Lecture 10

Diffraction

1 Introduction

It is quite often the case that no line-of-sight path exists between a cell phone and a basestation. In other words there are no basestations that the customer can see as they are all behind obstacles such as hills or buildings. Yet communication is typically still possible. How does a field make it from transmitter to receiver when there is no line-of-sight path? Usually the reason is the phenomenon of *diffraction* whereby a field can "bend" around obstacles.

Unlike reflection and path loss, that can be understood in geometrical terms, diffraction is a purely wave phenomenon. It is a strong function of wavelength. Long-wavelength (low-frequency) fields more readily "bend" than short-wavelength (high-frequency) fields. As the wavelength gets really small (e.g., light) diffraction becomes negligible and communication is typically possible only via line-of-sight paths.

Even "simple" diffraction problems are mathematically messy. We are going to analyze a "canonical" diffraction problem – *knife edge diffraction* – and use our findings to develop practical models that can be applied in more realistic scenarios.

The knife-edge diffraction problem consists of a plane wave interacting with a perfectly conducting infinite half plane (a "screen"). This is illustrated below.

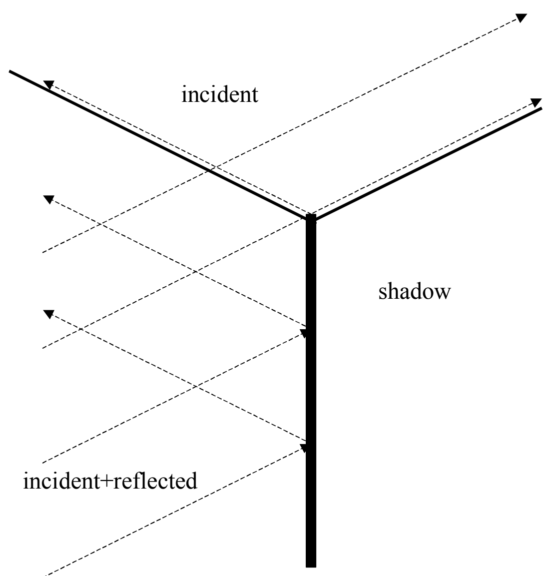

Fig. 1 Geometry of half-plane diffraction. Field is incident from the lower left. The conducting half-plane extends down the page infinitely far. The arrows represent the Poynting vector of the corresponding fields.

The plane wave is incident from the lower left of the figure. From a geometrical point of view where it hits the screen it is reflected, and where it misses the screen it continues on in a straight line. This results in space being divided into three regions: a region where both incident and reflected waves exist, a region where only the incident wave exists, and a shadow region where no wave exists. Because the field is a wave phenomenon, however, things are not quite this "clean." Most importantly, some of the field "leaks" into the shadow region. This difference between the real phenomenon and the straight-line geometry of Fig. 1 is what we refer to as diffraction. It is a very important phenomenon because if it did not exist it would be possible to receive a signal only when a direct line-of-sight path existed between transmitter and receiver. In a cellular phone system, especially in an urban environment, it is very often the case that users are in shadow regions where no line-of-sight path exists. In this case diffracted fields are typically the means by which communication takes place.

2 Simulation

Maxwell's equations can be solved numerically using a technique known as the finite-difference time-domain (FDTD) method. The following figure shows FDTD calculations of half-plane diffraction for four different wavelengths.

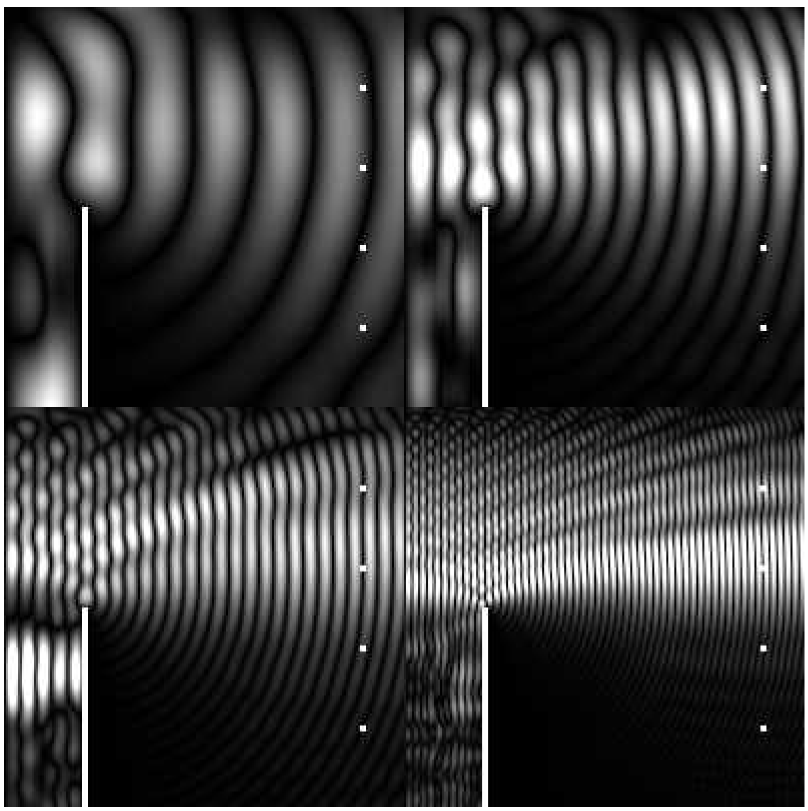

Fig. 2 Numerical calculations of diffraction by a half-plane for four different wavelengths. The plane wave is incident from the left.

In this case the field is incident directly from the left and the shadow region should be the region to the right of the screen. However, we can see that some of the field "leaks" into the shadow region. There is more leakage for the longer wavelengths than for the shorter wavelengths.

3 Theory

The idealized knife-edge diffraction problem is illustrated below. The theory is a bit mathematical and we will only give an overview of the development. (Refer to a course in optics or an optics text for more details.)

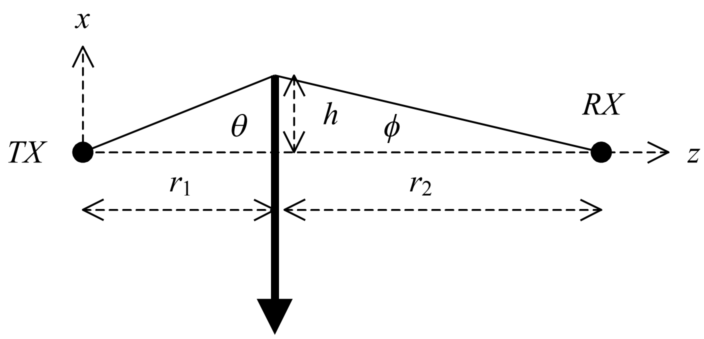

Fig. 3 Idealized knife-edge diffraction problem. The perfectly-conducting screen extends infinitely far down the page.

The transmitter (TX) will be the origin of our coordinate system and the line joining the transmitter and receiver (RX) will serve as our z axis. The screen is a distance r_1 from the TX, a distance r_2 from the RX, and it extends a height *h* above the *z* axis. If *h* is positive, RX is in the "shadow region" while if *h* is negative it is in the "incident region." The TX transmits a spherical wave described by

$$E_T = \frac{e^{-j\frac{2\pi}{\lambda}\sqrt{x^2+y^2+z^2}}}{-j\lambda\sqrt{x^2+y^2+z^2}} \approx \frac{e^{-j\frac{2\pi}{\lambda}z}}{-j\lambda z} e^{-j\frac{\pi}{\lambda z}(x^2+y^2)} \tag{1}$$

The approximation used to get the last expression is that $\sqrt{x^2+y^2+z^2} \approx z + (x^2+y^2)/(2z)$ which is valid when $z^2 \gg x^2+y^2$. At $z = r_1$ the field is zero for $x < h$ due to the screen. We will assume that for $x > h$ the field is simply the spherical wave (1). In other words the screen simply "cuts off" a part of the spherical wave. This is not exactly true – there are some edge effects – but it's a good approximation. Now we need to figure out how this truncated field propagates out to the RX.

Electromagnetic theory tells us that each point on the wavefront at $z = r_1$ can be considered as the source of a new spherical wave with amplitude given by the field amplitude at that point. (This concept is called *Huygen's principle*.) We then simply sum the contributions of all those spherical waves at the RX to get the received field. The result is given by the following equation:

$$\begin{aligned} E_R &= \int_{-\infty}^{\infty}\int_{h}^{\infty} \left[\frac{e^{-j\frac{2\pi}{\lambda}r_1}}{-j\lambda r_1} e^{-j\frac{\pi}{\lambda r_1}(x^2+y^2)} \right] \frac{e^{-j\frac{2\pi}{\lambda}r_r}}{-j\lambda r_r} e^{-j\frac{\pi}{\lambda r_2}(x^2+y^2)} \, dx\, dy \\ &= -\frac{e^{-j\frac{2\pi}{\lambda}(r_1+r_2)}}{\lambda^2 r_1 r_2} \int_{-\infty}^{\infty} e^{-j\frac{\pi}{\lambda}\left(\frac{1}{r_1}+\frac{1}{r_2}\right)y^2} dy \int_{h}^{\infty} e^{-j\frac{\pi}{\lambda}\left(\frac{1}{r_1}+\frac{1}{r_2}\right)x^2} dx \end{aligned} \tag{2}$$

Looking at the first line, the expression in the brackets is just the transmitted spherical wave at $z = r_1$. This then acts as the amplitude of a new spherical wave that then propagates a distance $\sqrt{x^2+y^2+r_2^2} \approx z + (x^2+y^2)/(2r_2)$ to the RX. That spherical wave accounts for the rest of integrand. We then sum up – integrate – all these contributions. The field is zero for $x<h$, so our integral over x is from h to infinity. In the second line we rearrange things to put it in a more convenient form.

The *y* integral can be done in closed form to get

$$\int_{-\infty}^{\infty} e^{-j\frac{\pi}{\lambda}\left(\frac{1}{r_1}+\frac{1}{r_2}\right)y^2}dy=(1-j)\sqrt{\frac{\lambda}{2}\left(\frac{r_1 r_2}{r_1+r_2}\right)} \tag{3}$$

With the change of variable $t=x\sqrt{\frac{2}{\lambda}\left(\frac{1}{r_1}+\frac{1}{r_2}\right)}$ the *x* integral becomes

$$\int_{h}^{\infty} e^{-j\frac{\pi}{\lambda}\left(\frac{1}{r_1}+\frac{1}{r_2}\right)x^2}dx=\sqrt{\frac{\lambda}{2}\left(\frac{r_1 r_2}{r_1+r_2}\right)}\int_{u}^{\infty} e^{-j\frac{\pi}{2}t^2}dt \tag{4}$$

where

$$u=h\sqrt{\frac{2}{\lambda}\left(\frac{r_1+r_2}{r_1 r_2}\right)} \tag{5}$$

is called the *diffraction parameter*. Now the received field can be written

$$E_R=\frac{e^{-j\frac{2\pi}{\lambda}r}}{-j\lambda r}\frac{1+j}{2}\int_{u}^{\infty} e^{-j\frac{\pi}{2}t^2}dt \tag{6}$$

with $r=r_1+r_2$. In the absence of the screen the received field would simply be the transmitted field $e^{-j\frac{2\pi}{\lambda}r}/(-j\lambda r)$. With the screen present we have

$$\frac{E_R}{E_T}=\frac{1+j}{2}\int_{u}^{\infty} e^{-j\frac{\pi}{2}t^2}dt \tag{7}$$

We define the *diffraction gain* as the power ratio of the screen present to the screen absent cases. This is

$$G_d=\left|\frac{E_R}{E_T}\right|^2=\frac{1}{2}\left|\int_{u}^{\infty} e^{-j\frac{\pi}{2}t^2}dt\right|^2 \tag{8}$$

It is usually convenient to write

$$\begin{aligned}\int_{u}^{\infty} e^{-j\frac{\pi}{2}t^2}dt&=\int_{0}^{\infty} e^{-j\frac{\pi}{2}t^2}dt-\int_{0}^{u} e^{-j\frac{\pi}{2}t^2}dt\\&=\frac{1}{2}(1-j)-[C(u)-jS(u)]\end{aligned} \tag{9}$$

where

$$
\begin{aligned}
C(u) &= \int_0^u \cos\left(\frac{\pi}{2}t^2\right)dt \\
S(u) &= \int_0^u \sin\left(\frac{\pi}{2}t^2\right)dt
\end{aligned} \tag{10}
$$

are called the *Fresnel integrals* and are plotted below.

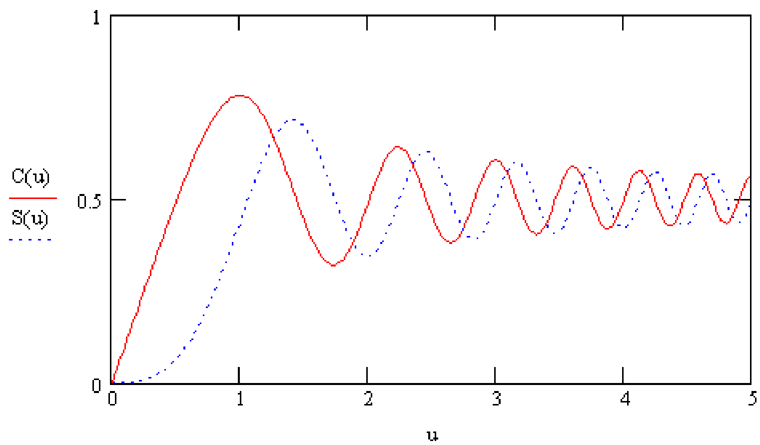

Fig. 4 Numerical calculation of the Fresnel integrals.

We then have

$$
G_d(u) = \frac{1}{2}\left[\left(\frac{1}{2} - C(u)\right)^2 + \left(\frac{1}{2} - S(u)\right)^2\right] \tag{11}
$$

We see that the diffraction gain is a function of the diffraction parameter u alone. For large values of u the Fresnel integrals behave as

$$
\begin{aligned}
C(u) &\rightarrow \frac{1}{2} + \frac{1}{\pi u}\sin\left(\frac{\pi}{2}u^2\right) \\
S(u) &\rightarrow \frac{1}{2} - \frac{1}{\pi u}\cos\left(\frac{\pi}{2}u^2\right)
\end{aligned} \tag{12}
$$

These expressions are accurate to better than 2% for $u > 2$. An approximation good to better than 1.5 dB is

$$
G_{d,\text{dB}} = \begin{cases} 0 & u \le -1 \\ -\dfrac{13}{2}(u+1) & -1 < u < 1 \\ 20\log\dfrac{1}{\sqrt{2}\pi u} & u \ge 1 \end{cases} \tag{13}
$$

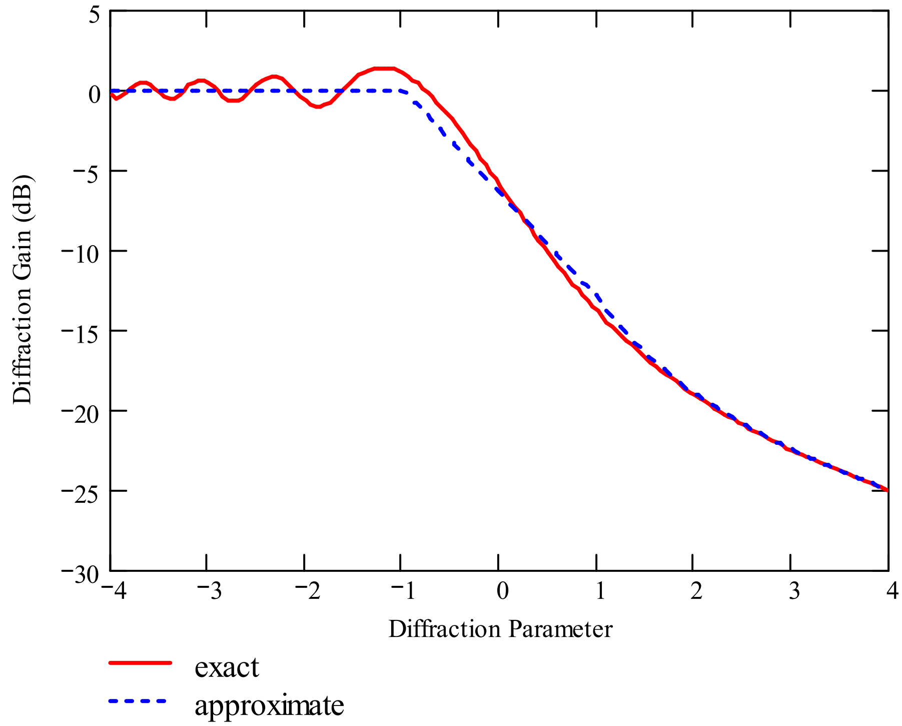

Fig. 5 Comparison of exact and approximate diffraction gain expressions.

Note that in Fig. 3, the angles θ and ϕ are given by $\tan\theta = h/r_1$, $\tan\phi = h/r_2$. If we call the distance between TX and RX $r = r_1 + r_2$, then we can write the diffraction parameter as

$$\begin{aligned} u &= h\sqrt{\frac{2}{\lambda}\left(\frac{r_1+r_2}{r_1 r_2}\right)} \\ &= \pm\sqrt{\frac{2r}{\lambda}\frac{h}{r_1}\frac{h}{r_2}} \\ &= \pm\sqrt{\frac{2r}{\lambda}\tan\theta\tan\phi} \end{aligned} \tag{14}$$

We use the "+" if $h, \theta \geq 0$ and "−" otherwise.

4 References

1. Born, M. and E. Wolf. *Principles of Optics*. Pergamon Press 1980. ISBN 0-08-026481-6. See section 8.7 "Fresnel Diffraction at a Straight Edge."