

Chapter 6: The Diffraction Efficiency of Gratings

For most of our discussions in this course, we will be worrying about where the light goes. But every once in a while, we have to worry about how much light gets there too. After all, no hologram means much if it is so dim that no one can see it! This brings us to the subject of *diffraction efficiency*, what determines it, and how it might be maximized.

We begin by considering the diffraction efficiency of a few simple gratings, for the cases of absorbing and phase-retarding gratings (unbleached and bleached holograms), gradually developing some general rules. This chapter can offer only a very preliminary pass at understanding the theoretical underpinnings of the brightness and contrast of holographic images, but it will be a good start!

Thin Gratings:

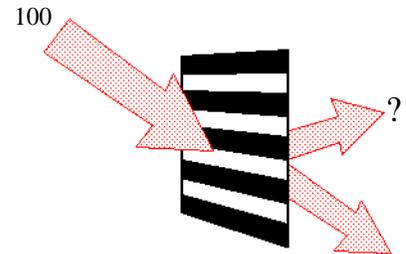
At the outset, we have to add that all of this chapter's remarks will be limited to "thin" gratings and holograms. That is to say that the thickness of the diffracting structures is small enough, compared to the grating spacing, d , and that angular selectivity effects (the Bragg angle effects¹) are not significant. This situation is often described by the "thickness parameter," Q (which is also a function of the incident illumination angle, θ_0 , and the wavelength, λ) being much less than unity (strong modulations of absorption or refractive index can also increase the apparent thickness of the grating). If the hologram thickens, the diffraction efficiency generally increases when it is properly angled, but the theories describing these conditions become very complicated. We will visit this special domain when we talk about reflection holograms, later in the term.

the thickness parameter :

$$Q = \frac{2\pi \cdot \lambda \cdot \text{thickness}}{n \cdot d^2 \cdot \cos \theta_0} \ll 1$$

Definition of Diffraction Efficiency

Usually, we mean by diffraction efficiency (designated as DE_{+1} , or η_{+1} "eta - plus-one") the ratio of the intensities of the desired (usually the plus-first order) diffracted beam and the illuminating beam, measured when both beams are large enough to overfill the area of the detector being used. We ignore any power losses due to reflections at the surfaces of the grating or hologram; practical DE measurements also take these surface reflections (approximately 4%/surface, for uncoated glass) into account. This chapter's simple mathematical models overlook polarization effects, large-diffraction-angle effects, and many other subtleties of rigorous electromagnetic theory, but their conclusions can generally be extended into those domains by more detailed math, so these results serve as useful guides nonetheless.



The local intensity of a beam of light is proportional to the time-average of the square of the magnitude of its electric field. Thus the ratio of the intensity of the output and input beams is equal to the square of the magnitude of the electric field transmittance of the hologram. Because the output field consists of several beams that eventually separate, we are interested in accounting for them one by one. That means breaking the transmittance pattern down into components that correspond to each of the beams. Finding the amplitude of those transmittance components is the principal concern of the rest of this chapter.

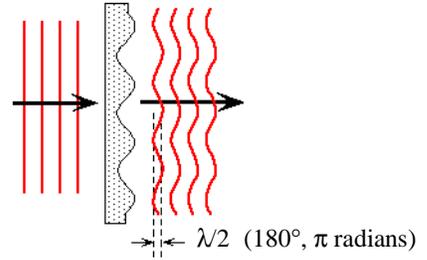
Transmission Patterns

We describe a grating or hologram by its two-dimensional *transmittance* pattern. By transmittance we mean the ratio of the electrical wave fields just after and just before the grating at the same x,y location. In the simplest case, the wave is simply attenuated, so that its electric field amplitude diminishes. This is called an *amplitude* transmittance grating.

Note that this is **not** the transmittance that one might usually think of measuring with a photographic light meter, because a light meter responds to the *intensity* of a beam, which is proportional to the *square* of the wave amplitude!

$$t_{\text{amp}}(x, y) = \frac{E_{\text{out}}(x, y)}{E_{\text{in}}(x, y)}$$

In so-called *coherent* optical systems (where the illumination is monochromatic and from a point-like source, generally from a laser), the delays that a wave encounters in passing through a grating are also very important. For example, if they are great enough to retard the wave by half a cycle in some places, then those waves would cancel out waves from other parts! Retarding effects are described as variations of *phase* transmittance, with the amount of delay being measured in degrees or radians, or sometimes in terms of wavelengths (or fractions thereof). Phase delays can be caused by variations in either the local microscopic thickness of a grating (with the phase increasing with increasing thickness) or the local refractive index, n , or both.

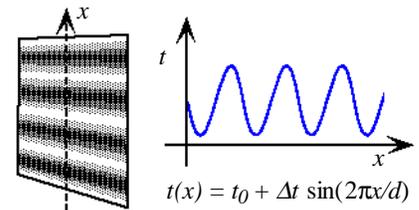


Usually, amplitude and phase transmittance variations, or modulations, are linked together in practical cases, but it is useful to first think of them as separate cases, with unbleached holograms being amplitude-only gratings, and bleached holograms being phase-only gratings.

For this discussion, the modulation patterns will vary with x only; that is, the patterns will be horizontal, with the “furrows” extending out of the page. The attenuation or phase delay variations of the pattern will be described as a graph and/or as an analytical expression, such as shown here. A few simple cases can help us find guidelines that will predict the behavior of a wider variety of holograms. However, true speckley-object-beam holograms (such as of 3-D objects) have a random transmittance pattern that requires a more complex analysis. The patterns we are talking about now are for “gratings” that have no randomness—images of single points, in effect.

Sinusoidal transmittance grating:

Here, the amplitude-only transmittance is a perfectly-smoothly-varying sinusoidal function of position, an ideal simplest case. Such a pattern could be produced by a low-contrast interference pattern exposure, for example. The most striking geometrical property of diffraction by such a grating is that there are only two output beams, the $m= +1$ and $m= -1$ orders, on either side of the straight-through $m= 0$ beam. In the far-field pattern, we see only one spot of light on each side of the zero-order beam for each sinusoidal component of the grating’s amplitude transmittance; it acts as a kind of Fourier transformer! This property can readily be proven by matching the amplitudes and phases of sets of waves on both sides of the grating (boundary condition matching).



The intensity of each of the $m= \pm 1$ beams varies as the square of the amount of swing of the sinusoidal modulation, and is given by the formula to the right. Because the transmittance has to stay between zero and one, the maximum value of Δt is 0.5 (only possible if $t_0=0.5$ also), and the maximum value of the diffraction efficiency is then 6.25%. Four times that amount emerges in the straight-through beam, and the rest is absorbed in the grating, gently warming it. This low maximum DE is not very encouraging for the brightness of display holograms! Unbleached holograms can be bright enough to be impressive under controlled lighting conditions, but it is usually quite difficult to consistently produce the maximum possible diffraction efficiency.

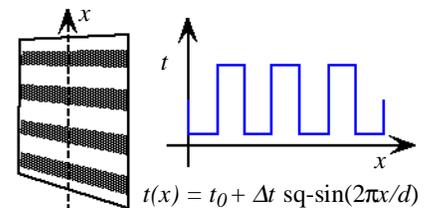
$$DE_{m=0} = t_0^2$$

$$DE_{m=\pm 1} = \left(\frac{\Delta t}{2}\right)^2 = 6.25\%_{\max}$$

$$DE_{|m|>1} = 0$$

Square-wave transmittance grating

Often, there are non-linearities in the exposure response of photographic materials that distort the purely sinusoidal nature of a transmittance pattern, in much the way that “fuzz boxes” can distort electrical guitar sounds. An extreme is a “hard-clipped” sine wave, which we will refer to here as a square wave or “squared-up sine wave” (that is, it is “high” 50% of the time, and “low” the other 50%) denoted as sq-sin θ (unpronounceable).



Such a grating can be considered as a summation of many ideal sinusoidal gratings, one with the same period as the square wave, and then gratings with integer fractions of that period (or multiples of that spatial frequency—the “higher harmonics,” one might say). Each sinusoid diffracts two beams of light, so that many points of light now appear in a straight line alongside the straight-through beam. But in spite of the energy going into the extra beams, the first-

order beams are brighter than before! This is because the “fundamental sinusoidal component” of a square wave has a magnitude that is larger than the magnitude of the square wave itself by a factor of $4/\pi$. So, we get transmittance values greater than unity and less than zero for that particular grating component, a physical paradox.

The application of Fourier theory produces these predictions of the diffraction efficiency: more light in the first order image, by 62% (when $\Delta t = 0.5$) giving over ten percent diffraction efficiency, plus some higher orders. Note that there are no *even* orders though; this depends on the grating being exactly 50/50 open/closed. Where does the extra total energy come from? Only one-half of the grating is dark, in the highest *DE* (black/clear) case, and therefore only 50% of the total energy gets absorbed, versus 62.5% in the sinusoidal case. So, it looks like non-linearities can work to our advantage! Unfortunately, in more complex images non-linearities produce noise images that strongly degrade the desired image.

Square-Wave Phase Grating

One reason for dwelling on the square-wave grating is that it offers a good introduction to simple phase-only gratings. Such gratings work by retarding the wavefronts as a function of position, and the results are hard to analyze for most modulation shapes. But if the grating comprises only two phase levels, such as 0 and π , the results follow from the same analysis used for square-wave amplitude-only gratings.

Phase-only gratings absorb no light energy, so that the total amount of diffracted light can reach 100% when summed over all the orders. The modulation possible for the fundamental transmittance component becomes twice what it was in the amplitude-only transmission case, ranging from +1 to -1 in effect (when $\Delta\phi = \pi/2$), so that the maximum diffraction efficiency can quadruple to over forty percent!

Sine-Wave Phase Grating

Now we will come almost full circle, from sinusoid to square and back: the depth of the phase-retarding structure varies smoothly, exactly as a sinusoid (the refractive index might vary sinusoidally instead, which is more common in bleached holograms). This turns out to be one of the few cases where the diffraction efficiency can be calculated analytically without much trouble, even though the link between phase and complex transmittance becomes highly nonlinear for only moderate modulations. For small phase modulations, the results should resemble those for sinusoidal amplitude gratings, although the phases of the first-order diffracted waves are different by 90° from the unbleached case (which hardly ever matters).

The diffraction efficiencies are expressed in terms of zero- and first-order Bessel functions of the first kind, which are a lot like cosine and sine functions except that they damp down for large $\Delta\phi$, are not strictly periodic, and the maxima of J_1 do not lie at the minima of J_0 . Nevertheless, the general behavior is as expected: as the modulation increases the zero-order beam weakens and the first-order beams strengthen to a maximum *DE* of 33.8% (when $\Delta\phi = 0.59\pi$). Because a sinusoidal-phase grating is a distorted sinusoid in amplitude transmittance terms, higher order beams begin to appear too, each described by a higher-order Bessel function, $J_m^2(\Delta\phi)$. For a more detailed look at this question, see the Collier, Burckhardt & Lin book ².

Generalized Gratings

If the transmittance variation is neither smoothly sinusoidal nor step-wise constant, the diffraction efficiency can be difficult to compute even within the limited accuracy of this simple thin-hologram approach. However, there are a *few* things we can say in general that help tie the just-previous results together, and extend them in interesting ways toward image holograms. These ideas are simple to comprehend for amplitude gratings, a little harder for phase-only

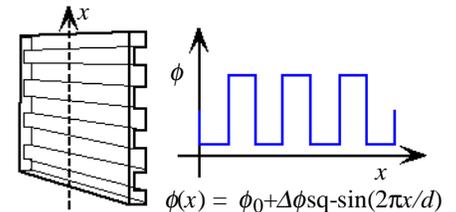
$$DE_{m=0} = t_0^2$$

$$DE_{m=\pm 1} = \left(\frac{2\Delta t}{\pi}\right)^2 = 10.1\%_{\max}$$

$$DE_{m=\text{even}} = 0$$

$$DE_{m=\text{odd}} = \frac{1}{m^2} DE_{+1}$$

$$\sum_{m \neq 0} DE_m = \Delta t^2 = 24\%_{\max}$$



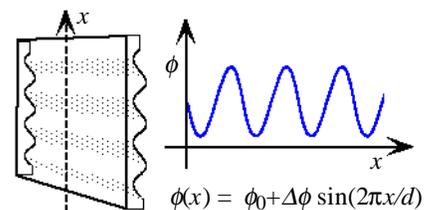
$$DE_{m=0} = \cos^2 \Delta\phi$$

$$DE_{m=\pm 1} = \left(\frac{2}{\pi} \sin \Delta\phi\right)^2 = 40.5\%_{\max}$$

$$DE_{m=\text{even}} = 0$$

$$DE_{m=\text{odd}} = \frac{1}{m^2} DE_{+1}$$

$$\sum_{m \neq 0} DE_m = \sin^2 \Delta\phi = 100\%_{\max}$$



$$DE_{m=0} = J_0^2(\Delta\phi)$$

$$DE_{m=\pm 1} = J_1^2(\Delta\phi) = 33.8\%_{\max}$$

$$\sum_{m \neq 0} DE_m = 1 - J_0^2(\Delta\phi) = 100\%_{\max}$$

gratings, and the general mixed case requires a lot of dabbling in the unit circle of complex variable mathematics.

The fraction of the optical power transmitted at each point of the grating is given by the magnitude-squared of the amplitude transmittance at that point, or $|t(x,y)|^2$. That power finds its way into the variously diffracted beams, so that the first number we can find is the sum over all the orders, including the zero order, of the diffraction efficiencies.

The amplitude of the zero-order beam by itself is given by the average of the transmittance over the entire hologram area, which we call \bar{t} (this was t_0 in the previous amplitude-only examples; in general, it is a complex number, $t_0 e^{i\phi_0}$). The power in the zero-order beam is the magnitude-squared of that average, or t_0^2 in this case.

$$DE_{m=0} = |\bar{t}_{\text{avg}}|^2$$

$$\sum_{m \neq 0} DE_m = \overline{|t(x)|^2} - DE_0 = \overline{|\delta t(x)|^2}$$

$$= \text{var } t = \sigma_t^2$$

The difference between the total diffracted power and the power in the zero-order beam must be the total power in all the diffracted beams! If the transmittance is a constant over the hologram area, then the average of the square of the magnitude of the transmittance will be equal to the square of the magnitude of the average transmittance, and the diffracted power will be zero, as expected. If the transmittance fluctuates as a function of position, diffraction begins. The difference between the average of the magnitude-squared and the magnitude-squared of the average is termed the *variance* of the random fluctuations, or the square of their *standard deviation*. This is equal to the sum of the diffraction efficiencies in all the orders!

Telling how much power goes into any specific order, and into the $m=+1$ order in particular, is trickier. If the fluctuating transmittance pattern can be decomposed into various spatial frequency components, then the variance can be interpreted as a sum over a power spectrum, where each component of the power spectrum corresponds to the power in one diffracted order (an application of the Wiener³-Khinchine theorem of communication theory). That precise decomposition required finding the Fourier transform of the transmittance fluctuations, which is beyond the scope of this discussion.

There are some interesting special cases, though. For example, can you think of a transmittance pattern that diffracts all of the light into one of the first-order beams?

Thick Gratings

All of the above discussion assumes that the modulation of the grating has been crammed into a layer that is infinitely thin. In real holograms, the emulsion is several wavelengths of light thick (silver-halide emulsion thickness of 5-7 μm is typical), and the modulation is spread over fringe surfaces that are wide enough to act something like mirrors—that is, they may diffract light more into the +1 order than the -1 order, or vice versa. This angular selectivity is usually called the “Bragg effect,” and its analysis would take us far beyond the mathematical scope of this course^{4,5,6}. In general, the trend is to increase the diffraction efficiency of one beam at the expense of the others, and to make the hologram quite sensitive to its angle to the illuminating beam. Volume reflection holograms, described in later chapters, are at the opposite extreme, where the emulsion layer is considered as being extremely thick in the simplest case.

References:

1. The Bragg angle is the angle (for each wavelength of light) at which selection effects due to the thickness of a hologram maximize its diffraction efficiency. Named after the Braggs, father and son: Bragg, Sir William Henry, 1862-1942, British physicist. He shared a 1915 Nobel Prize with his son Sir William Lawrence Bragg (1890-1971, 25 at the time) for the analysis of x-ray spectra and the structure of crystals.
2. R.J. Collier, C.B. Burkhart, and L.H. Lin, *Optical Holography*, Academic Press, San Diego, 1971; Section 8.5.
3. Wiener (wè'nér), Norbert, 1894-1964, American mathematician and educator (MIT); b. Columbia, Mo. Known for founding the theory of *cybernetics* and for his many contributions to the development of computers, Wiener also did research in probability theory and the foundations of mathematics. He was one of the few child prodigies whose later lives fulfilled their early promise.
4. H. Kogelnik, “Coupled Wave Theory for Thick Hologram Gratings,” *Bell System Technical Journal* **48**, 2909-2947 (1969).

-
- ⁵. R.J. Collier, C.B. Burkhardt, and L.H. Lin, *Optical Holography*, Academic Press, San Diego, 1971; Chapter 9.
- ⁶. P.M. Hariharan, *Optical Holography*, Cambridge University Press, Cambridge, England, 1984; Chapter 4.