6.014 Lecture 6: Multipath, Arrays, and Frequency Reuse

A. Superposition of phasors

This lecture focuses on the superposition of duplicate waves at receivers, where the multiplicity of waves may have originated from multiple reflectors in the environment or from multiple transmitting antenna elements. Superposition of waves is easiest to understand when only one narrow band is considered at a time; we approximate such bands here as pure monochromatic sinusoids. The simplest case is illustrated in Figure L6-1, where the waves A and B are duplicates and superimpose in phase to yield A+B with double amplitude and quadruple power, and superimpose 180^o out of phase to yield A+C with zero amplitude and zero power. When these two equal-amplitude waves superimpose 90^o out of phase, we obtain A+D with double power and amplitude $2^{0.5}$A. Equation L6-1 shows how two equal-amplitude sinusoids with phase offset ϕ combine to yield a double-amplitude wave at the same ω, but phase-shifted by $\phi/2$ and multiplied by the constant $\cos(\phi/2)$, which can be >0, 0, or <0. For our case A+B, $\phi = 0$ and we produce a double-amplitude wave. For A+C, $\phi = 180^o$ and $\cos(\phi/2) = 0$; for A+D, $\phi = 90^o$ and $\cos(\phi/2) = 2^{-0.5}$.

A convenient way to think about such superposition of waves is in terms of phasors $\underline{E}$ characterized by their real and imaginary parts, as suggested graphically in Figure L6-2, where phasors represent the waves A, B, C, D, and A+D. The physical significance of the phasor $\underline{E}$ is defined by: $E(t) = \mathrm{Re}\{\underline{E}\, e^{j\omega t}\}$. The significance of the real and imaginary parts of $\underline{E}$ follow from

$$E(t) = \mathrm{Re}\{\underline{E}\, e^{j\omega t}\} = \mathrm{Re}\{[\mathrm{Re}\{\underline{E}\} + j\,\mathrm{Im}\{\underline{E}\}][\cos \omega t + j \sin \omega t]\} \qquad (1)$$

The real part of $\underline{E}$ thus corresponds to the amplitude of the $\cos \omega t$ term, and the imaginary part corresponds to $-\sin \omega t$. This correspondence is consistent with the phasors plotted in the bottom figure.

We can also represent the phasor $\underline{E}$ by its equivalent:

$$\underline{E} = |\underline{E}|e^{j\phi} = |\underline{E}|\cos \phi + j|\underline{E}| \sin \phi = \mathrm{Re}\{\underline{E}\} + j\,\mathrm{Im}\{\underline{E}\} \qquad (2)$$

where ϕ is the angle in the figure between the real axis and the phasor. Thus the phasor $\underline{E}e^{j\omega t}$ rotates counter-clockwise as time advances (see the direction of the arrow for ϕ in the figure). A phasor $\underline{E}$ that has been delayed θ radians would be represented by $|\underline{E}|e^{-j\theta}$.

B. Antenna arrays

To achieve desired antenna directional characteristics, either arrays, lenses, or reflectors are commonly used. Arrays usually consist of a set of duplicate small antennas, each located differently but often with the same orientation. The amplitudes and phases of the currents with which they are driven can be different. For example, if a single reference transmitting element "i" of the antenna array driven by current $\underline{a}_i$ produces the electric field $\underline{a}_i\mathbf{E}_i \exp\{-jkr_i\}$ at distance r_i (note: boldface indicates vectors here), then the total electric field in that direction θ,ϕ and at that distance r is:

$$\underline{\mathbf{E}}(r, \theta, \phi) = \Sigma_i\, \underline{a}_i\, \mathbf{E}_i \exp\{-jkr_i\} \tag{3}$$

If all elements are identical and oriented the same, then $\mathbf{E}_i = \mathbf{E}$, where $\mathbf{E}$ characterizes the basic radiating element, and is called the "element factor".

$$\underline{\mathbf{E}}(r, \theta, \phi) = \mathbf{E}\,(\Sigma_i\, \underline{a}_i \exp\{-jkr_i\}) = (\text{element factor } \mathbf{E})(\text{array factor}) \tag{4}$$

where the array factor characterizes the spatial distribution of radiating elements and the amplitudes and phases of the currents with which they are excited.

Consider the antenna pattern that results from two vertical (z-directed) dipole antennas arranged $\lambda/2$ apart along the y axis, as illustrated in Figure L6-2. Clearly the radiation from these two dipoles arrives in phase at receivers anywhere in the x-z plane, and the two beams cancel anywhere along the y axis. The pattern in the x-y plane is sketched in the same figure, and exhibits the expected maximum along the x axis and perfect null along the y axis. At the angle $\phi = \sin^{-1}0.5$ (from the x axis) the two rays arrive $\lambda/4$ out of phase, which results in half the power available at the maximum (see Figure L1-1(bottom)). This is the array factor. The element factor in the x-y plane is simply a circle, as illustrated, because a vertical dipole is isotropic in its equatorial plane. The antenna pattern in the x-y plane that is produced when these two dipoles are excited 180^o out of phase is also illustrated, and it is again clear that the two rays will now cancel along the x axis and add perfectly along the $\pm$y axis. The half-power angle $\phi = \sin^{-1}0.5$ remains the same, and the two lobes of the antenna pattern are now circles rather than resembling ellipses.

Figure L6-3A shows the 8-lobe pattern that results in the x-y plane when these two z-directed dipoles are arranged along the x axis 2λ apart. Clearly the two rays add in phase along both the x and y axes, and reach a maximum at another angle $\phi = \cos^{-1}0.5$. There are perfect nulls between the maxima because the two rays have equal magnitude. One such null angle is illustrated: $\theta = \cos^{-1}(1.5\lambda/2\lambda)$. A more interesting pattern results when the two dipoles are $\lambda/4$ apart and excited 90^o out of phase, as illustrated in Figure L6-3B. The two rays add coherently along the +x axis, and the rays cancel along the –x axis because the two 90^o phase shifts add in that direction. The half-power direction is along the $\pm$y axis because there the relative phase difference between the two rays is 90^o.

When the two identically oriented dipoles are excited unequally, they can never produce a null, no matter what the relative phase, because two unequal phasors can not cancel perfectly. Figure L6-3C illustrates this case where two out-of-phase dipoles $\lambda/2$ apart add coherently along the x axis, but can not perfectly cancel along the $\pm$y axis. Figure L6-3D shows the pattern from a linear array D meters long that is uniformly excited—all elements are in phase with equal amplitude currents. Clearly all phasors add coherently to produce a maximum along the ±x axes in the x-y plane. The first null is readily found if there is an even number of elements, because we can group them in pairs that, in the direction θ_{null} of the first null, are $\lambda/2$ out of phase and therefore cancel. All such offset pairs cancel in this same direction, and therefore the entire antenna produces a null in that direction. In the figure the first and fourth elements cancel in the direction $\theta_{firstnull} = \sin^{-1}[(\lambda/2)/(3D/5)] \Rightarrow \sim\sin^{-1}[(\lambda/2)/(D/2)] = \sin^{-1}(\lambda/D) \cong \lambda/D$ radians for large values of D/λ. Similarly the second and fifth, and the third and sixth elements cancel in that same direction.

One way to produce the equivalent of a second radiating element is to introduce a mirror that produces an image of the source, as illustrated in Figure L6-4E; the image is 180^{o} out-of-phase. Mirror images will be discussed further later.

C. Multipath

Multipath originates when a transmitter radiating in all directions produces reflections from objects like buildings and trees which arrive at the receiver with independent amplitudes and phases so as to interfere constructively or destructively. Because the rays may reflect from objects that alter the polarization, the powers received on two orthogonal polarizations typically vary somewhat independently. Monochromatic signals exhibit fading, the statistics of which depend on the time variations along the various paths. If the line-of-sight path is clear, then the reflections typically cause only minor fluctuations in strength. Urban cellular phones often have no line of sight, so only reflections and diffraction provide signal, and multipath can then produce deep fading.

The time constant characterizing such fading depends on the rate of change of the various paths relative to $\lambda/2$. The longer the paths relative to a wavelength, the smaller the fractional change in length required to accomplish this $\lambda/2$ drift, and the faster the fades. Besides the obvious fading experienced as cellar phones enter tunnels or elevators, there is also the fading of FM radio signals as automobiles move through marginal reception areas. For example, the sharp threshold of FM signals between good reception and static makes such radios an excellent detector of signal nulls. It is not unusual in a city to have only multipath FM reception dominated by only two or three rays of comparable magnitude. In this case, as the automobile inches forward, perhaps at a traffic light, the wavelength is evident in the distance $\sim\lambda$ that the automobile moves between transitions to static. At higher automobile speeds this effect is manifest as quasi-periodic clicks in the FM signal. Simple geometric considerations reveal this distance as a function of the directions of arrival of the interfering beams.

If the transmitter, receiver, or mid-path reflector is moving, then there can also be a small Doppler shift in frequency f_D:

$$f_D \text{ Hz} = (dL/dt)/\lambda \text{ cycles per second} = v/\lambda = f_o v/c \text{ Hz, so} \qquad (5)$$

$$f_D = f_o(1 - v/c) \text{ Hz} \tag{6}$$

When the doppler shift is upward we sometimes say the signal is "blue-shifted", and when the shift is downward, "red-shifted"; these terms have an astronomical origin and refer to apparent color shifts in celestial objects approaching or moving away from the earth.

Most signals of interest are not monochromatic, however, and occupy some bandwidth that may be affected differently by multipath at different frequencies. Consider two rays that interfere at the receiver and have pathlengths that differ by D [m]. Then frequencies near f_o separated by Δf can both experience nulls if $D/\lambda = f_o/\Delta f$.

In general, we can represent a multipath environment as a linear system with multiple delayed impulse responses, as suggested in Figure L6-5. For example, the system frequency response $\underline{H}(f)$ for to a system impulse response h(t) corresponding to two equal amplitude signals delayed by t_1 and t_2 is:

$$\underline{H}(f) = \int_{-\infty}^{+\infty} [\delta(t - t_1) + \delta(t - t_2)]\, e^{-j\omega t}\, dt = e^{-j\omega t1} + e^{-j\omega t2} = \\ = e^{-j\omega (t1 + t2)/2}\, [e^{j\omega (t1 - t2)/2} + e^{-j\omega (t1 - t2)}], \text{ and} \tag{7}$$

$$|\underline{H}(f)|^2 = [2\cos(\omega[t_1 - t_2]/2)]^2 \tag{8}$$

This yields nulls when $\omega_n[t_1 - t_2]/2 = (2n + 1)\pi/2$, and therefore nulls occur at frequencies $f_n = \omega_n/2\pi = (n + ½)/(t_1 - t_2)$, and therefore for two paths corresponding to delays of t_1 and t_2 seconds, the Δf between nulls is:

$$\Delta f = 1/(t_2 - t_1) \text{ Hz} \tag{9}$$

E. Frequency reuse

Many communications systems are seriously limited by the available bandwidth for wireless communications. The over-the-air spectrum must be shared and it has finite width. The frequencies most favored are below 1 GHz because thye diffract around objects better, but this bandwidth could not begin to satisfy current demand without reusing this bandwidth many times. The simplest form of reuse occurs when, for example, the same frequency is allocated by the Federal Communications Commission (FCC) only to radio or TV stations that are separated more than a hundred miles or so. More powerful transmitters with taller antennas must be spaced farther apart than weak stations. Moreover, poor engineering practices currently mandate that channels adjacent to allocated TV channels be kept vacant because of out-of-band interference. Hence Boston has VHF TV channels 2, 4, 5, and 7, but not 3, 6, or 8 (channels 4 and 5 are not adjacent in frequency). Out-of-band interference is not severe, so adjacent cities will usually use the alternate channels.

Cellular telephone base stations often utilize array antennas to achieve frequency reuse. Figure L6-6A illustrates a typical face of a cellular base station, with 3 or 4 elements and a combining circuit that forms the various desired beams. Three such faces arranged in a triangle, as seen from top view in Figure L6-6B, might produce for example, two sets of antenna lobes—the A set and the B set. Since these two sets overlap in certain directions, they would typically operate within two different sub-bands within the allocated bandwidth. Some users could then use both bands, and others could use only one. Since the different faces of the antenna can be connected to different receivers and transmitters, the same frequency could then be used by three different users simultaneously. Clearly the design of such antennas to maximize reuse requires some thought and could be tailored to the distribution of users within the local environment.

Another form of frequency re-use is employed for satellite communications systems where the antenna in space has multiple beams pointed at different places across the globe. Densely populated areas are generally served by smaller antenna beams so fewer users have to share its frequency allocation. The same frequencies can then be reused in another antenna beam that is not adjacent. Figure L6-7A illustrates a few such beams in North America, and Figure L6-7B illustrates how three arrays of beams are sufficient to provide full coverage without adjacent beams overlapping. That is, the degree of reuse can be ~one-third the number of antenna beams.

Satellite communications systems below ~1-2 GHz are bothered by variable polarization rotation in the ionosphere due to faraday rotation, motivating the use of circular polarization to minimize such effects. Above ~6 GHz rain attenuation can prevent communications from time to time, motivating more powerful links with greater signal-to-noise ratio margins, or spatial diversity protected by redundant links using multiple ground stations.

F. Wave interference for lithography

Figure L6-8A suggests how lithography of silicon wafers requires delicate masks through which light shines to alter photoresist in patterns that can be etched away to create integrated circuits. Recently the requirements for such masks are so severe that interference patterns are sometimes used to create the desired result. This is particularly simple if only periodic gratings are desired. For example, an excimer laser operating at a standard wavelength of 148 nanometers (0.14 microns) can be made to interfer with itself at large angles of incidence, producing strong nulls spaced approximately $\lambda/2$, or ~74 nm. This patterned light can then expose the photoresist, leading to a pattern of periodic stripes. Doing this in two dimensions can produce arrays of small pillars that can each code one bit of information magnetically, and therefore form a memory with ~2.3 GB/cm^2.