Lecture 33

High Frequency Solutions, Gaussian Beams

When the frequency is very high, the wavelength of electromagnetic wave becomes very short. In this limit, the solutions to Maxwell's equations can be found approximately. These solutions offer a very different physical picture of electromagnetic waves, and they are often used in optics where the wavelength is short. So it was no surprise that for a while, optical fields were thought to satisfy a very different equations from those of electricity and magnetism. So it came as a surprise that when it was revealed that in fact, optical fields satisfy the same Maxwell's equations as the fields from electricity and magnetism!

In this lecture, we shall seek approximate solutions to Maxwell's equations or the wave equations when the frequency is high or the wavelength is short. High frequency approximate solutions are important in many real-world applications. This is possible when the wavelength is much smaller than the size of the structure. This can occur even in the microwave regime where the wavelength is not that small, but much smaller than the size of the structure. This is the case when microwave interacts with reflector antennas for instance. It is also the transition from waves regime to the optics regime in the solutions of Maxwell's equations. Often times, the term "quasi-optical" is used to describe the solutions in this regime.

In the high frequency regime, a plane wave can be approximated by a ray[1] which simplifies its solution in many instances. For instance, ray tracing can be used to track how these rays can propagate, bounce, or "richohet" in a complex environment. In fact, it is now done in a movie industry to give realism to simulate the nuances of how light ray will bounce around in a room, and reflecting off objects.

33.1 Tangent Plane Approximations

We have learnt that reflection and transmission of waves at a flat surface can be solved in closed form. The important point here is the physics of phase matching. Due to phase

[1]We shall learn in a later lecture why this is the case.

matching, we have the law of reflection, transmission and Snell's law [53].[2]

When a surface is not flat anymore, there is no closed form solution. But when a surface is curved, an approximate solution can be found. This is obtained by using a local tangent-plane approximation when the radius of curvature is much larger than the wavelength. Hence, this is a good approximation when the frequency is high or the wavelength is short. It is similar in spirit that we can approximate a spherical wave by a local plane wave at the spherical wave front when the wavelength is short compared to the radius of curvature of the wavefront.

When the wavelength is short, phase matching happens locally, and the law of reflection, transmission, and Snell's law are satisfied approximately as shown in Figure 33.1. The tangent plane approximation is the basis for the geometrical optics (GO) approximation [31, 181]. In GO, light waves are replaced by light rays. The reflection and transmission of these rays at an interface is estimated using the tangent plane approximation. This is also the basis for lens or ray optics from which lens technology is derived. It is also the basis for ray tracing for high-frequency solutions [182, 183].

Many real world problems do not have closed-form solutions, and have to be treated with approximate methods. In addition to geometrical approximations mentioned above, asymptotic methods are also used to find approximate solutions. Asymptotic methods imply finding a solution when there is a large parameter in the problem. In this case, it is usually the frequency. Such high-frequency approximate methods are discussed in [184–188].

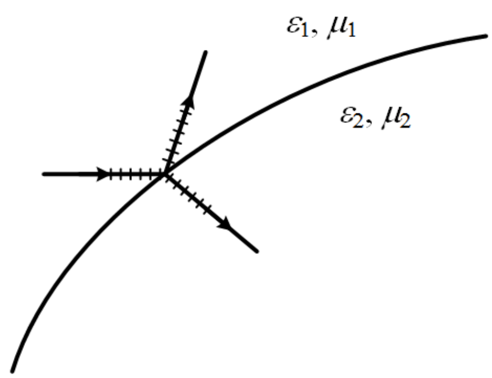

Figure 33.1: In the tangent plane approximation, the surface where reflection and refraction occur is assumed to be locally flat. Hence, phase-matching is approximately satisfied, and hence, the law of reflection, transmittion, and Snell's law are satisfied.

33.2 Fermat's Principle

Fermat's principle (1600s) [53,189] says that a light ray follows the path that takes the shortest time between two points.[3] Since time delay is related to the phase shift, and that a light ray

[2]This law is also known in the Islamic world in 984 [180].

[3]This eventually give rise to the principle of least action.

can be locally approximated by a plane wave, this can be stated that a plane wave follows the path that has a minimal phase shift. This principle can be used to derive law of reflection, transmission, and refraction for light rays. It can be used as the guiding principle for ray tracing.

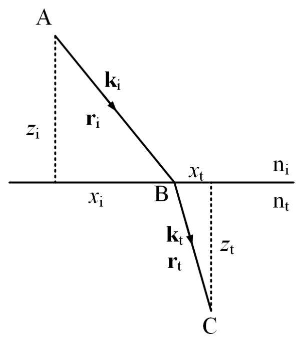

Figure 33.2: In Fermat's principle, a light ray, when propagating from point A to point C, takes the path of least delay.

Given two points A and C in two different half spaces as shown in Figure 33.2. Then the phase delay between the two points, per Figure 33.2, can be written as[4]

$$P = \mathbf{k}_i \cdot \mathbf{r}_i + \mathbf{k}_t \cdot \mathbf{r}_t \tag{33.2.1}$$

In the above, $\mathbf{k}_i$ is parallel to $\mathbf{r}_i$, so is $\mathbf{k}_t$ is parallel to $\mathbf{r}_t$. As this is the shortest path with minimum phase shift or time delay, according to Fermat's principle, another other path will be longer. In other words, if B were to move to another point, a longer path with more phase shift or time delay will ensue, or that B is the stationary point of the path length or phase shift. Specializing (33.2.1) to a 2D picture, then the phase shift as a function of x_i is stationary. In this Figure 33.2, we have $x_i + x_t = \text{const}$. Therefore, taking the derivative of (33.2.1) or the phase change with respect to x_i, assuming that $\mathbf{k}_i$ and $\mathbf{k}_t$ do not change as B is moved slightly, one gets[5]

$$\frac{\partial P}{\partial x_i} = 0 = k_{ix} - k_{tx} \tag{33.2.2}$$

The above yields the law of refraction that $k_{ix} = k_{tx}$, which is just Snell's law. It can also be obtained by phase matching. Notice that in the above, only x_i is varied to find the stationary point and $\mathbf{k}_i$ and $\mathbf{k}_t$ remain constant.

[4]In this course, for wavenumber, we use k and β interchangeably, where k is prevalent in optics and β used in microwaves.

[5]If we write $P = k_i r_i + k_t r_t$, and let $r_i = \sqrt{x_i^2 + z_i^2}$, and $r_t = \sqrt{x_t^2 + z_t^2}$, and take the derivative with respect to to x_i, one would also get the same answer.

33.2.1 Generalized Snell's Law

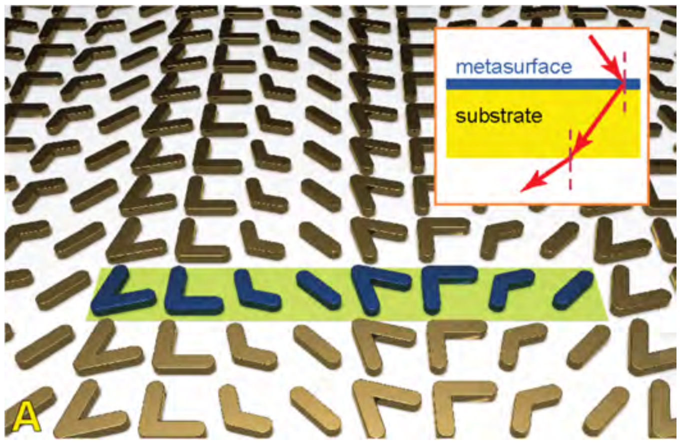

Figure 33.3: A phase screen which is position dependent can be made using nano-fabrication and design with commercial software for solving Maxwell's equations. In such a case, one can derive a generalized Snell's law to describe the diffraction of a wave by such a surface (courtesy of Capasso's group [190]).

Metasurfaces are prevalent these days due to our ability for nano-fabrication and numerical simulation. One of them is shown in Figure 33.3. Such a metasurface can be thought of as a phase screen, providing additional phase shift for the light as it passes through it. Moreover, the added phase shift can be controlled to be a function of position due to advent in fabrication technology and commercial software for numerical simulation.

To model this phase screen, we can add an additional function $\Phi(x, y)$ to (33.2.1), namely that

$$P = \mathbf{k}_i \cdot \mathbf{r}_i + \mathbf{k}_t \cdot \mathbf{r}_t - \Phi(x_i, y_i) \tag{33.2.3}$$

Now applying Fermat's principle that there should be minimal phase delay, and taking the derivative of the above with respect to x_i, one gets

$$\frac{\partial P}{\partial x_i} = k_{ix} - k_{tx} - \frac{\partial \Phi(x_i, y_i)}{\partial x_i} = 0 \tag{33.2.4}$$

The above yields that the generalized Snell's law [190] that

$$k_{ix} - k_{tx} = \frac{\partial \Phi(x_i, y_i)}{\partial x_i} \tag{33.2.5}$$

It yields the fact that the transmitted light can be directed to other angles due to the additional phase screen.

33.3 Gaussian Beam

We have seen previously that in a source free space

$$\nabla^2\mathbf{A} + \omega^2\mu\varepsilon\mathbf{A} = 0 \tag{33.3.1}$$

$$\nabla^2\Phi + \omega^2\mu\varepsilon\Phi = 0 \tag{33.3.2}$$

The above are four scalar equations with the Lorenz gauge

$$\nabla \cdot \mathbf{A} = -j\omega\mu\varepsilon\Phi \tag{33.3.3}$$

connecting $\mathbf{A}$ and Φ. We can examine the solution of $\mathbf{A}$ such that

$$\mathbf{A}(\mathbf{r}) = \mathbf{A}_0(\mathbf{r})e^{-j\beta z} \tag{33.3.4}$$

where $\mathbf{A}_0(\mathbf{r})$ is a slowly varying function while $e^{-j\beta z}$ is rapidly varying in the z direction. (Here, $\beta = \omega\sqrt{\mu\epsilon}$ is the wavenumber.) This is primarily a quasi-plane wave propagating predominantly in the z-direction. We know this to be the case in the far field of a source, but let us assume that this form persists less than the far field, namely, in the Fresnel as well. Taking the x component of (33.3.4), we have[6]

$$A_x(\mathbf{r}) = \Psi(\mathbf{r})e^{-j\beta z} \tag{33.3.5}$$

where $\Psi(\mathbf{r}) = \Psi(x, y, z)$ is a slowly varying envelope function of x, y, and z.

33.3.1 Derivation of the Paraxial/Parabolic Wave Equation

Substituting (33.3.5) into (33.3.1), and taking the double z derivative first, we arrive at

$$\frac{\partial^2}{\partial z^2}\left[\Psi(x,y,z)e^{-j\beta z}\right] = \left[\frac{\partial^2}{\partial z^2}\Psi(x,y,z) - 2j\beta\frac{\partial}{\partial z}\Psi(x,y,z) - \beta^2\Psi(x,y,z)\right]e^{-j\beta z} \tag{33.3.6}$$

Consequently, after substituting the above into the x component of (33.3.1), making use of the definition of ∇^2, we obtain an equation for $\Psi(\mathbf{r})$, the slowly varying envelope as

$$\frac{\partial^2}{\partial x^2}\Psi + \frac{\partial^2}{\partial y^2}\Psi - 2j\beta\frac{\partial}{\partial z}\Psi + \frac{\partial^2}{\partial z^2}\Psi = 0 \tag{33.3.7}$$

where the last term on the right-hand side of (33.3.6) cancels with the term coming from $\omega^2\mu\varepsilon\mathbf{A}$ of (33.3.1). When $\beta \to \infty$, or in the high frequency limit,

$$\left|2j\beta\frac{\partial}{\partial z}\Psi\right| \gg \left|\frac{\partial^2}{\partial z^2}\Psi\right| \tag{33.3.8}$$

[6] Also, the wave becomes a transverse wave in the far field, and keeping the transverse component suffices.

In other words, (33.3.7) can be approximated by

$$\frac{\partial^2\Psi}{\partial x^2}+\frac{\partial^2\Psi}{\partial y^2}-2j\beta\frac{\partial\Psi}{\partial z}\approx 0 \tag{33.3.9}$$

The above is called the paraxial wave equation. It is also called the parabolic wave equation.[7] It implies that the $\boldsymbol{\beta}$ vector of the wave is approximately parallel to the z axis, or β_z to be much greater than β_x and β_y, and hence, the name.

33.3.2 Finding a Closed Form Solution

A closed form solution to the paraxial wave equation can be obtained by a simple trick.[8] It is known that

$$A_x(\mathbf{r})=\frac{e^{-j\beta|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \tag{33.3.10}$$

is the solution to

$$\nabla^2 A_x+\beta^2 A_x=0 \tag{33.3.11}$$

if $\mathbf{r}\neq\mathbf{r}'$. One way to ensure that $\mathbf{r}\neq\mathbf{r}'$ always is to let $\mathbf{r}'=-\hat{z}jb$, a complex number. Then (33.3.10) is always a solution to (33.3.11) for all $\mathbf{r}$, because $|\mathbf{r}-\mathbf{r}'|\neq 0$ always. Then, we should next make a paraxial approximation to the solution (33.3.10) by assuming that $x^2+y^2\ll z^2$. By so doing, it follows that

$$\begin{aligned}|\mathbf{r}-\mathbf{r}'|&=\sqrt{x^2+y^2+(z+jb)^2}\\&=(z+jb)\left[1+\frac{x^2+y^2}{(z+jb)^2}\right]^{1/2}\\&\approx(z+jb)+\frac{x^2+y^2}{2(z+jb)}+\ldots,\qquad |z+jb|\to\infty\end{aligned} \tag{33.3.12}$$

And then using the above approximation in (33.3.10) yield

$$A_x(\mathbf{r})\approx\frac{e^{-j\beta(z+jb)}}{4\pi(z+jb)}e^{-j\beta\frac{x^2+y^2}{2(z+jb)}}=e^{-j\beta z}\Psi(\mathbf{r}) \tag{33.3.13}$$

By comparing the above with (33.3.5), we can identify

$$\Psi(x,y,z)=A_0\frac{jb}{z+jb}e^{-j\beta\frac{x^2+y^2}{2(z+jb)}} \tag{33.3.14}$$

where A_0 is used to absorbed the constant to simplify the expression. By separating the exponential part into the real part and the imaginary part, viz.,

$$\frac{x^2+y^2}{2(z+jb)}=\frac{x^2+y^2}{2}\left(\frac{z}{z^2+b^2}-j\frac{b}{z^2+b^2}\right) \tag{33.3.15}$$

[7]The paraxial wave equation, the diffusion equation and the Schrodinger equation are all classified as parabolic equations in mathematical parlance [34, 44, 191, 192].

[8]Introduced by Georges A. Deschamps of UIUC [193].

and writing the prefactor in terms of amplitude and phase, viz.,

$$\frac{jb}{z+jb} = \frac{1}{\sqrt{1+z^2/b^2}} e^{j\tan^{-1}\left(\frac{z}{b}\right)} \tag{33.3.16}$$

we then have

$$\Psi(x,y,z) = \frac{A_0}{\sqrt{1+z^2/b^2}} e^{j\tan^{-1}\left(\frac{z}{b}\right)} e^{-j\beta\frac{x^2+y^2}{2(z^2+b^2)}z} e^{-b\beta\frac{x^2+y^2}{2(z^2+b^2)}} \tag{33.3.17}$$

The above can be rewritten as

$$\Psi(x,y,z) = \frac{A_0}{\sqrt{1+z^2/b^2}} e^{-j\beta\frac{x^2+y^2}{2R}} e^{-\frac{x^2+y^2}{w^2}} e^{j\psi} \tag{33.3.18}$$

where A_0 is a new constant introduced to absorb undesirable constants arising out of the algebra, and

$$w^2 = \frac{2b}{\beta}\left(1+\frac{z^2}{b^2}\right), \qquad R = \frac{z^2+b^2}{z}, \qquad \psi = \tan^{-1}\left(\frac{z}{b}\right) \tag{33.3.19}$$

For a fixed z, the parameters w, R, and ψ are constants. It is seen that the beam is Gaussian tapered in the x and y directions. Hence, w is the beam waist which varies with z, and it is smallest when $z = 0$, or $w = w_0 = \sqrt{\frac{2b}{\beta}}$.

And the term $\exp(-j\beta\frac{x^2+y^2}{2R})$ resembles the phase front of a spherical wave where R is its radius of curvature. This can be appreciated by studying a spherical wave front $e^{-j\beta R}$, and make a paraxial wave approximation, namely, $x^2 + y^2 \ll z^2$ to get

$$\begin{aligned} e^{-j\beta R} = e^{-j\beta(x^2+y^2+z^2)^{1/2}} &= e^{-j\beta z\left(1+\frac{x^2+y^2}{z^2}\right)^{1/2}} \\ &\approx e^{-j\beta z - j\beta\frac{x^2+y^2}{2z}} \approx e^{-j\beta z - j\beta\frac{x^2+y^2}{2R}} \end{aligned} \tag{33.3.20}$$

In the last approximation, we assume that $z \approx R$ in the paraxial approximation.

The phase ψ changes linearly with z for small z, and saturates to a constant for large z. Then the phase of the entire wave is due to the $exp(-j\beta z)$ in (33.3.13). A cross section of the electric field due to a Gaussian beam is shown in Figure 33.4.

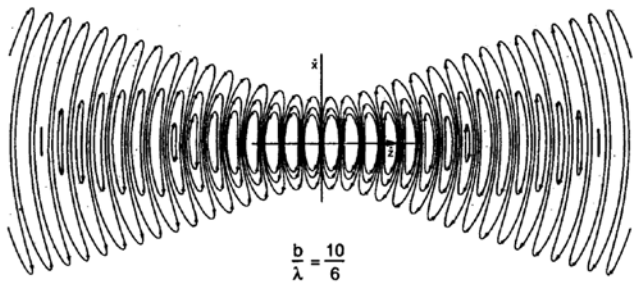

Figure 33.4: Electric field of a Gaussian beam in the x-z plane frozen in time. The wave moves to the right as time increases; here, $b/\lambda = 10/6$ (courtesy of Haus, Electromagnetic Noise and Quantum Optical Measurements [75]).

33.3.3 Other solutions

In general, the paraxial wave equation in (33.3.9) is a partial differential equation which can be solved by the separation of variables just like the Helmholtz wave equation. Therefore, in general, it has solutions of the form[9]

$$\Psi_{nm}(x,y,z) = \left(\frac{2}{\pi n! m!}\right)^{1/2} 2^{-N/2} \left(\frac{1}{w}\right) e^{-(x^2+y^2)/w^2} e^{-j\frac{\beta}{2R}(x^2+y^2)} e^{j(m+n+1)\Psi} \cdot H_n\left(x\sqrt{2}/w\right) H_m\left(y\sqrt{2}/w\right) \quad (33.3.21)$$

where $H_n(\xi)$ is a Hermite polynomial of order n. The solutions can also be express in terms of Laguere polynomials, namely,

$$\Psi_{nm}(x,y,z) = \left(\frac{2}{\pi n! m!}\right)^{1/2} \min(n,m)! \frac{1}{w} e^{-j\frac{\beta}{2R}\rho^2} - e^{-\rho^2/w^2} e^{+j(n+m+1)\Psi} e^{jl\phi} (-1)^{min(n,m)} \left(\frac{\sqrt{2}\rho}{w}\right) L^{n-m}_{\min(n,m)}\left(\frac{2\rho^2}{w^2}\right) \quad (33.3.22)$$

where $L_n^k(\xi)$ is the associated Laguerre polynomial.

These gaussian beams have rekindled recent excitement in the community because, in addition to carrying spin angular momentum as in a plane wave, they can carry orbital angular momentum due to the complex transverse field distribution of the beams.[10] They harbor potential for optical communications as well as optical tweezers to manipulate trapped nano-particles. Figure 33.5 shows some examples of the cross section (xy plane) field plots for some of these beams. They are richly endowed with patterns implying that they can be used to encode information. These lights are also called structured lights [195].

[9]See F. Pampaloni and J. Enderlein [194].

[10]See D.L. Andrew, Structured Light and Its Applications and articles therein [195].

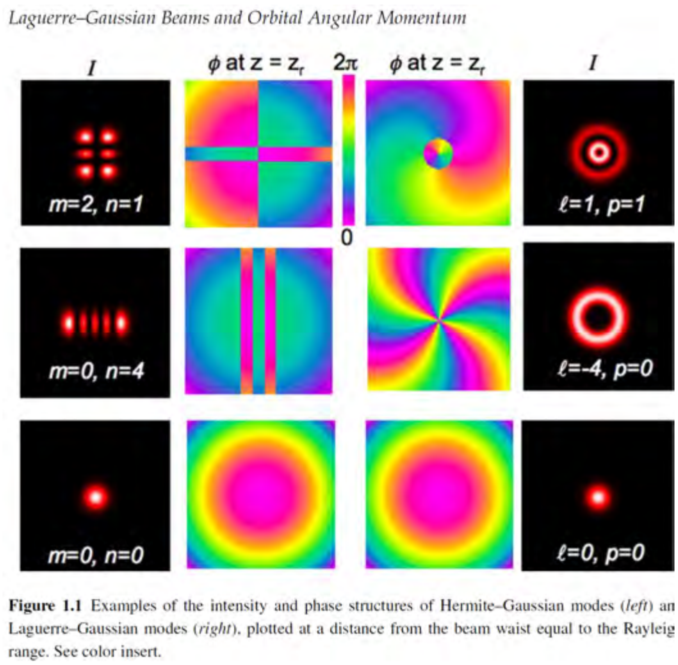

Figure 33.5: Examples of structured light. It can be used in encoding more information in optical communications (courtesy of L. Allen and M. Padgett's chapter in J.L. Andrew's book on structured light [195]).