla\psi\left(\varphi_1\psi = \text{scalar functions}\right)$

⇨ $\iint_s (\varphi\nabla\psi\cdot\hat{n})\,ds = \iiint_V \nabla\cdot(\varphi\nabla\psi)\,dV = \iiint_V \left(\varphi\nabla^2\psi + \nabla\varphi\cdot\nabla\psi\right)dV$

We could just as well let $\vec{F} = \psi\nabla\varphi$, giving

$$\iint_s \psi\nabla\varphi\cdot\hat{n}ds = \iiint_V \left(\psi\nabla^2\varphi + \nabla\psi\cdot\nabla\varphi\right)dV$$

Subtracting the two yields Green's theorem

$$\iint_s (\varphi\nabla\psi - \psi\nabla\varphi)\cdot\hat{n}ds = \iiint_V \left(\varphi\nabla^2\psi - \psi\nabla^2\varphi\right)dV$$

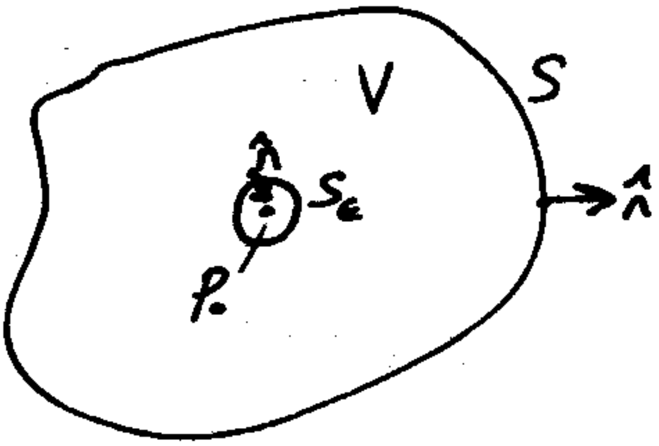
This is often written in a slightly different way, using

$$\nabla\psi\cdot\hat{n} = \frac{\partial\varphi}{\partial n} \quad (\hat{n} = \text{unit normal to surface S})$$

⇨ (*) $\iint_s \left(\varphi\frac{\partial\psi}{\partial n} - \psi\nabla\frac{\partial\varphi}{\partial n}\right)ds = \iiint_V \left(\varphi\nabla^2\psi - \psi\nabla^2\varphi\right)dV$

The general problem: suppose we know $E = E_0\varphi(\vec{r})e^{i\omega t}$ on some surface S. Given φ on S, we want to know φ at some point P_0 inside S.

(φ, ψ = solutions of Helmholtz's eqn. in V)



S_ε = small surface enclosing P_0 (spherical)

V = volume bounded by S on outside and S_ε on inside.

Let $\psi\left(\overrightarrow{r_{01}}\right) = \dfrac{e^{-i\vec{k}\cdot\overrightarrow{r_{01}}}}{r_{01}}, \overrightarrow{r_{01}}$ measured from P_0.

This is our "Green's function." It is clearly a spherical wave centered on P_0.

In volume V, the field has no sources =>

$$\left(\nabla^2 + k^2\right)\psi = 0 \quad \left(\nabla^2 + k^2\right)\varphi = 0$$

⇨ $$\iiint_V \left(\varphi\nabla^2\psi - \psi\nabla^2\varphi\right)dV = \iiint_V \left(\varphi\psi k^2 - \psi\varphi k^2\right)dV = 0$$

⇨ $\iint_{S_\varepsilon} + \iint_S = 0$ (integral over total surface bounding V)

(*) $$-\iint_{S_\varepsilon}\left(\psi\frac{\partial\varphi}{\partial n} - \varphi\frac{\partial\psi}{\partial n}\right)ds = \iint_S\left(\psi\frac{\partial\varphi}{\partial n} - \varphi\frac{\partial\psi}{\partial n}\right)ds$$

First consider the integral over S_ε = small sphere of radius ε centered on P_0 :

$$\iint_{S_\varepsilon}\left(\varphi\frac{\partial\psi}{\partial n} - \psi\frac{\partial\varphi}{\partial n}\right)ds = \iint_{S_\varepsilon}\left[\varphi\frac{\partial}{\partial n}\left(\frac{e^{-i\vec{k}\cdot\overrightarrow{r_{01}}}}{r_{01}}\right) - \frac{e^{-i\vec{k}\cdot\overrightarrow{r_{01}}}}{r_{01}}\frac{\partial\varphi}{\partial n}\right]ds$$

Now $$\frac{\partial}{\partial n}\left(\frac{e^{-i\vec{k}\cdot\overrightarrow{r_{01}}}}{r_{01}}\right) = \nabla\left(\frac{e^{-i\vec{k}\cdot\overrightarrow{r_{01}}}}{r_{01}}\right)\cdot\hat{n} = -\left(-\frac{1}{r_{01}{}^2} - \frac{ik}{r_{01}}\right)e^{-i\vec{k}\cdot\overrightarrow{r_{01}}}$$

$$\hat{n} = -\overrightarrow{r_{01}}$$

If we are on S_ε , then $r_{01} = \varepsilon$, and we can introduce the solid angle element

$$d\Omega = \frac{ds}{\varepsilon^2}$$

Then

$$\iint_{S_\varepsilon}\left(\frac{1}{\varepsilon^2} + \frac{ik}{\varepsilon})e^{-ik\varepsilon}\varphi - \frac{e^{-ik\varepsilon}}{\varepsilon}\frac{\partial\varphi}{\partial n}\right)\varepsilon^2 d\Omega = \iint_{S_\varepsilon}\left[\left(1+ik\varepsilon\right)\varphi - \varepsilon\frac{\partial\varphi}{\partial n}\right]e^{-ik\varepsilon}d\Omega$$

Now let $\varepsilon \to 0$ (S_ε shrinks to an infinitesimal volume around P_0).The only nonzero term in the integral is $\iint_{S_\varepsilon}\varphi d\Omega = 4\pi\varphi(P_0)$

Thus we have ,from (*) Ｐ．３５７，

$$\varphi\left(P_0\right) = \frac{1}{4\pi}\iint_S\left[\frac{e^{-i\vec{k}\cdot\overrightarrow{r_{01}}}}{r_{01}}\frac{\partial\varphi}{\partial n} - \varphi\frac{\partial}{\partial n}\left(\frac{e^{-i\vec{k}\cdot\overrightarrow{r_{01}}}}{r_{01}}\right)\right]ds$$

This is the Helmholtz-Kirchhoff integral theorem.

It says we know the field at P_0 given its value (and thus its derivative $\frac{\partial\varphi}{\partial n}$ also) on any surface S containing P_0 .

Now suppose we have a point source at P_s , and we want the field at P_0 , but there is an opaque screen containing an aperture between them:

We take our surface S enclosing p_0 to be as shown: