Lecture 10: Holography

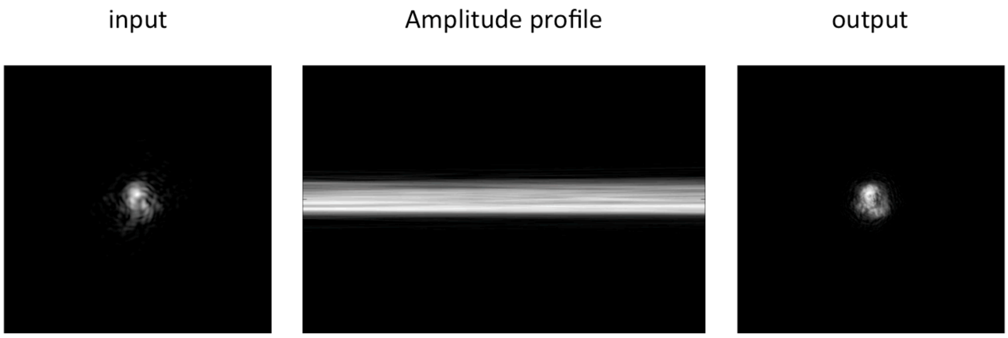
<u>Outline</u>

1. Amplitude and Phase
2. Interferometers and phase measurements
3. Holography
4. Holograms with plane waves
5. Holograms with spherical waves
6. Reflection holograms
7. Holographic recording materials
8. Digital holography
9. Phase conjugation
10. Holographic filters

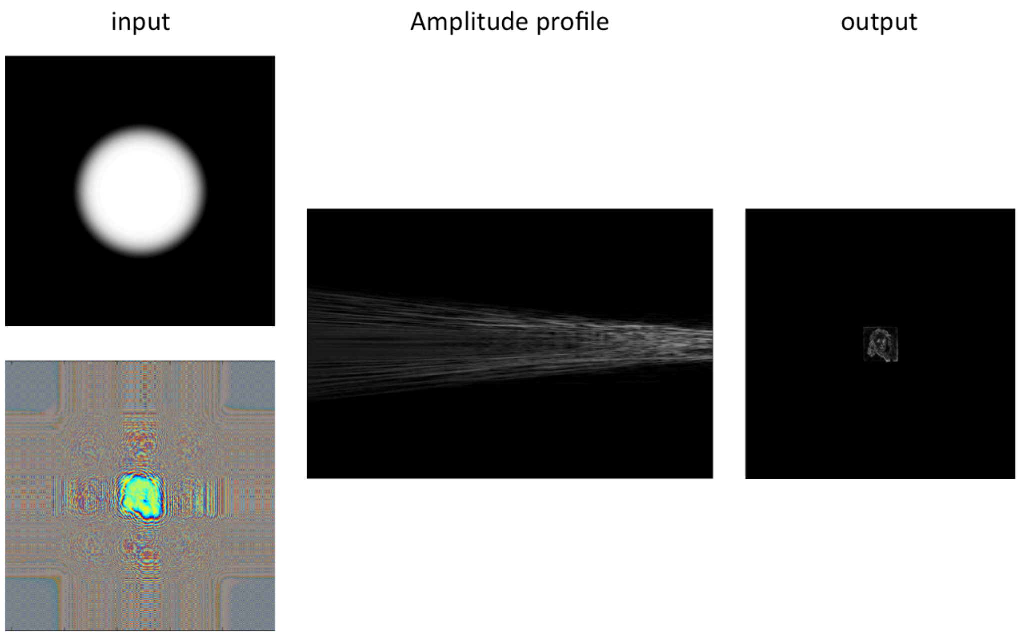
The word holography means in Greek complete recording. Complete in the sense that both the amplitude and phase of the optical field are recorded. In conventional photography the recorded signal is proportional to the light intensity recorded at a plane at $z=z_0$:

$$t(x,y)=\alpha\, I(x,y,z_0)=\alpha\, \overline{(\vec{E}\times\vec{H})\bullet\vec{dS}}\;=\alpha\, A(x,y,z_0)A^*(x,y,z_0)$$

where $E(x,y,z,t)=|A(x,y,z)|\cos(\omega t-kz-\varphi(x,y,z))$, $A(x,y,z)=|A(x,y,z)|\,e^{j\phi(x,y,z)}$. Therefore in a conventional photograph or recording in an electronic camera all the phase information is gone. How important is the phase? In a focused image plane, where the 2D intensity profile of a distant object is replicated on the photographic film or the CCD camera, the phase is not needed. The retinas in our eyes also discard the phase when they detect intensity. However if we are not in a focused image plane, most of the information in the blurred image is actually in the phase. In other words if we had a choice to record either the amplitude or phase of A(x,y) and then retrieve the information about what was in the object we would generally get much more information out of the phase. In Figure 1 we see the face of Newton in and out of focus and the reconstruction of the field at the image focus plane from phase-only and amplitude-only information.



(a)



(b)

Figure 1

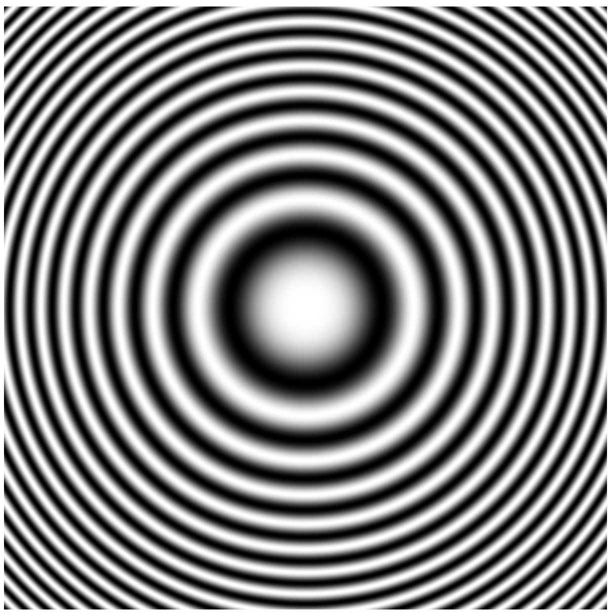


Figure 3

The recorded pattern in this case is of the form

$$I_{\mathrm{det}}(x,y) \sim \cos[\phi_1 - \beta(x^2 + y^2)]$$

Source coherence

If we do not have a purely monochromatic source then our ability to observe interference is generally impaired. Consider for example the Michelson interferometer with a source whose light amplitude can be described by $E(t)=a(t)e^{j\omega_0 t}$ where a(t) is the slow varying temporal envelope of the source. Suppose that the two paths of the interferometer differ by a distance d and both mirrors are perfectly flat. Then the signal detected is

$$I_{\det}=\int|a(t)|^2\,dt+\int|a(t-\tau)|^2\,dt+\frac{1}{2}\mathrm{Re}\{\int a(t)a^*(t-\tau)dt\}\cos\omega\tau=R(\tau)\cos\omega\tau$$

where $\tau=d/c$ and $R(\tau)$ is defined as the correlation function of the source or the temporal coherence function. We know from the correlation theorem that the Fourier transform of the correlation is the power spectrum. See Figure 4. The wider the bandwidth (*W*) of the light source (while light source) the shorter its correlation length. The coherence length is $\Delta d=c\,\Delta\tau=\dfrac{c}{W}$. In order to observe interference between two optical beams the difference in the optical paths they follow to reach the detector must be less than the coherence length of the source.

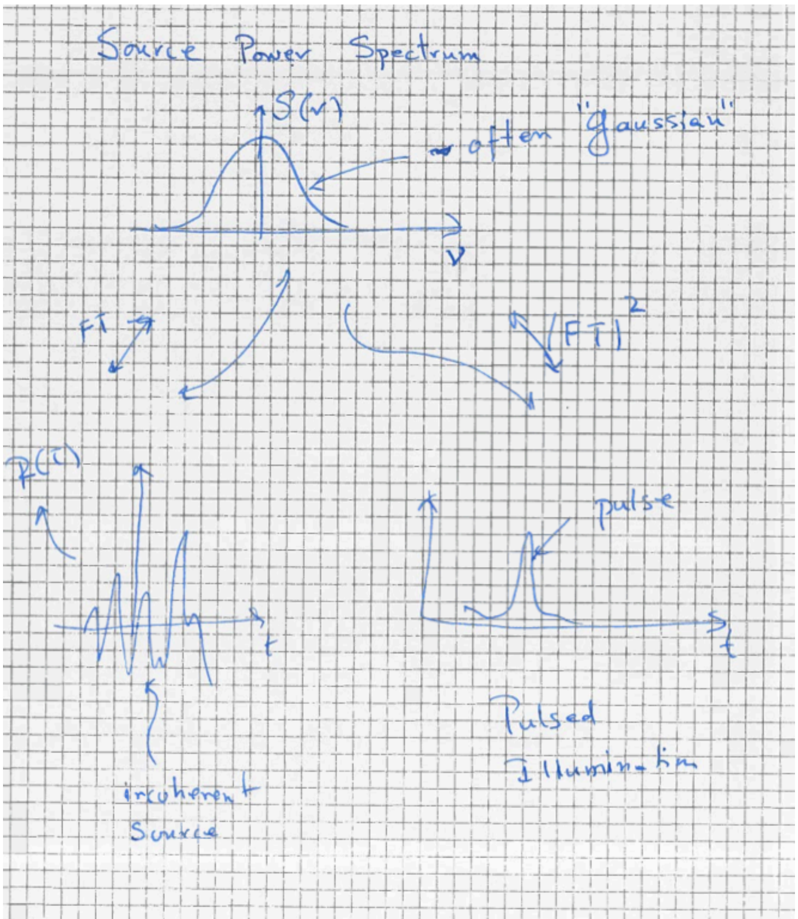


Figure 4
Broadband illumination

Wavefront shaping

A diagram of the Shack Hartman interferometer is shown in Figure 8. An array of small lenses (a lenslet array) is used. Each lenslet samples the local wavefront which appears as a plane wave to the small aperture of the lens. The local direction of the plane wave is mapped to position in the back focal plane of the lenslet. Measurement of the position is an estimate of the local phase gradient from which we can obtain the phase profile by integration.

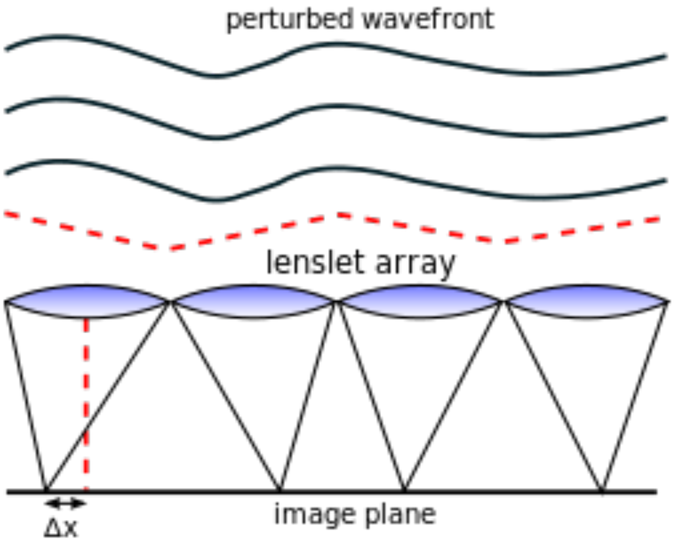
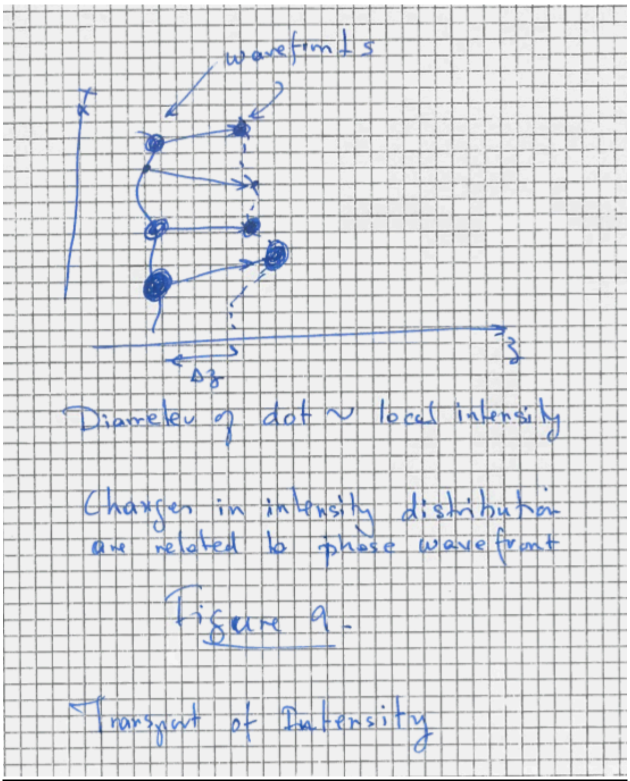


Figure 8. Shack-Hartman interferometer

Transport of intensity

Interferometric detection is complex, requiring coherent sources and mechanical stability. The Shack-Hartman interferometer is a way to address this but it has low resolution. Another way to do this through a method called "Transporation of Intensity". The idea is that the phase can be retrieved from the change in intensity. See Figure 9. For an analytical derivation of the Transport of Intensity equation see the Appendix to Lecture 10 notes.



Holography

The classic holographic recording and read-out set up is shown in Figure 10. An object is illuminated and the light reflected is directed towards the holographic recording medium. A portion of the light from the laser is diverted and made incident on the recording medium separately. This is the reference beam and it interferes with the object or signal beam to record the hologram. The transmittance of the film is proportional to the exposing light intensity:

$$t_{hol}(x,y) = |R+S|^2 = |R|^2 + |S|^2 + RS^* + R^*S = |R|^2 + |S|^2 + 2|RS|\cos(\phi_r - \phi_s)$$

Notice that the phase of the signal beam is now preserved in the interference term. If the phase of the reference is known we can imagine measuring the phase of the signal as we normally would do in interferometry. Often we use a plane wave as the reference in which case

$$R = e^{-j(k_{xr}x + k_z z)} \quad and \quad \phi_r = k_{xr}x + \phi_0$$

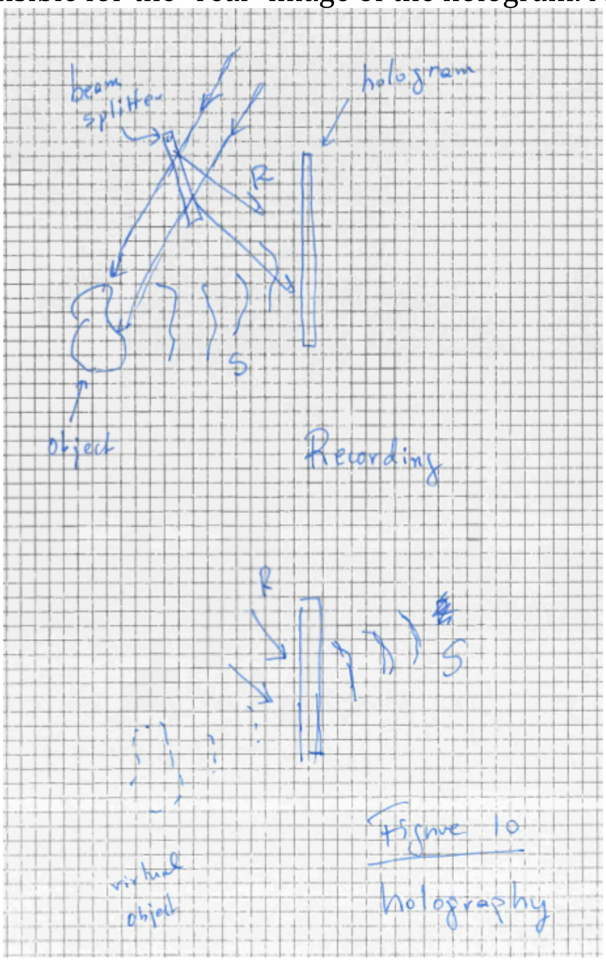
and we can recognize the 3rd term of the hologram as an amplitude and phase modulated spatial carrier:

$$t_{hol}(x,y) = |R|^2 + |S|^2 + 2|R||S|\cos(k_{xr}x - \phi_s)$$

The amplitude and phase is non other than the amplitude and phase of the signal beam. In order to recover the full complex field from the holographic recording we can digitize the hologram and demodulate in the computer the carrier and retrieve the phase and amplitude. This is called digital holography. In conventional holography the complex field is reconstructed optically by Illuminating the hologram (which is simply a thin transparency) with the same reference beam R to obtain:

$$tR = R|R|^2 + R|S|^2 + |R|^2 S + R^2 S^*$$

The third term is the term of interest. It is the signal beam multiplied by $|R|^2$ which is a constant (no spatial dependence) for beams like a plane wave or a spherical wave. Therefore the light that is diffracted away from the hologram due to this term is in principle indistinguishable from the original signal beam. Therefore if the recording signal S was the light scattered from a 3D object, the reconstruction of the hologram will reproduce the same optical field which will give the 3D sensation. The first two beams describe the read-out light (the plane wave) that goes through the holographic medium. It is slightly distorted by the low frequency modulation by $|S|^2$. The fourth term looks a bit unusual. If we think of the hologram as a grating and the 3rd term as the +1 diffraction order then the 4th term is the -1 diffraction order. Therefore the light will be directed on the other side of the zero order beam (the light that goes straight through). This term is responsible for the “real” image of the hologram. More on this later.

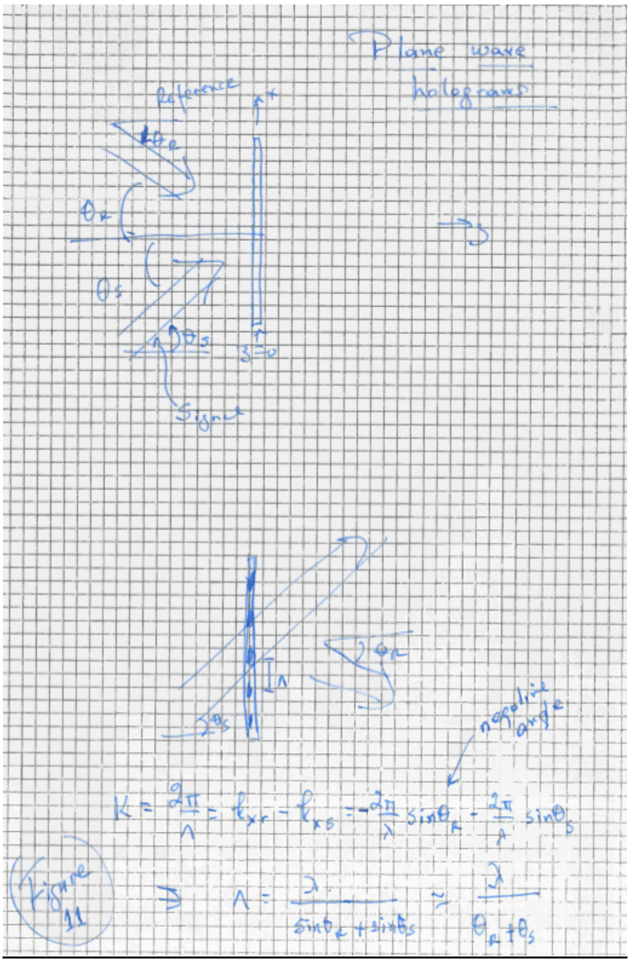


Plane wave holograms

Perhaps the simplest hologram is formed by selecting both R and S as plane waves. See Figure 11. Assuming the hologram is located at z=0 without loss of generality, the transmittance of the hologram can be written as

$$t_{hol}(x,y) = \left| R_0 e^{-jk_{xr}x} + S_0 e^{-jk_{xs}x} \right|^2 = \left| R_0 \right|^2 + \left| S_0 \right|^2 + 2R_0 S_0 \cos[(k_{xr} - k_{xs})x]$$

The holographic recording in this case is a sinusoidal amplitude grating. If the amplitude of the signal beam increases then the diffraction of the light into the first order increases in proportion. If the phase of the signal beam shifts, the recorded grating picks up the same phase shift. In this way the signal beam (the plane wave) is reconstructed by the hologram.



Spherical wave holograms

The hologram recording formed by interfering a spherical signal wave with a plane wave reference is shown in Figure 12. The recording is a chirped grating. In this geometry we can trace the nature of the “conjugate” beam. While the regular reconstruction is a diverging spherical beam that when viewed by an observer it will appear to emanate from behind the hologram, the conjugate diffracted beam is a converging wave the come to a focus in *front* of the recording medium. An arbitrary object can be thought of as collection of point sources. Therefore the formation of a real object in front of an object happens also for an arbitrary object not only a point source. The object formed in this way however appears inverted in depth. It is called the pseudoscopic image.

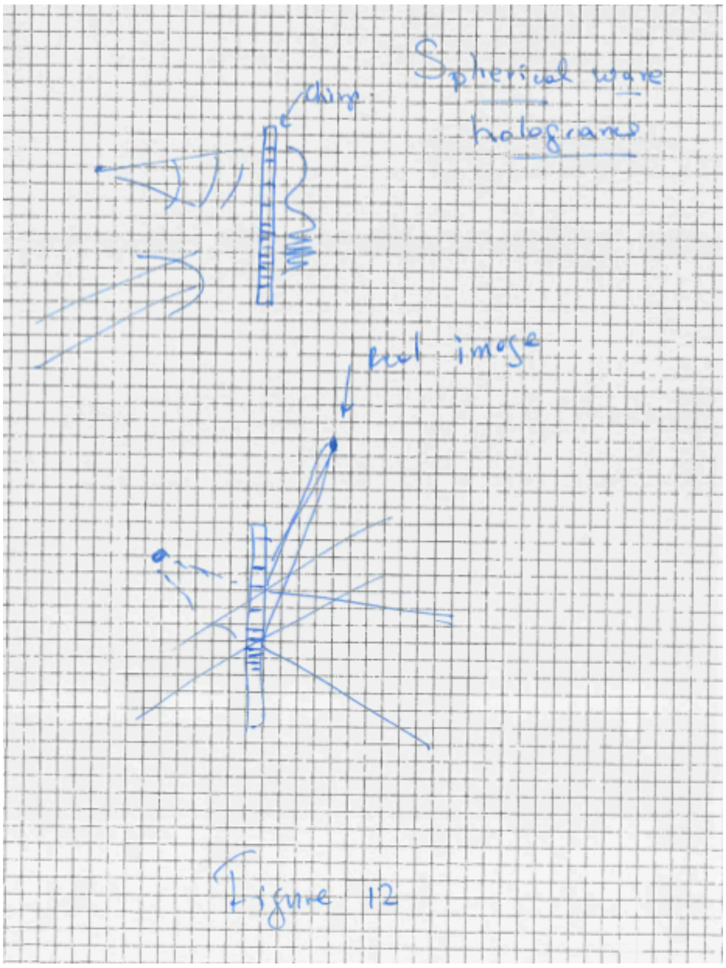
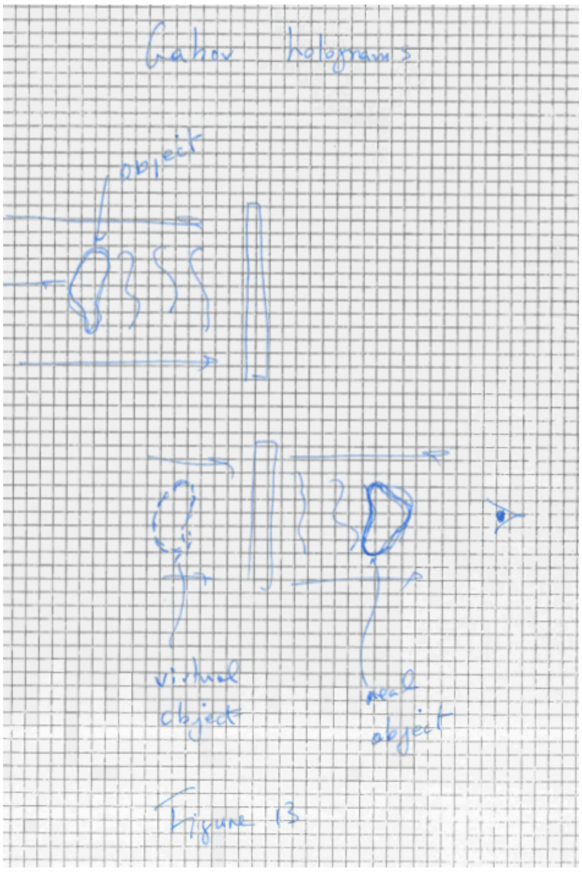
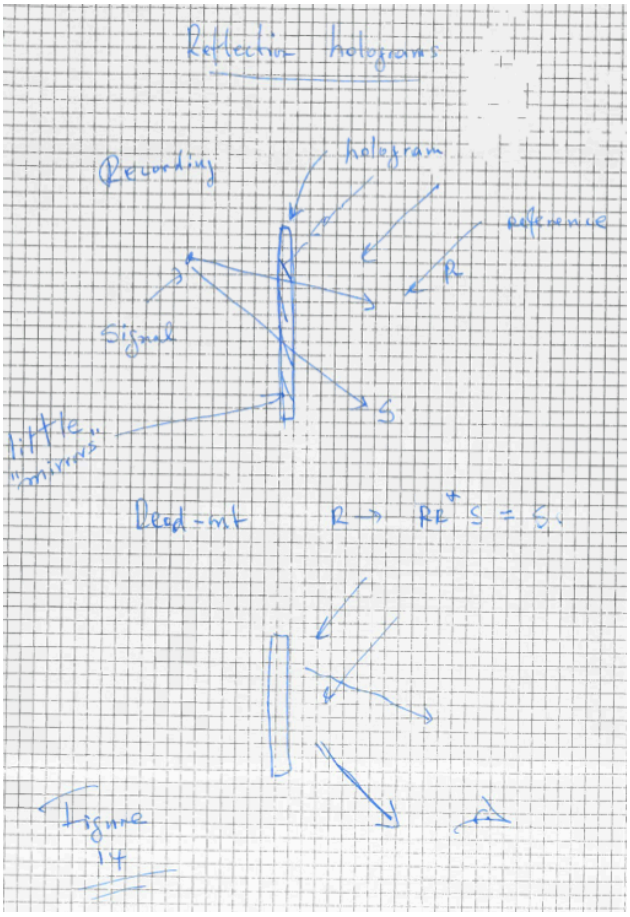


Figure 12

A nice property of the hologram formed as in Figure 11 is that the 3 diffracted beams are angularly separated and can be viewed independently. This is called the off-axis holography and was invented by Leith and Upatnieks. Holography itself was invented in 1947 by Gabor. He demonstrated on-axis holography shown in Figure 13 . In this case the holographic image needs to be viewed in the strong background of the read-out beam. The advantage is that it requires minimal coherence from the light source and in 1947 there were no lasers. Leith and Upatkiens did their holograms with a laser. Gabor got the Nobel prize for holography in 1971.



We can also form a hologram in the reflection geometry. This was invented by Denysuk, a Russian scientist. See Figure 14 .   This is actually the most common type of display since it can be viewed in reflection hologram and it can be read-out with incoherent (even white) light. The reason is the wavelength selectivity of reflection gratings which act as wavelength filters.



Holography versus interferometry

How is holography different from interferometry? After all a hologram is nothing but an interferometric recording of a complex field. True enough. It is also true that a hologram is the perfect interferometer. It records amplitude and phase of the field and in principle makes all necessary information available. Indeed the two concepts are closely related. One difference is that holography does not aim to only record and measure the phase but also to find a way to reproduce the complete field. Gabor's original motivation for the invention of holography was lensless electron microscopy. In other words if we can record the complete complex field then we do not need lenses and other classical optical elements because we can manipulate the complex field directly. At Gabor's time the manipulation could only be done optically, nowadays it is can also be done digitally. It is probably this idea of post-detection complex field manipulation that captures best the idea of holography.

Appendix

Start with $A(\vec{r}) = \sqrt{I(\vec{r})}e^{i\phi(\vec{r})}$.

First calculate

$\vec{\nabla}_\perp A = \frac{1}{2}I^{-1/2}\left(\vec{\nabla}_\perp I\right)e^{i\phi} + iI^{1/2}\left(\vec{\nabla}_\perp \phi\right)e^{i\phi}$

$\nabla^2_\perp A = -\frac{1}{4}I^{-3/2}\left|\vec{\nabla}_\perp I\right|^2 e^{i\phi} + \frac{1}{2}I^{-1/2}\left(\nabla^2_\perp I\right)e^{i\phi} + \frac{i}{2}I^{-1/2}\vec{\nabla}_\perp I \cdot \vec{\nabla}_\perp \phi e^{i\phi} +$

$\frac{i}{2}I^{-1/2}\vec{\nabla}_\perp I \cdot \vec{\nabla}_\perp \phi e^{i\phi} + iI^{1/2}\left(\nabla^2_\perp \phi\right)e^{i\phi} - I^{1/2}\left|\vec{\nabla}_\perp \phi\right|^2 e^{i\phi}$.

Multiply by A^* to obtain

$A^*\nabla^2_\perp A = -\frac{1}{4}I^{-1/2}\left|\vec{\nabla}_\perp I\right|^2 + \frac{1}{2}\nabla^2_\perp I + i\,\vec{\nabla}_\perp I \cdot \vec{\nabla}_\perp \phi + iI\,\nabla^2_\perp \phi - I(\vec{r})\left|\vec{\nabla}_\perp \phi\right|^2$.

Next calculate

$i4\pi\kappa\frac{\partial A}{\partial z} = i2\pi\kappa I^{-1/2}\frac{\partial I}{\partial z}e^{i\phi} - 4\pi\kappa I^{1/2}\frac{\partial \phi}{\partial z}e^{i\phi}$.

Multiply by A^* to obtain

$i4\pi\kappa A^*\frac{\partial A}{\partial z} = i2\pi\kappa\frac{\partial I}{\partial z} - 4\pi\kappa I\frac{\partial \phi}{\partial z}$.

From paraxial Helmoholtz equation, we have

$$A^*\left(\nabla^2_\perp + i4\pi\kappa\frac{\partial}{\partial z}\right)A = 0.$$

or, combining terms from above,

$-\frac{1}{4}I^{-1/2}\left|\vec{\nabla}_\perp I\right|^2 + \frac{1}{2}\nabla^2_\perp I + i\,\vec{\nabla}_\perp I \cdot \vec{\nabla}_\perp \phi + iI\,\nabla^2_\perp \phi - I(\vec{r})\left|\vec{\nabla}_\perp \phi\right|^2 +$

$i4\pi\kappa A^*\frac{\partial A}{\partial z} = i2\pi\kappa\frac{\partial I}{\partial z} - 4\pi\kappa I\frac{\partial \phi}{\partial z} = 0$.

Both imaginary and real parts must equal 0. Look at imaginary part only:

$\vec{\nabla}_\perp I \cdot \vec{\nabla}_\perp \phi + I\,\nabla^2_\perp \phi + 2\pi\kappa\frac{\partial I}{\partial z} = 0$.

Hence

$$2\pi\kappa\frac{\partial}{\partial z}I = -\vec{\nabla}_\perp I \cdot \vec{\nabla}_\perp \phi - I\,\nabla^2_\perp \phi = -\vec{\nabla}_\perp \cdot I\,\vec{\nabla}_\perp \phi.$$