Interference

The interference of light is both fascinating and subtle. No one realised it was happening until Thomas Young made the brilliant experiment in the early 19th century that will be described in this chapter. Yet, interference is responsible for the vibrant colours on starling's wings, on beetle exo-skeletons, on fish scales, on oxidized titanium jewelry, in supernumerary rainbows, in oil on water and other 'everyday' phenomena. Interference is used to measure large and small distances exceedingly accurately, to make holograms and very much more besides. This practical discipline is called **interferometry**. Interference also tells us something quite basic and quite deep about how light is produced.

Interference is very odd. Imagine a light wave producing more or less uniform illumination on a screen. Now add a second wave that on its own produces equal illumination. With them both together 'interfering', there are places on the screen where you get four times the illumination produced by one of the waves. This almost seems like something for nothing. How can all this be going on and no one noticed until barely 200 years ago?

'Ordinary' illumination

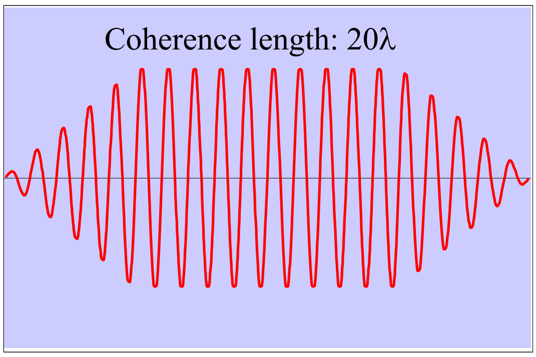
It's true that we don't notice the interference of light in ordinary circumstances. Why not? The answer takes you very quickly into the subject of what's really going on when sources emit light.

First, light is a very rapid fluctuation of electric field. Typically, the period of a light wave is about 10^{-15}s. Neither our eyes, nor photographic film nor any of the modern detectors can follow changes in electric field that are this quick. What detectors 'see' is the energy associated with the light wave, or strictly speaking the time average energy, averaged over many, many periods of oscillation. We introduced this idea in the previous section and wrote that the irradiance, I, depended on the average of the square of the electric field:

$$I \propto \langle E^2 \rangle \quad .$$

You need to remember this.

Secondly, almost all everyday light sources (the sun, filament bulbs, fluorescent tubes) don't emit long wavetrains of light. They emit short packets, each point on the source emitting light quite independently of the next point or indeed of the same region a brief time later. The accompanying figure shows such a packet of about 20 wavelengths long, or 20 periods, which is just the same thing. Now each of these packets acts quite independently. Each produces its own brief burst of illumination. There are billions of these arriving at any surface we're looking at and the total illumination is just the sum of the illumination provided by each one. In symbol form:



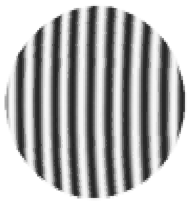
Temporal coherence: 20 periods

$$I_{total} = I_1 + I_2 + I_3 + I_4 + \ldots..$$

This is how everyday illumination works. When a ceiling full of lights shines onto a desk, the total illumination is the sum of the illuminations produced separately by each light. This state of affairs is called **incoherent illumination,** and it's what we're used to. It is the background against which you have to interpret the effects of interference.

What is interference?

Interference usually shows itself as regular **fringe patterns**, a series of light and dark areas of illumination. Sometimes these are straight-line fringes, sometimes circular, sometimes wiggly. Sometimes the neighbouring lines are equi-spaced, sometimes not. Unfortunately, not all optical fringe patterns are caused by interference (for example Moiré fringe patterns aren't), but most are.

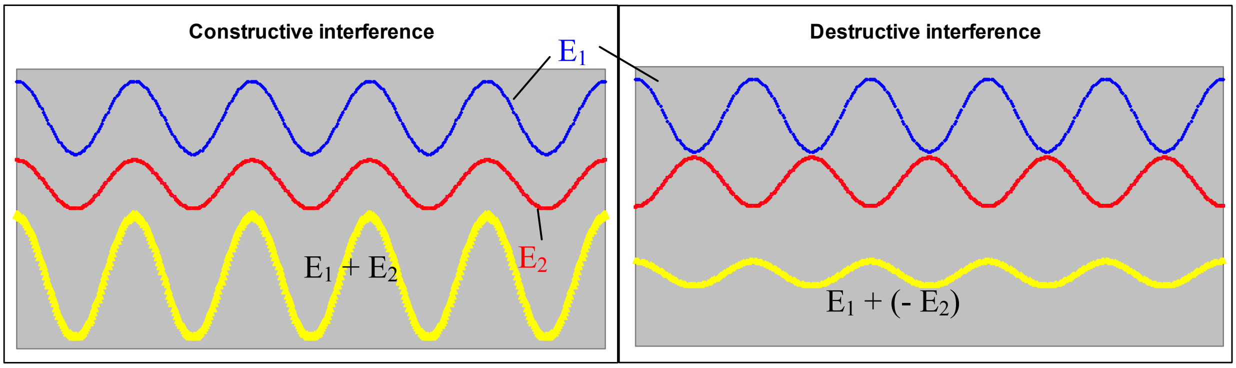


The essential difference between interference and normal circumstances is that with interference the light waves add together *first* and *then* the average value of the total gives the irradiance seen.

How two light waves behave when added together depends on the relative phase of the waves. The two extreme circumstances are:

1) the light waves have just the same phase, which leads to **constructive interference**
2) the light waves are 180° (π radians) out-of-phase, which leads to **destructive interference**

Of course, there can be any phase difference in between these extremes.

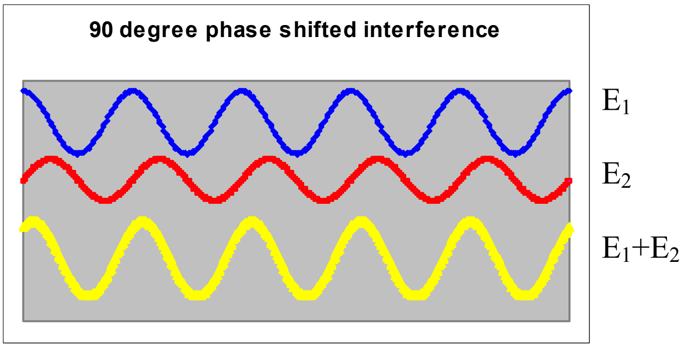


The accompanying diagrams show what happens. The blue wave (E1) has amplitude 3 units and the red wave (E_2) has amplitude 2 units. The constructive interference has an amplitude of 5 units and the destructive interference an amplitude of 1 unit.

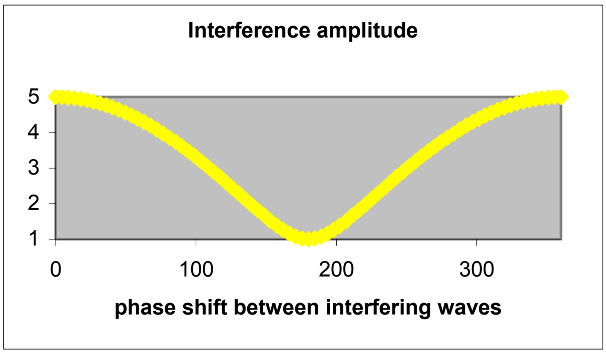
Intermediate phase shifts

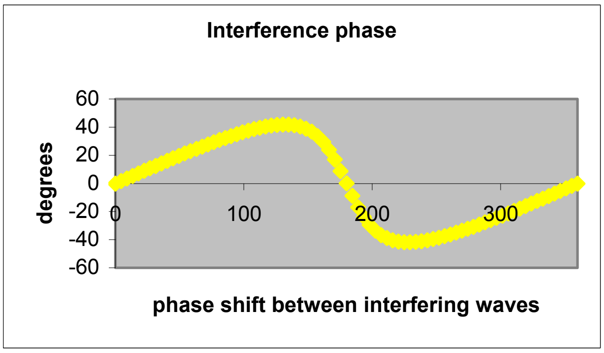
If the two interfering waves are neither completely in phase nor completely out-of-phase, the result is still a single wave. Its amplitude is somewhere between the extreme cases and the resulting interference wave is not in-phase with either wave.

Look at the accompanying sketches. The one alongside shows the interference of the same two waves as above, only this time the red wave is 90° out-of-phase with the blue one.



The second set of sketches shows what happens when the phase shift of the red wave (E_2) is varied through 360° relative to the blue wave (E_1). The amplitude varies smoothly between the maximum of 5 for constructive interference to the minimum of 1 for destructive interference. The phase shift of the interference wave relative to the blue wave (E_1) also changes. Since the irradiance of the interference depends on the square of the amplitude and does not depend on the phase, then the question of the phase of the interference wave hardly ever comes up. We'll forget about it from now on.

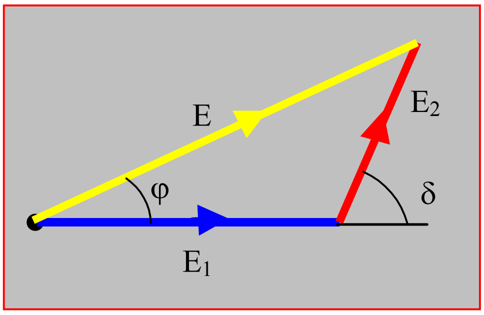




To recap: ***Constructive interference*** occurs when the two waves are in phase, or differ in phase by a multiple of 360°. The total irradiance is proportional to $(E_1 + E_2)^2$, as expected. Hence two waves of equal irradiance constructively interfere to give light of **four** times the irradiance of either wave. **Destructive interference** occurs when the two waves are π radians out of phase (i.e. 180°), or an odd multiple of π radians. The total irradiance is proportional to $(E_1 - E_2)^2$, as expected. For two waves of equal irradiance (only) the result is no irradiance at all.

Mathematically speaking

What we've done above is to use the computer to work out the sum of two waves. What's actually going on? It's always a good idea to try to understand what's happening, because the use of computers isn't supposed to be a substitute for understanding. In this case, a nice visual idea comes to our aid. It's called a **phasor**. The amplitude and phase of each wave of interest is represented by a line on a diagram. This line is called a phasor. The length of the line is proportional to the amplitude, and the angle that the line makes to the horizontal axis represents its phase. That's almost it. There's one more idea, which is that to add phasors you put them end-to-end, just as you add vectors in a vector diagram.

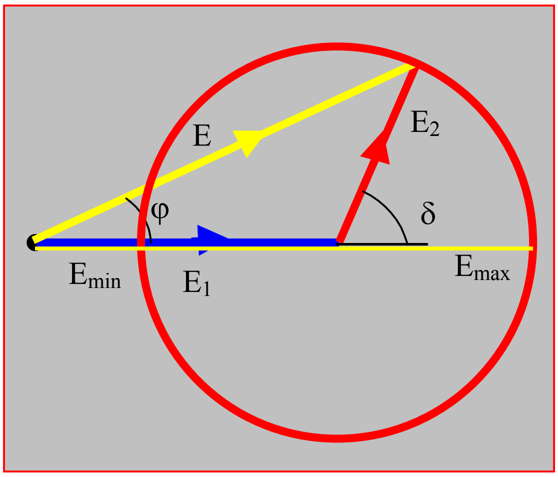


Phasors have a good mathematical justification, so this isn't just a piece of hand-waving analogy. They really work. Look at the phasor diagram for the addition of our waves. Our first wave had amplitude E_1 (3 units in the illustrations) and we represent it by a horizontal phasor. The second wave had amplitude E_2 (2 units in our example) and suppose it differs in phase by an angle δ, as shown. The sum of the two is represented by the yellow line, amplitude E, and its phase angle is given by φ in the diagram. If you know the cosine rule that relates the square of one side of a triangle to the squares of the other two sides and the angle between them (it is an extension of the theorem of Pythagoras), then you'll see immediately that the following must be the case:

$$E^2 = E_1^2 + E_2^2 + 2E_1E_2\cos(\delta) \quad .$$

In terms of irradiance, I, the result becomes:

$$I = I_1 + I_2 + 2\sqrt{I_1}\sqrt{I_2}\cos(\delta) \quad .$$



This result is central to the workings of interference. The first two terms on the right-hand side are just what you'd expect for incoherent illumination. The third term is the crucial interference term that makes all the difference when the light waves have a constant phase difference (δ) between them. It is final term that gave the variation in interference amplitude shown earlier.

You can easily see how the diagrams drawn earlier for the whole range of phase differences δ running between 0 and 360° have come about. As δ varies over its complete range, the phasor for E_2 goes around in a circle. The maximum value of E is clearly (3 + 2) = 5 and the minimum value (3 – 2) = 1. Later on, phasors will let us see what's going on optically in a more complicated situation.

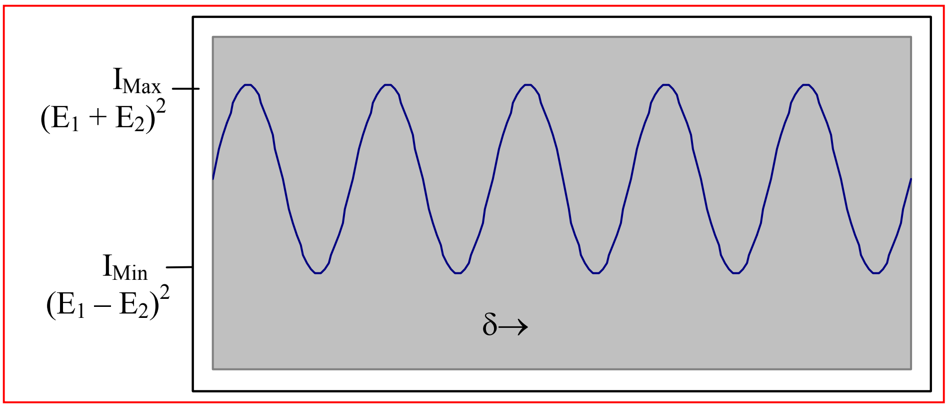
When the two waves have equal amplitudes and hence equal intensities, $I_1 = I_2 = I_o$, say, the expression for the irradiance simplifies:

$$\begin{aligned} I &= 2I_o + 2I_o\cos\delta \\ &= 2I_o(1+\cos\delta) \\ &= 4I_o\cos^2(\delta/2) \quad . \end{aligned}$$

[This result follows because $\cos^2\theta + \sin^2\theta = 1$ **and** $\cos2\theta = \cos^2\theta - \sin^2\theta = 2\cos^2\theta - 1$; hence $1 + \cos2\theta = 2\cos^2\theta$. Equivalently, $1 + \cos\delta = 2\cos^2(\delta/2)$].

Fringes formed by the interference of two waves are called **2-beam interference fringes**. Their intensity pattern varies as $\cos^2$, as you can see above.

Visibility



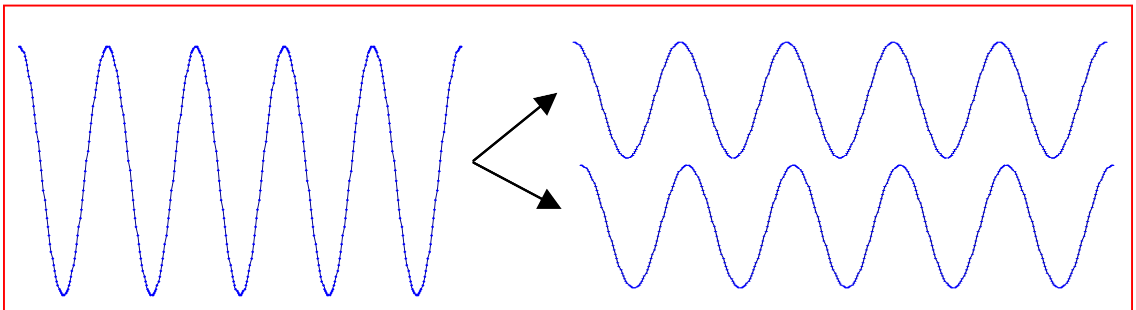
The visibility of an interference pattern is controlled by the ratio of the maximum to minimum irradiance in the pattern. This leads to the definition of the visibility (quotient) V, as a percentage, given by

$$V = \frac{(I_{max} - I_{min})}{(I_{max} + I_{min})} \times 100\% \quad .$$

In the best possible case, the two beams have equal intensity; $I_{min} = 0$ and the visibility is 100%.

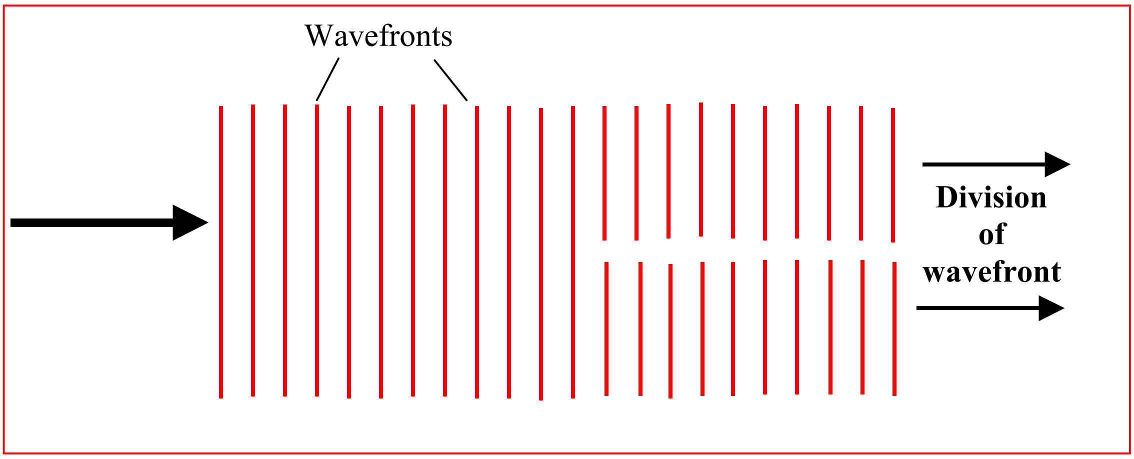
Light waves interfere with themselves

For two light waves to interfere, they must stay in step with each other. The technical word is that they must be **coherent**. Now this just doesn't happen with separate wavetrains issuing from light sources. Firstly, they must be **monochromatic**, namely of a single colour. Light waves of different colours or frequencies don't stay in step. Secondly, most light waves don't last long enough to interfere with other light waves. Earlier, a wavetrain was shown 20 periods long. Even if it were 10^5 periods long, it would be finished in about 10^{-10} seconds; gone before the next wavetrain of exactly the same wavelength arrived at exactly the same point. Even lasers, which produce exceptionally long wavetrains, don't normally produce light of a stable enough phase and frequency that you can see two lasers interfere with each other.



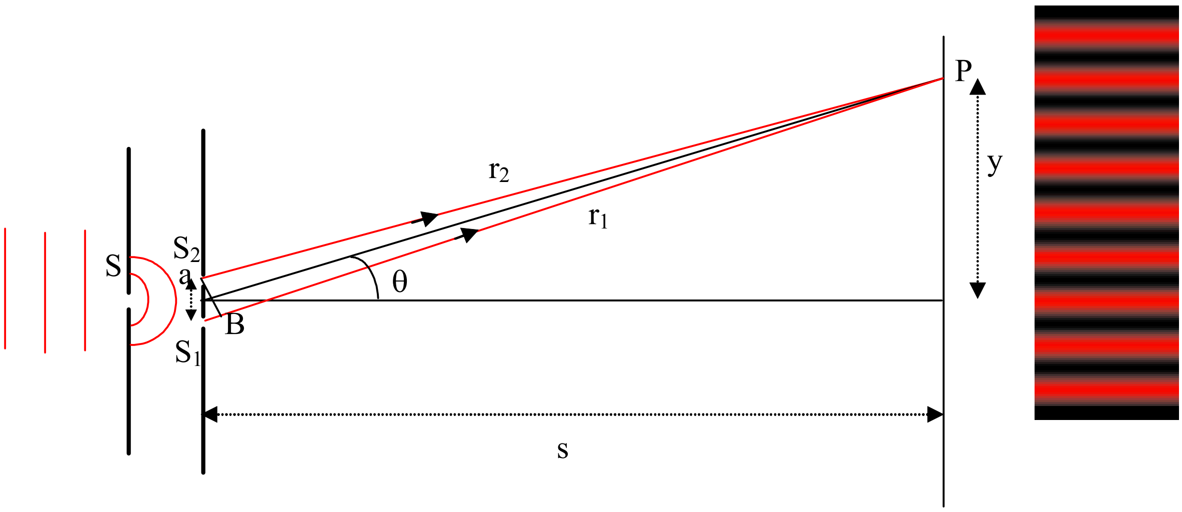
Division of amplitude

The only way to get light interfering is to make two wavetrains by dividing one original one, and then bring the two together again. There are two recognized ways of doing this: **division of amplitude** and **division of wavefront**. You'll see them at work soon.



Young's Experiment

Young's experiment was one of the world's great experiments. It is simple, easy to follow the correct theory, gives a clear visual result and gives deep insight. Young's experiment is an archetypal arrangement to illustrate interference of light (and other waves). Young used pinholes but it is usually done nowadays with slits. Variations in the experiment shed light on



light's subtle characteristics.

S_1, S_2 are two coherent sources a distance 'a' apart, generated from the original diffraction source S *by division of wavefront*. Constructive interference occurs at P when the path lengths S_1P and S_2P differ by a whole number ('m') of wavelengths, i.e.

$$m\lambda = S_1P - S_2P = S_1B = a.\sin\theta \approx a\theta \, ,$$

since the fringe pattern will usually be spread over a small range of θ. θ varies as P moves up the screen. 'y' measures the distance of P up the screen.

$$\theta \approx \tan\theta = y/s \ .$$

- Bright **fringes occur at positions y = s θ = msλ/a** . *Hecht* labels these positions y_m, giving

$$\boxed{y_m = \frac{ms\lambda}{a}}$$

- The distance between neighbouring fringes is clearly $\Delta y = s\lambda/a$. For example, with $\lambda = 500$ nm, a = 1 mm and s = 3m, the spacing between neighbouring fringes is 1.5 mm.

- The irradiance of the interference fringe pattern is that for 2-beam interference from waves of equal amplitude, i.e.

$$I \propto \cos^2(\delta/2) = \cos^2(kay/2s) \quad ,$$

since the phase difference between the two waves at P is $\delta = k(r_1 - r_2) = kay/s$.

Rôle of diffraction

Diffraction causes the light to be spread out from S in the first place, producing the coherent sources S_1 and S_2. Diffraction again causes the light to spread out from S_1 and S_2 and hence the irradiance of the pattern on the screen is multiplied by the irradiance of the diffraction pattern from a slit. The result is that the $\cos^2$ fringes don't go on for ever but fade out as the slit diffraction pattern fades with θ. They are most clearly visible only within the central maximum of the slit diffraction pattern, which is much the most intense part (see later in the course). With two pin-holes, one can clearly see the two circular diffraction patterns and the interference fringes present where they overlap.

Conclusions

- The interference pattern can be seen if S is got rid of and S_1, S_2 illuminated by a laser beam, which has a *transverse coherence* at least as big as *a*.

- By measuring the distance between neighbouring fringes, the wavelength of light can be deduced easily, even though it is more than ten times smaller than the finest division on a micrometer. It is also easy to show that red light has a longer wavelength than green light or blue light: the red fringes can be seen to be larger.

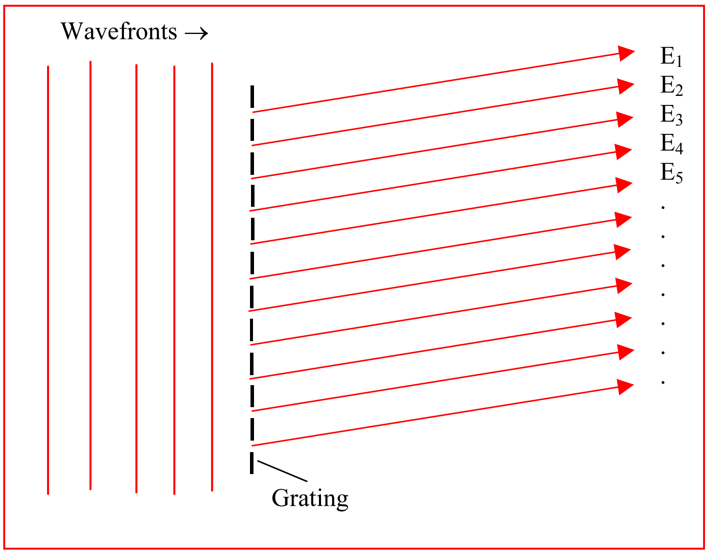
- White light does produce a fringe pattern, consisting of a white central peak surrounded by a coloured band on either side, since the blue fringes form closer than the red. The colours wash out after a few fringes into a uniform dull grey.

- By putting a thin wedge of transparent material across the two slits, an additional phase lag is added in one of the beams. (Alternatively, with a little care a thin wedge of material can be inserted just after S_1 alone). As this phase lag is increased, the fringe pattern will be found to disappear when the phase lag exceeds the coherence length of the source. In this way, coherence length can be measured.

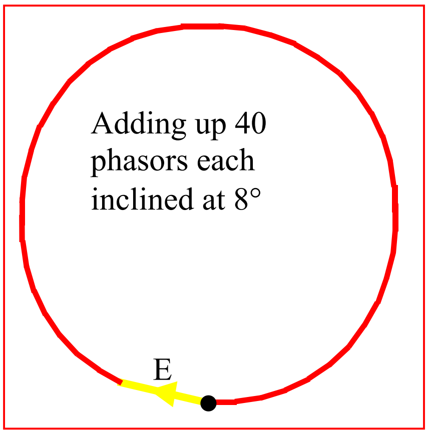
- Consider what happens when the light irradiance is reduced by so much that only one photon is present in the equipment at one time. How can the wavefront then be divided? Surely the photon has to go through one slit *or* the other and interference won't then occur? In fact, interference continues to occur. Pursue this on your own.

- There is nothing really in the experiment that ties it to optics. Waves of all kinds interfere and the mathematics is the same. Hence this experiment can be used to show convincingly that electrons, neutrons and other particles have wave properties, when the dimensions of the slits are appropriate and a suitable detector is used. You have seen that the wavelength associated with a 'particle' of momentum p is given by de Broglie's relationship $p = h/\lambda$. Pursue this on your own.

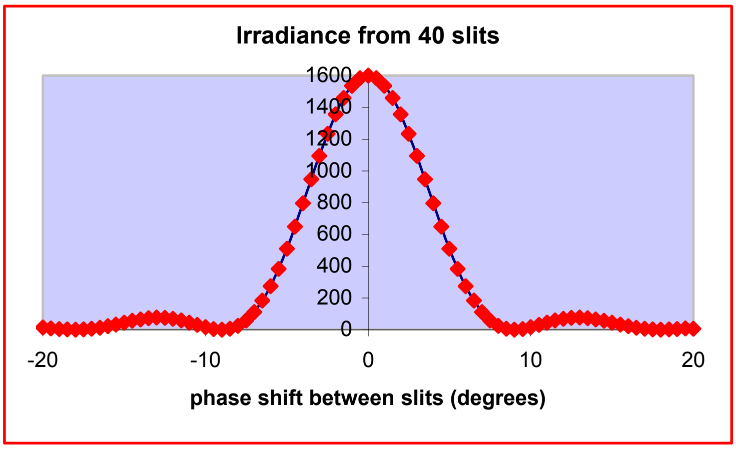
The idea behind Young's slits is just the idea behind a **diffraction grating**. This device consists 'simply' of thousands of equispaced narrow slits rather than just two slits. It works by interference, in spite of its name. It is at the heart of almost all **spectrometers**, instruments designed to spread out the spectral content of a light beam and allow the wavelengths of the component light to be measured. Diffraction gratings spread out the spectra exceedingly well. You can understand how they do this by thinking of the associated phasor diagram.



As an example, we'll consider the numbers involved when the diffraction grating has 40 slits, the effects of which can be shown and calculated. The diagram above shows very clearly how the diffraction grating works by *division of wavefront*. Suppose the phase difference in the waves going out from successive slits is exactly 360°. All the waves E_1, E_2, E_3, E_4, E_5, …… add up to give a total, E equal to 40 times the contribution from one slit, if there are 40 slits, or 'n' in general if there are 'n' slits. Now suppose the waves from successive slits are just 8° degrees out-of-phase. It's not much. Add up the waves from 40 slits. Instead of getting the resulting E forty times as large as each phasor, you can see from the phasor diagram above that the result is pretty small.



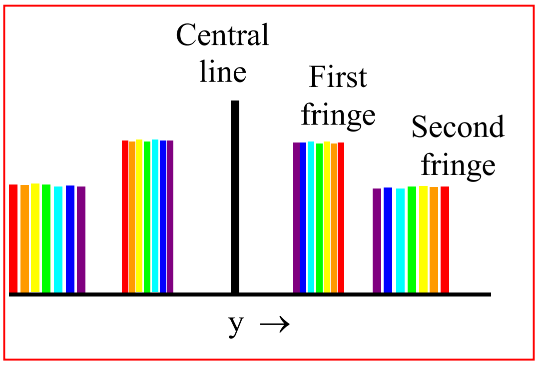
The next diagram shows the irradiance produced by the addition of 40 waves as the phase shift between each wave changes. The maximum is very sharply located. Just 4° of phase shift is enough to drop the irradiance to a half. If you were adding just two waves together, 4° phase shift would make almost no difference. The consequence is that instead of getting $\cos^2$ fringes you get fringes that are almost spikes. Remember that the phase shift has to be 360° before the next fringe appears on the diagram. 9° of phase shift reduces the irradiance to zero, as you would expect (9×40 = 360 and the phasor diagram has curled into a complete circle). The more 'slits' there are in the diffraction grating, the sharper the spike.



You can even deduce the angular width Δθ of each line. To achieve 9° of phase shift between neighbouring slits in our grating of 40 lines needs a change of angle Δθ such that $a\Delta\theta/\lambda = 9/360$. In general, $\Delta\theta = \lambda/(aN)$ where there are N slits each separated by a distance a. For example, if $a = 10\lambda$, making a about 0.005 mm, and there are 2000 slits (making the grating about 10mm wide), then $\Delta\theta$ = 1/20,000 radians or 2.9×10^{-3} degrees ≡ 10″ arc; not much.

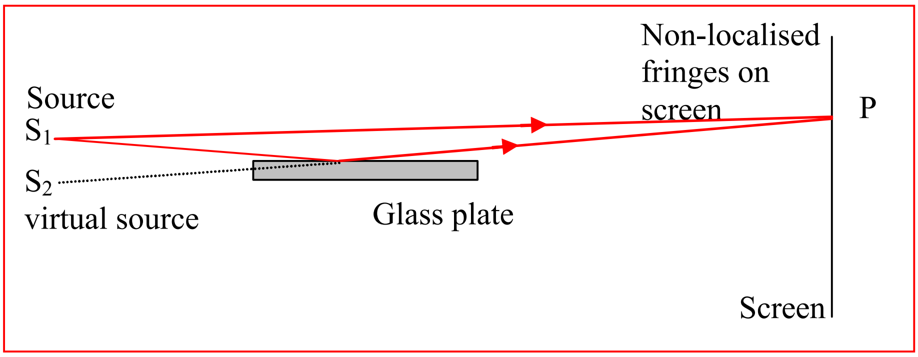
Since the irradiance depends on E^2, the irradiance produced by a grating increases as the square of the number of lines. So not only is the fringe sharper with, say, 4000 slits compared with 40 but it is 100^2 (=10,000) times as intense. Actual gratings may well have from 10,000 to 100,000 slits.

The final detail that makes the grating show the spectrum of the incident light is that each wavelength produces its fringe at a slightly different angle. Remember the Young's slits result that the angular position of the 'm'th fringe is just $\propto m\lambda/a$. 'a' here is just the slit separation. Hence, the range of wavelengths produces the spectral spread given by the diffraction grating.



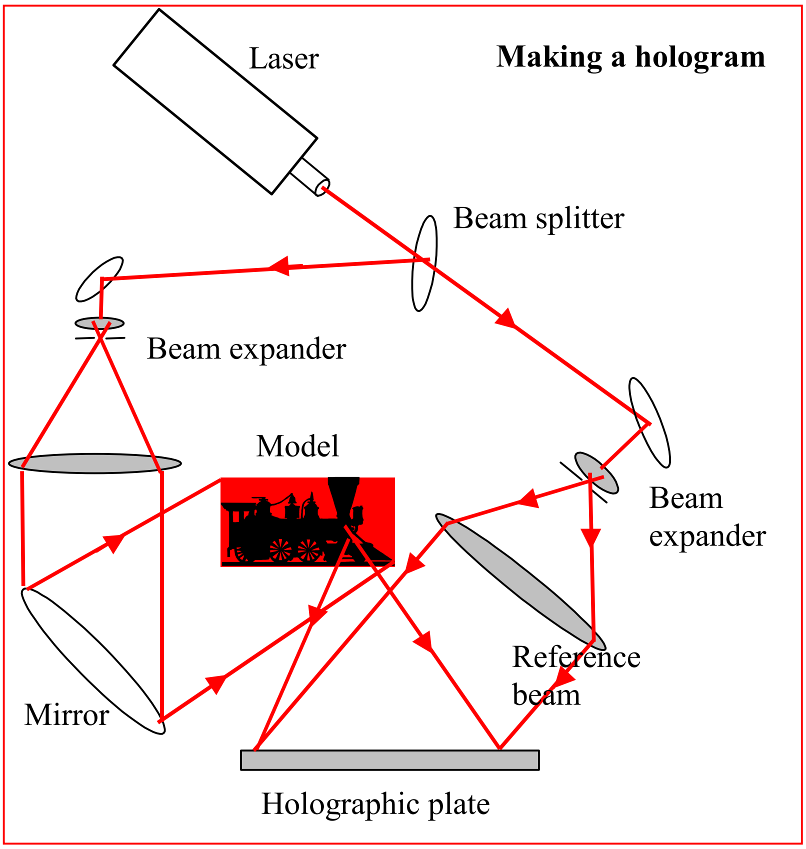
Lloyd's Mirror

Dr H Lloyd investigated this method of producing interference in the mid 19th century. It still has two points of interest. First, Lloyd was able to show that external reflection at a glass-air surface introduced a phase change of π radians (180°). Secondly, Lloyd's interference arrangement has all the ingredients needed for making a hologram. It took a century, though, before this was appreciated, since Gabor did not conceive of the idea of a hologram until 1959.



Lloyd produced 2-beam interference fringes by getting the wave directly from the source to interfere with part of the wavefront reflected from a glass plate. If S_1 is a horizontal slit or pin-hole, the second 'slit' S_2 is the mirror image of S_1. Just as in Young's experiment, the fringes are non-localized in that they appear on any screen. They are formed by *division of wavefront*. Just as in Young's slits, the closer together are the two effective sources S_1 and S_2, then the broader will be the fringe pattern.

Lloyd observed that when the path lengths from S_1 and S_2 are equal, a dark fringe occurs. This implied that the reflected light at the glass suffered a phase change of π radians. There is a dark fringe in the centre of some other fringe patterns (PX2505 students will meet *Newton's rings*) but you can't tell in this experiment which of the two reflections is contributing the π phase change. Lloyd showed that it had to be the reflection at a boundary where the refractive index increased.



A hologram records the interference between a direct (laser) beam and a beam reflected from the object (*Hecht* chpt. 14). Look at the schematic diagram here. Start at the laser, top left. The laser light is divided by the beam splitter. Part of it, on the left, is collected and expanded so that it illuminates the model. The model then generates reflected waves that reach the holographic plate. The other part of the laser beam is expanded to form a reference beam, which also lands on the plate. The two beams illuminating the plate interfere, forming a complex fringe pattern that is the hologram. This complex pattern is recorded when the plate is covered with an extremely fine grain photographic emulsion and exposed correctly.

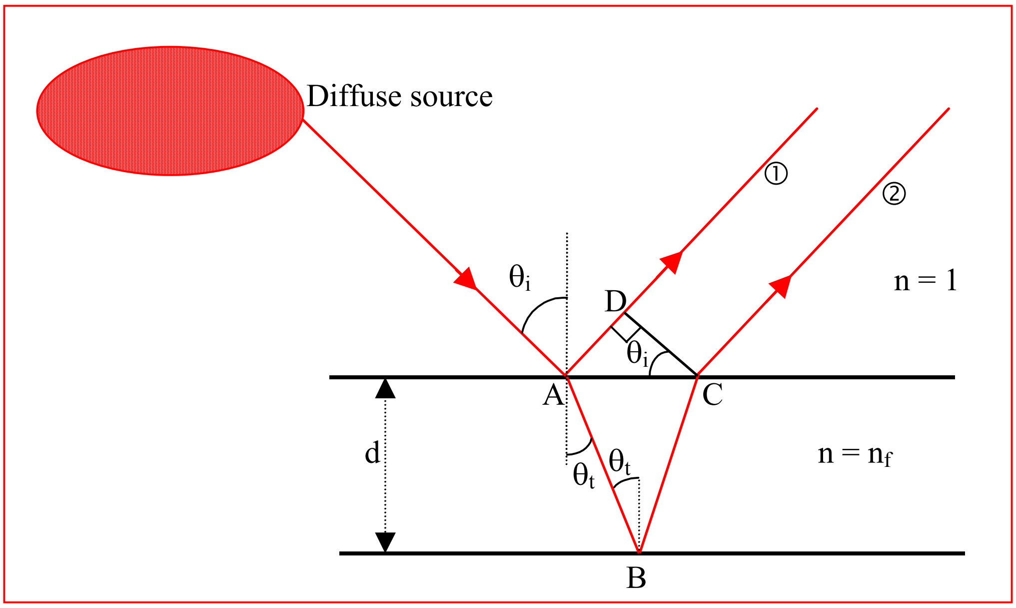
In Lloyd's mirror, this arrangement of interference fringes formed between a reference beam and a reflected beam is just what you have. A point source and its reflection produce simple $\cos^2$ fringes. When one source is a reflection from a complex object, the interference fringes are much more complicated. The fringes are exceedingly fine and have to be recorded on special very high resolution film, capable of recording in the region of 1000 lines per mm. Diffraction is needed to view a hologram. That's another, more complicated, story.

Thin film fringes

You can see **fringes of constant inclination** in patches of oil, on oxidized titanium jewelry, on some insect exo-skeletons and birds' wings. Coated lens optics uses interference from thin films of low refractive index to cancel the usual reflection at an interface, at least for one wavelength.

From an incoherent source, you cannot get interference between light coming from two different points on the source. In general, interference is obtained from light emitted by a single point on the source when that light suffers *division of amplitude*, part of the light being reflected at one surface and part at a following surface. The following general derivation gives the interference condition for many different interference arrangements involving reflection.

The extra path length between ray ② and ray ① can be found in terms of the difference in the optical path lengths along the corresponding paths. Optical path lengths are denoted by square brackets [].



Extra path length $= [ABC] - [AD]$
$= 2\ AB\ n_f - AC \sin\theta_i$
$= 2\ AB\ n_f - 2\ AB \sin\theta_t\ n_f \sin\theta_t$, from Snell's law: $\sin\theta_i = n_f \sin\theta_t$
$= 2\ n_f\ d\ (1-\sin^2\theta_t)/\cos\theta_t$, since $AB = d/\cos\theta_t$
$= 2\ n_f\ d \cos\theta_t$ since $1-\sin^2\theta_t = \cos^2\theta_t$.

Remember Lloyd's mirror result that if light is reflected at an interface between an optically less dense and a more dense medium (the higher refractive index) then it suffers a phase change of π. Hence the total path difference is:

$$2d\ n_f \cos\theta_t - \lambda/2 \quad .$$

When the path difference is $m\lambda - \lambda/2$, we will obtain destructive interference, corresponding to a dark fringe, which therefore occurs when:

$$2d\, n_f \cos\theta_t = m\lambda$$

i.e.

$$2d \cos\theta_t = m\lambda_f \text{, since } \lambda_f = \lambda/n_f$$

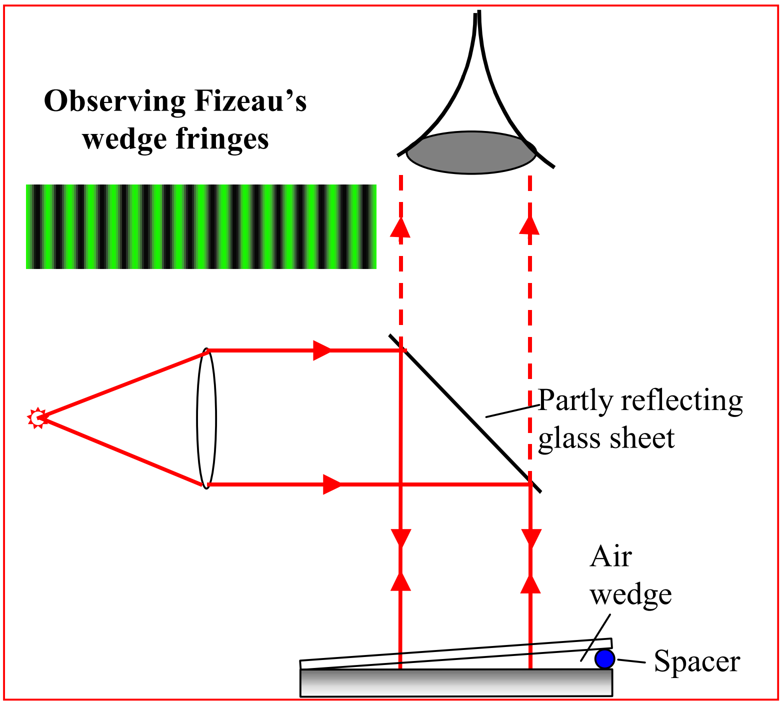
'm' is called the *order of the interference* fringe. For a constant 'd', the larger θ_t the smaller the order, i.e. the largest *order* occurs at normal incidence and the path difference decreases with increasing θ_t (or θ_i). The maxima are half-way between the minima. As before, 2-beam interference produces $\cos^2$ fringes.

A fringe of constant order has constant "d $\cos\theta_t$". If the film giving rise to the interference has a constant thickness, then the fringe will spread along the line of constant θ_t, i.e. constant θ_i. These are the **fringes of constant inclination**. If the source is extended, the fringes can be seen over a large area of film. See the diagrams in *Hecht*.

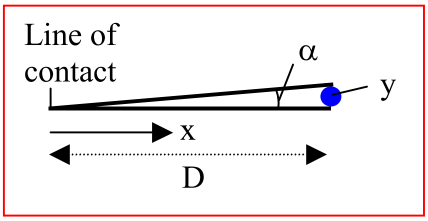
At normal incidence, $\cos\theta_t$ is constant around a circle away from the exact normal and you get circular fringes (Haidinger's fringes and circular Michelson interferometer fringes, see later).

Fringes of constant optical thickness

If $\cos\theta_t$ is effectively constant, then fringes will occur if the film thickness varies, or more exactly $d \times n_f$ is constant. You then get **fringes of constant optical thickness**. Examples of these are Newton's rings (see the PX2505 lab), Fizeau's wedge fringes and Michelson interferometer fringes with inclined mirrors, see later.



We'll look at one example, namely Fizeau's wedge fringes. You can see easily how they come about, and they are simple to set up. They can be used to measure very thin objects, the thickness of a hair or less, shown in the accompanying diagram as a *spacer*. The figure shows a common arrangement. Light from a monochromatic source is shone down via a glass plate onto a wedge-shaped film, made for example by laying one flat glass plate on top of another and then supporting one end with a thin spacer. The fringes are observed vertically, through the glass plate. In practice they are usually so close together that you use a low power microscope to see them.



Fizeau's wedge fringes are straight lines parallel to the line of contact, since the optical path increases linearly with distance x from the line of contact. The wedge angle α is given by

$$\alpha = y/D$$

From the general theory above (or 'obviously', as mathematicians are fond of saying) the fringe spacing is given by

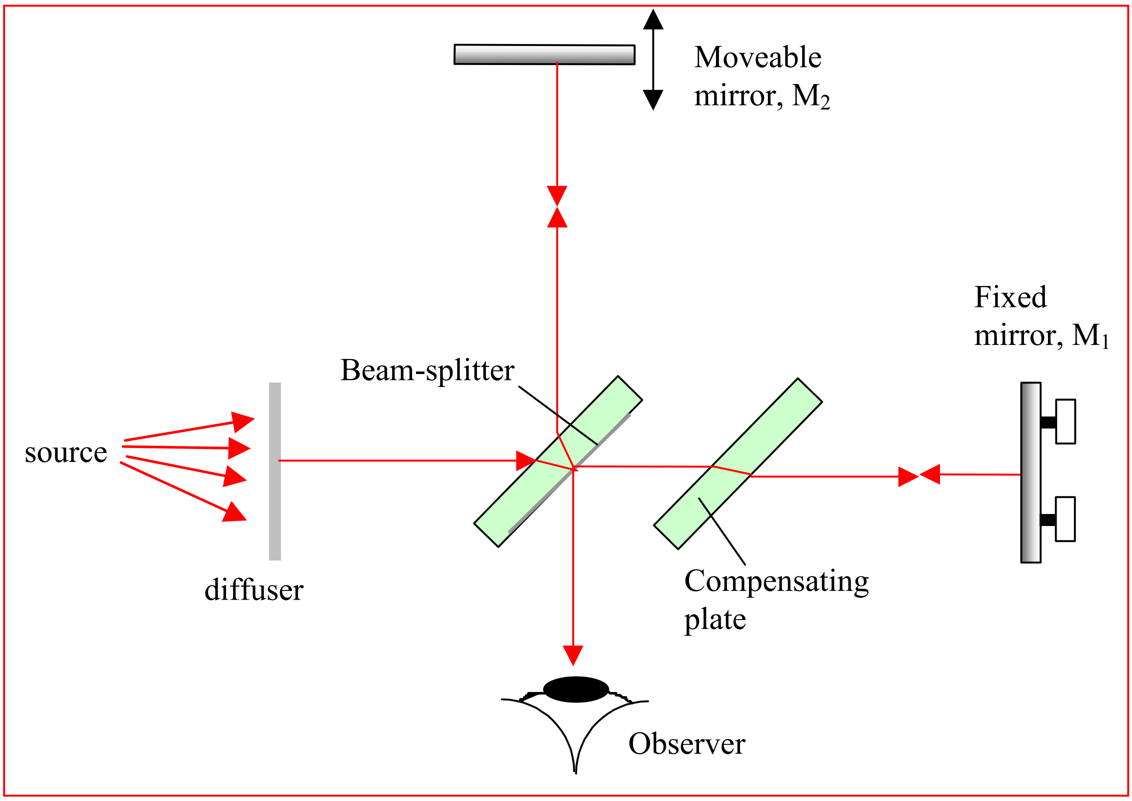
$$\Delta x = \lambda_f/2\alpha \quad .$$

In practice the angle has to be pretty small for the fringes to be visible to the naked eye. E.g. if the distance between neighbouring fringes is $\Delta x = 0.1$ mm for $\lambda_f = 500$ nm, then the angle $2\alpha = 5\times10^{-3}$ radians and hence $\alpha<0.2°$. If the spacer is a hair of thickness 75 μm, then D, the length of the glass plates, needs to be at least 30 mm long.

Michelson interferometer

The Michelson **interferometer** is one of the classic instruments of physical science. I've heard it said (by me) that anyone who considers themselves educated should know a little about the Michelson interferometer. In millennia to come, when the plays of Shakespeare are just a few hundred forgotten megabits in a cryogenic archive, the Michelson interferometer will still be well known. It was invented for the specific purpose of the Michelson-Morely experiment (see *Hecht* section 9.8) but can be considered an archetypal interferometer from which many others are derived. *Hecht*'s section 9.4 gives good diagrams of the Michelson interferometer. What it's used for will be described after we've explained what it is.

The device consists of a *diffuse source*, *beam-splitter*, two *mirrors* at the end of two perpendicular arms, a *compensating plate* in one arm to compensate for the thickness of the beam-splitter that is necessarily in the other arm. See the diagram below. The components



are optically simple. However, the equipment is very sensitive to vibration because a movement of one mirror of only 0.1 μm will produce an easily discernable movement of the fringes. Hence, careful mechanical construction is needed. Although the mirrors are physically perpendicular, they are optically parallel in that the image of M_1 in the beam-splitter is parallel to M_2 (and vice-versa).

What you see

When you first set up the interferometer, you usually see nothing in the way of fringes. This is because the mirrors are so out of parallel that any fringes are too fine to be visible, or the path difference, which is 2× the difference in the length of the arms, is greater than the coherence length of the source. The mirrors are brought into close alignment by superimposing the multiple reflections of a point source.

If the two mirrors are *optically parallel* but the arms are of different length, you see *fringes of constant inclination*, i.e. concentric circles (provided the path difference is less than the coherence length of the source). The bigger the difference in arm length, the more circles are seen within the field of view. Why is this?.

If the mirrors are *slightly inclined* by a single tilt, then Fizeau's *wedge fringes* are seen, i.e. parallel lines running in the direction of the optical contact line between the mirrors. These fringes are best seen with the path lengths almost equal and the relative mirror tilt very small.

Coloured w*hite-light fringes* are seen in either configuration when there is no path difference between the arms. This allows the zero condition of the interferometer to be found.

Use of the Michelson interferometer

1) Quantitative work is usually done by replacing the eye with a photo-detector and focusing the centre of the ring pattern onto the detector.

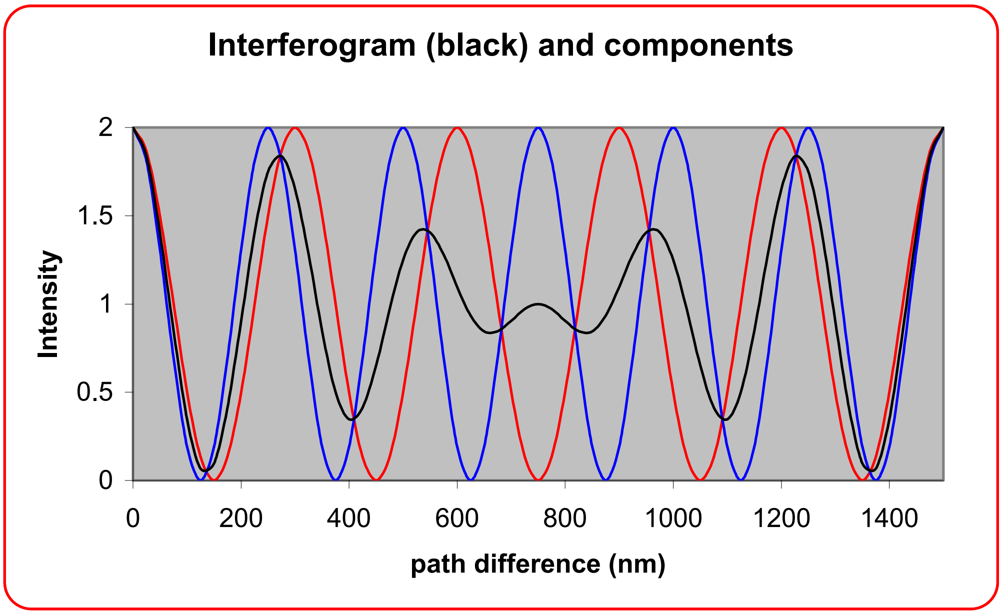
2) As one mirror is moved back, accurately parallel to itself, fringes pass at the rate of one per $\lambda/2$ mirror movement. Hence linear motion can be very accurately recorded (and converted into digital counting).

   a) Michelson counted the number of fringes of a particularly coherent cadmium (Cd) emission line in precisely 1 metre (there are over 3×10^6). The metre is now defined in terms of time of travel of light but is effectively measured by counting known wavelengths of stable light sources.
   b) The interferometer is a highly accurate optical ruler, particularly effective over modest distances. Maintaining parallelism of the mirrors during tracking is essential and is mechanically demanding. The ruler has a built-in zero, by virtue of the white-light fringes.
   c) The interferometer can be made the basis of very accurate ($\lambda/10$) position sensing. In a particular application, it may not be necessary to know the position in mm or μm but to sense only positional changes to high accuracy.

3) Replacing one mirror by a reflecting optical component allows a comparison of the surface profile of the component against a 'true' flat used for the second mirror.

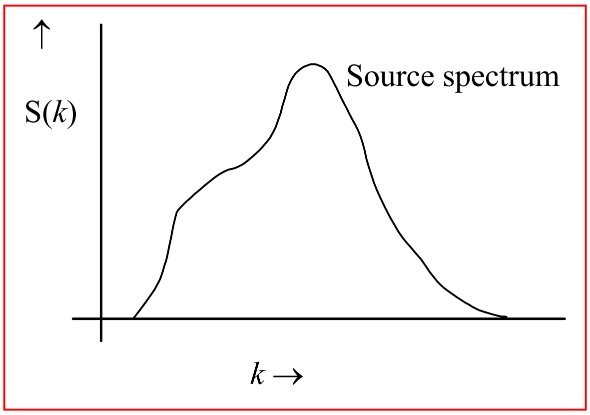
4) If the source of light is truly monochromatic, the irradiance pattern will vary sinusoidally with the mirror separation $\Delta\ell$ (remember $2I_o(1+\cos\delta) = 4I_o\cos^2\delta/2$). The pattern $I(\Delta\ell)$, which is the irradiance recorded as one mirror is moved, is called an *interferogram*.

5) If the source is not monochromatic, then there is a periodic variation within the interferogram for each frequency component present in the light. Suppose, for example, that the source has a red line of wavelength 600 nm and a blue line of wavelength 500 nm. The red set of fringes varies as $\cos(2\pi.2\Delta\ell/600)$, with $\Delta\ell$ in nm, and the blue line as $\cos(2\pi.2\Delta\ell/500)$. Both fringe patterns are in synchronism at $\Delta\ell = 0$ but gradually get out of phase as $\Delta\ell$ increases. They come back into synchronism again when $\Delta\ell = 1{,}500$ nm. This can be seen in the next sketch. In short, 6 fringes of $\lambda = 500$ nm will fit into the same path difference as 5 fringes from $\lambda = 600$ nm. By counting the path length until the fringes come into coincidence again, the ratio of their wavelengths can be found.



It is well-known that the bright yellow sodium D line contains two closely separated wavelengths. A prism spectrometer will not resolve these two lines. Yet a comparatively simple experiment with the Michelson interferometer can find the spectral line separation, even though it is less than 1 nm. This is done by moving one mirror and counting how many yellow fringes pass as the fringes change from excellent visibility (when the two fringe systems coincide) to poor visibility when they are out-of-phase, and back to excellent visibility again.

6) In general, the source will have a continuous spectrum of wavelengths. This is better expressed as a continuous spectrum of wavenumbers, k, because, as you can see from the previous sections, it is the wavenumber $k = 2\pi/\lambda$ that controls the interferogram. Let the spectrum of the source be $S(k)$, meaning that the strength of the source in the wavenumber range k to $k + dk$ is $S(k)dk$.



The following argument is for the mathematical to follow.

The interferogram I depends on the optical path difference between the arms, written as x = $2\Delta\ell$. The contribution to I(x) from a small range of wavenumbers d*k* is

$$d\mathrm{I}(x) = 2S(k)(1+\cos(kx))dk \quad .$$

[Remember the interference expression $2I_o(1 + \cos\delta)$ for two sources of equal intensity].

Hence, the full inteferogram is the integral of this expression across the source spectrum:

$$\begin{aligned}\mathrm{I}(x) &= \int 2S(k)(1+\cos(kx))dk \\ &= 2S_0 + 2\int S(k)\cos(kx)dk\ , \quad \text{where}\ \mathrm{S}_0 = \int S(k)dk\end{aligned}$$

$$\therefore \left(\mathrm{I}(x) - 2S_0\right) = 2\int_0^{\infty} S(k)\cos(kx)dk$$

i.e. $$S(k) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left(\mathrm{I}(x) - 2S_0\right)\cos(kx)dx \quad .$$

The spectrum S(*k*) is given by the Fourier transform of the variation of the interferogram (subtracting off the average). Hence, if we can record the fringe pattern over the whole range of x (= $2\Delta\ell$), then using standard Fourier transform techniques we can recover the source spectrum. In other words, we have a spectrometer, usually called a **Fourier transform spectrometer**.

The Fourier Transform (FT) Spectrometer is a powerful instrument.

a) Light from all parts of the spectrum is recorded simultaneously. The FT spectrometer is therefore good for weak sources. This advantage is known as the *multiplex advantage*.
b) Light from the full area of the mirrors is used, not simply light coming through a narrow slit. This is known as the *throughput advantage*.
c) The optical components are flat mirrors and a beam splitter – no expensive lenses are needed with their requirements of good optical homogeneity and accurate curved surfaces.
d) Since transmission through thick optical material is not needed, the FT spectrometer is a prime candidate for operating outside the visible wavelength range. Because of the stringent mechanical requirements that are related to the wavelength of the radiation used, the FT spectrometer has particularly made an impact in IR and far IR spectroscopy.
e) In practice you have to stop tracking the mirror after a certain range of mirror movement. This limits the resolution of the spectrum. By tracking over a small range, a low resolution spectrum can be quickly collected. By tracking over a longer range, very good resolution spectra can be collected without any new optical components. This is in marked contrast to the situation with prism and grating spectrometers.
f) With modern computing, the Fourier transform can be evaluated in 'real time', as the mirror is scanned, to give a rapid read-out. Repeated scanning improves the signal-to-noise ratio. The mirror position is tracked using laser light fringes (often He-Ne) from a small part of the mirror to ensure consistency of repeated tracking and accuracy in $\Delta\ell$.

Modern infra-red spectrometry is about identifying simple and complex molecular groups and determining the extent of their natural vibrations and rotations. The subject underpins a huge area of analytic science in the chemical and materials industries, in environmental studies and in a range of more exotic subjects such as investigations of the occurrence of molecules in space. The Fourier transform spectrometer is at the cutting edge of the tools available. It is almost magical how a high-resolution spectrum is revealed by the simple movement of just a plain mirror in a device with no prisms or diffraction grating. It works, brilliantly, thanks to our understanding of the interference of light.

End of Interference

JSR