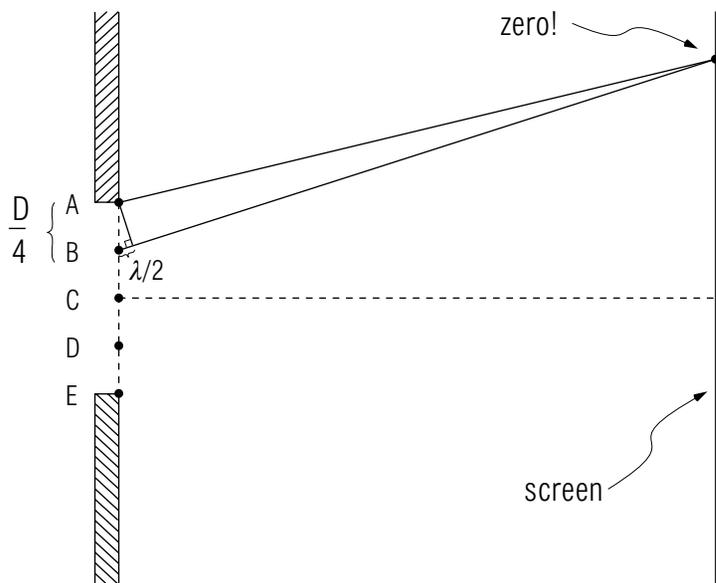


RESOLUTION AND DIFFRACTION



RESOLUTION AND DIFFRACTION

by
Fred Reif and Jill Larkin

CONTENTS

- A. Resolution
- B. X-Rays and the Structure of Matter
- C. Diffraction and Shadow Formation
- D. Diffraction Pattern from a Slit
- E. Ray Approximation for Waves
- F. Summary
- G. Problems

Title: **Resolution and Diffraction**

Author: F. Reif and J. Larkin, Department of Physics, Univ. of Calif., Berkeley.

Version: 5/1/2002

Evaluation: Stage 0

Length: 1 hr; 52 pages

Input Skills:

1. Vocabulary: coherent waves, incoherent waves (MISN-0-431).
2. Describe the interference far from N coherent sources in phase (MISN-0-231).

Output Skills (Knowledge):

- K1. Vocabulary: resolution, diffraction, ray.
- K2. Describe the diffraction of a wave by a hole or a slit.
- K3. Describe x-ray methods for determining the structure of matter.
- K4. State Huygen's principle.
- K5. Describe the interference effects responsible for brightness and shadow.
- K6. State the conditions under which waves may be approximated by rays.

Output Skills (Problem Solving):

- S1. Relate the approximate resolvable distance between two points to the wavelength of the waves used for the observations, to the detector size, and the detector position relative to the points.
- S2. For waves emanating from a hole or source of any size: (a) use interference arguments to describe qualitatively the resultant intensity at large distances from the source or hole; (b) relate the angular width of the emerging central beam to the wavelength and the size of the hole (exactly for a long slit and approximately for a circular hole).

THIS IS A DEVELOPMENTAL-STAGE PUBLICATION
OF PROJECT PHYSNET

The goal of our project is to assist a network of educators and scientists in transferring physics from one person to another. We support manuscript processing and distribution, along with communication and information systems. We also work with employers to identify basic scientific skills as well as physics topics that are needed in science and technology. A number of our publications are aimed at assisting users in acquiring such skills.

Our publications are designed: (i) to be updated quickly in response to field tests and new scientific developments; (ii) to be used in both classroom and professional settings; (iii) to show the prerequisite dependencies existing among the various chunks of physics knowledge and skill, as a guide both to mental organization and to use of the materials; and (iv) to be adapted quickly to specific user needs ranging from single-skill instruction to complete custom textbooks.

New authors, reviewers and field testers are welcome.

PROJECT STAFF

Andrew Schnepf	Webmaster
Eugene Kales	Graphics
Peter Signell	Project Director

ADVISORY COMMITTEE

D. Alan Bromley	Yale University
E. Leonard Jossem	The Ohio State University
A. A. Strassenburg	S. U. N. Y., Stony Brook

Views expressed in a module are those of the module author(s) and are not necessarily those of other project participants.

© 2002, Peter Signell for Project PHYSNET, Physics-Astronomy Bldg., Mich. State Univ., E. Lansing, MI 48824; (517) 355-3784. For our liberal use policies see:

<http://www.physnet.org/home/modules/license.html>.

MISN-0-432

RESOLUTION AND DIFFRACTION

- A. Resolution
- B. X-Rays and the Structure of Matter
- C. Diffraction and Shadow Formation
- D. Diffraction Pattern from a Slit
- E. Ray Approximation for Waves
- F. Summary
- G. Problems

Abstract:

The interference of waves has far-reaching implications and provides an understanding of some fundamental questions. In particular, we shall use the present unit to examine these questions: When one uses light or other kinds of waves to observe various objects, what limitations exist on how much detailed information can be obtained about such objects? Despite such limitations, how can one obtain detailed information about the arrangement of atoms in molecules of chemical or biological interest? Since sound and light are both waves, why is it that sound seems to spread out in all directions while light seems to travel along straight lines?

SECT.

A RESOLUTION

To use light or other kinds of waves to make observations of objects, the waves emanating from various parts of these objects are allowed to enter some detector (such as the eye). The information contained in the waves arriving at the detector is then used to infer information about the objects. In the following, we shall be primarily interested in *visual* observations so that the waves emanating from the objects are *light* waves. Our comments will, however, be generally applicable to all kinds of waves.

► *Resolution of two points*

Suppose that light waves emanate from two different points, e.g., from two different small parts of a bacterial cell. What then must be the minimum distance d_{\min} between these points so that observations of the waves from them allows one to recognize these points as distinct (i.e., as distinguishable from a single point)? We shall call d_{\min} the “minimum resolvable distance,” in accordance with the following meaning of “resolution”:

Def.	Resolution (or resolving power): The ability to use observations to recognize two different points as distinct.	(A-1)
------	--	-------

For example, to use visual observations to deduce information about different parts of a bacterial cell, these parts must be far enough apart so that they can be resolved.

► *Basic question*

To examine the resolution attainable by observations with some detector, we must answer this question: By what distance must two points be separated so that waves, emanating from these points and then entering a detector can be distinguished from waves emanating from a single point?

Note that the resolution of a given detector depends only on the waves *entering* the detector, irrespective of the complexity of the detector. (For example, the detector might be an eye, a camera, a microscope, a telescope, ...) Indeed, if the waves *entering* the detector cannot be distinguished from those originating from a single point, then nothing subsequently done to the waves in the detector can provide information distinguishable from the case of waves originating from a single point. (Of course, the converse is *not* true. Although the waves entering a detector

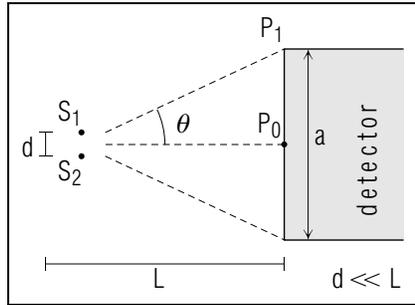


Fig. A-1: Two sources S_1 and S_2 observed by a detector

may be distinguishable from those originating from a single point, the detector may not be sufficiently well designed to reveal this information.)

► *Description of situation*

Let us then consider the situation illustrated in Fig. A-1. Spherical waves of wavelength λ emanate from two points S_1 and S_2 , separated by a distance d , which act thus as sources of the waves. The waves then enter a detector whose circular entrance (or “aperture”) has a diameter a . (For example, the aperture of an eye is its pupil. The aperture of a microscope is its “objective lens” close to the object under observation.) We need then to examine whether the waves w_1 and w_2 , arriving at the aperture of the detector after emanating from the two sources S_1 and S_2 , are distinguishable from the waves emanating from a single source.

► *Coincident sources*

Let us first examine the extreme case where the two sources S_1 and S_2 are so close together that they practically coincide. Then they can obviously *not* be distinguished from a single source emitting a single spherical wave. Indeed, in this case the relation between the phases of the waves w_1 and w_2 is the *same* at every point (equal to the relation between their phases at the sources, since each wave travels the same distance from the sources). This is why the resultant wave w , obtained by adding w_1 and w_2 , is merely equivalent to a single wave emanating from the position of the coincident sources.

► *Separated sources*

As we have just seen, the relation between the phases of the waves from two *coincident* sources is everywhere the same. (Thus, if the waves interfere constructively at one point, they also interfere constructively at all other points.) Hence two sources can only be recognized as distinct if the waves arriving from these sources at the detector are significantly different from the waves that would emanate from coincident sources. To be specific, we adopt the following approximate criterion of what we mean by “significantly different”: *The relation between the phases of the waves should differ by at least one-half cycle at different points of the detector aperture.* (If the waves interfere constructively at one point of the aperture, they should then interfere destructively at least at one other point.)

RESOLUTION OF A DETECTOR

Let us apply the preceding criterion to the situation illustrated in Fig. A-1. Here the point P_0 at the center of the detector aperture is equidistant from both sources. Hence the relation between the phases of the waves arriving at P_0 is the same as the relation between their phases at the sources. If the waves arriving at the aperture are to be significantly different from those originating from a single source, the relation between the phases of these waves at a point P_1 at the edge of the aperture should then approximately differ from that at the center point P_0 by at least a half cycle. This requires that the path difference of the edge point P_1 from the sources should differ at least by $\lambda/2$ from the zero path difference of P_0 from the sources. But the path difference of P from the sources is, by Relation (B-5) of Unit 431, equal to $d \sin \theta$, where θ is the angle indicated in Fig. A-1 (i.e., the angle subtended at the sources by the radius of the detector aperture). Thus the approximate requirement for distinguishability of the sources is that $d \sin \theta$ be at least as large as $\lambda/2$ i.e., that

$$d \sin \theta \geq \frac{\lambda}{2}. \quad (\text{A-2})$$

Hence $d \geq (\lambda/2)/\sin \theta$ and the *minimum* source separation d_{\min} , required for distinguishability of the sources, is approximately

$$d_{\min} \approx \frac{\lambda}{2 \sin \theta} \quad (\text{A-3})$$

► *Incoherent sources*

If the sources are incoherent, the relation between their phases varies in the course of time. But since the result Eq. (A-3) does *not* depend on the particular relation existing between the phases of the waves emitted by the sources, the preceding approximate argument remains valid at any time. Thus Eq. (A-3) is equally applicable to incoherent sources.

► *Resolution and λ*

The result Eq. (A-3) shows that the minimum resolvable distance d_{\min} depends crucially on the wavelength λ of the waves used to make the observations. In particular, one can resolve two points a smaller distance apart (i.e., the resolution is better) if the wavelength λ is smaller.

► *Resolution and detector size*

Suppose that observations are made with waves of some specified wavelength λ . Then better resolution should be provided by a detector with a larger aperture, since such a detector collects information about the waves over a larger area and can thus distinguish better whether these waves originate from two distinct points or not. The relation (A-3) is consistent with this conclusion. Indeed, if a detector has a larger aperture, the angle θ subtended by the detector in Fig. A-1 is larger. Hence $\sin\theta$ in the denominator of Eq. (A-3) is larger and the minimum resolvable distance d_{\min} is correspondingly smaller (i.e., the resolution of the detector is better).

► *Optimum resolution*

Suppose that the aperture of the detector is so large that the angle θ subtended in Fig. A-1 is 90° . Then $\sin\theta$ in Eq. (A-3) has its maximum value 1 and the minimum resolvable distance d_{\min} has its smallest possible value, $\lambda/2$. Thus

$$d_{\min} \approx \lambda/2 \quad (\text{A-4})$$

For example, the smallest wavelength of light visible to the human eye is about 4×10^{-7} m or 4000 Å (where $\text{Å} \equiv 10^{-10}$ m). By Eq. (A-4), the minimum resolvable distance which can be achieved with visible light *under the most optimal conditions* is then approximately $d_{\min} \approx \lambda/2 = 2000 \text{ Å}$. Note that this distance is very much larger than 1 Å, the typical separation of atoms in a molecule. Thus one cannot possibly use observations with visible light to distinguish the positions of individual atoms in a molecule.

► *Resolution vs. magnification*

Note that the limitation on resolution is a fundamental limitation due to the wavelength of the waves used to make observations. Hence this limitation cannot be overcome by any instrumentation, no matter how fancy. For example, suppose that light from two distinct small regions of a bacterial cell enters a microscope. If these regions of the cell are closer than $\lambda/2$, the light waves entering the microscope are indistinguishable from those originating from a single small region. Although the microscope may provide enormous magnification, the light waves finally leaving the microscope and entering the eye are then still indistinguishable from those originating from a single small region. Thus the magnification provided by the microscope helps the eye see a single enlarged region, but cannot help the eye see two distinct regions.

RESOLUTION OF A DETECTOR FOR DISTANT OBJECTS

Suppose that the diameter a of the aperture of a detector is much smaller than the distance L of the detector from the sources observed with this detector. (For example, this would certainly be the situation in the case of a telescope used to observe two distant objects.) Then the resolution of the detector is limited by the fact that it collects only a small fraction of the information provided by the waves emanating from the sources. Hence the minimum resolvable distance d_{\min} is much larger than the wavelength.

► *Resolvable distance*

To be explicit, if $a \ll L$, the angle θ subtended at the sources by the radius of the detector aperture is small. As is seen from Fig. A-2a, the sine of this angle is then approximately $\sin\theta = (a/2)/L$ since the radius of the aperture is $a/2$ and since the distance from the sources to the edge of the aperture is almost the same as the distance L to the center of the aperture. Hence the result Eq. (A-3) for the minimum resolvable distance then becomes

$$d_{\min} \approx \frac{\lambda}{2 \left(\frac{a}{2L} \right)} = \frac{\lambda}{\left(\frac{a}{L} \right)}$$

Multiplying both numerator and denominator by L then yields

$$\boxed{d_{\min} \approx \lambda \left(\frac{L}{a} \right) = L \left(\frac{\lambda}{a} \right)} \quad (\text{A-5})$$

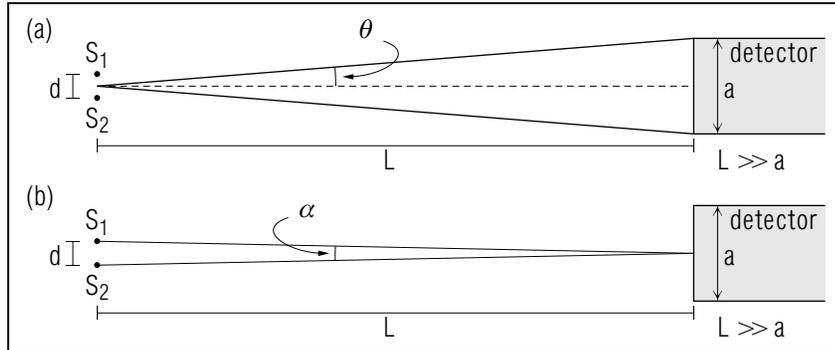


Fig. A-2: Detector with an aperture diameter a small compared to its distance L from two sources: (a) Angle θ subtended at the sources by the aperture radius; (b) Angle α subtended at the detector by the sources.

► *Resolvable angle*

The resolution of distant sources can also conveniently be expressed in terms of the angle α subtended at the detector by the sources. As seen from Fig. A-2b, this angle (measured in radians) is just $\alpha = d/L$. The result Eq. (A-5) specifies then the minimum resolvable angle α_{\min} subtended at the detector by the sources, i.e.,

$$\alpha_{\min} = \frac{d_{\min}}{L} \approx \frac{\lambda}{a} \quad (\text{A-6})$$

This minimum resolvable angle is thus determined by the ratio of the wavelength compared to the diameter of the detector aperture.

Resolution, Wavelength, and Detector Properties (Cap. 1)

A-1 *Acoustic microscope:* Recently an acoustic microscope has been developed which allows one to resolve distances as short as 10^{-6} m and thus to use sound waves in order to study and visualize details of a bacterial cell. (a) What must be approximately the largest possible wavelength of the sound waves required to achieve such a resolution? (b) What must be the corresponding frequency of these sound waves if these waves are used to observe cells suspended in water in which the speed of sound is 1500 m/s? (*Answer: 5*)

A-2 *Resolution of an optical microscope:* Suppose that an optical microscope is used with light of a given wavelength to observe bacterial cells. (a) If the objective lens of the microscope is designed to be at a given distance from the cell, is the resolution of the microscope better (i.e., is the minimum resolvable distance smaller) if the objective lens has a larger or smaller diameter? (b) For an objective lens of given diameter, is the resolution better if the objective lens is designed to be placed farther from, or closer to, the cell? (c) For a given diameter and position of the objective lens, is the resolution better if one observes the cell with red light (having a wavelength of 7×10^{-7} m) or with blue-green light (having a wavelength of 5×10^{-7} m)? (d) In the preceding case, what is the ratio d_b/d_r of the minimum resolvable distance d_b obtainable with blue-green light as compared with the minimum resolvable distance d_r obtainable with red light? (e) What is approximately the minimum resolvable distance obtainable with blue-green light if the objective lens has a diameter of 0.8 cm and is placed at a distance of 0.3 cm from the bacterial cell? (*Answer: 2*) (*Suggestion: [s-3]*)

A-3 *Oil immersion microscope:* A sample is observed under a microscope with light of a given color (i.e., of a given frequency). Suppose now that the region between the sample and the objective lens of the microscope is filled with transparent oil in which the speed V' of light is smaller than the speed V of light in air. (a) Is the wavelength λ' of the light in the oil then larger or smaller than the wavelength λ of the light in air? (b) Is the minimum resolvable distance d'_{\min} between two points in the sample then larger or smaller than the minimum resolvable distance d_{\min} in the absence of the oil? (c) If $V' = 0.6V$, express the minimum resolvable distance d'_{\min} (with the oil) in terms of d_{\min} . (The preceding comments illustrate the practical utility of oil immersion for improving the resolution of an optical microscope.) (*Answer: 1*)

A-4 *Observation of objects on the moon:* The moon is at a distance of approximately 4×10^8 m from the earth. Estimate the minimum distance by which two objects on the moon must be separated so that they can be recognized as distinct when they are observed from the surface of the earth under ideal conditions (no air turbulence or smog), with light having a wavelength of 5×10^{-7} m: (a) When the observations are made with the unaided eye, having a pupil diameter of 5 mm. (b) When the observations are made with the Mt. Palomar telescope having a diameter of 200 inch (i.e., about 5 meter). (*Answer: 6*)

A-5 *Satellite detection of pollution on the earth:* It is desired to use a satellite, orbiting around the earth at an altitude of 300 km above its surface, to detect sources of pollution (such as effluents from smokestacks) on the surface of the earth. If one wants to pinpoint pollution sources which may be no more than 5.0 m apart, what must be the approximate size of the lens of the TV camera used in the satellite to observe the earth? Neglect disturbing effects of the earth's atmosphere and assume that the wavelength of the light used to make the observations is 5.5×10^{-7} m. (*Answer: 4*)

More practice for this Capability: [p-1], [p-2], [p-3]

SECT.

B X-RAYS AND THE STRUCTURE OF MATTER

The properties of matter depend crucially on the spatial arrangement of the constituent atoms. For example, the biological functioning of protein molecules, such as enzymes, depends sensitively on the precise three-dimensional configuration of the atoms in these molecules. How then can one determine the relative positions of atoms in matter?

► *Required wavelength*

The typical separation between atoms in matter is about 1 \AA (i.e., 10^{-10} m). To distinguish between the positions of neighboring atoms, the observations of these atoms must thus be sufficiently refined to provide a minimum resolvable distance d_{\min} at least as small as 1 \AA . But we know from Eq. (A-4) that, under optimum conditions, d_{\min} is related to the wavelength λ of the waves used to make the observations so that $d_{\min} \approx (1/2)\lambda$. Hence the wavelength of the waves needed to make the observations must be at least as small as $\lambda \approx 2d_{\min} \approx 2 \text{ \AA}$. This wavelength is more than 1000 times smaller than that of visible light. Electromagnetic waves with such a small wavelength are called X-rays.*

* Such X-rays are produced by accelerating electrons to high energies and then letting these electrons strike a metal target. Because the electrons are very rapidly decelerated as they are stopped in the target, they emit electromagnetic waves with the very high frequencies (or correspondingly small wavelengths) characteristic of X-rays.

► *Interference from a molecule*

Figure B-1 illustrates schematically how one might, in principle, use X-rays to determine the arrangement of atoms in a molecule. A beam of X-rays is incident on a sample consisting of a single molecule. The sinusoidally varying electric field associated with the X-rays produces then a correspondingly varying force on, and thus a corresponding acceleration of, every electron in the sample.

Every such electron emits accordingly an electromagnetic wave of the same frequency (or wavelength) as that of the incident X-rays, and synchronized in phase with the incident X-rays. Hence the electrons in the sample act as coherent sources of X-rays and the waves emitted from them can interfere. If the intensity of the X-rays emerging from the sample is

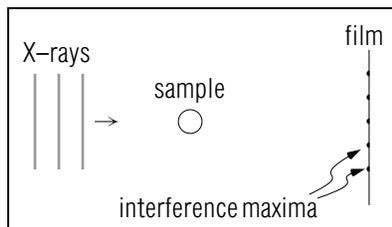


Fig.B-1: Experiment for studying the interference of X-rays passing through a sample.

observed with some detector (such as a photographic film), the intensities observed at various points exhibit thus characteristic intensity maxima due to interference (see Fig. B-1).

Observations of this interference pattern then allows one to deduce information similar to that which can be deduced in the simple case of two sources (as discussed at the end of text section C of Unit 431): (1) The relative positions of the observed interference maxima allow one to deduce information about the relative positions of the electrons in the atoms of the molecule. (2) The relative intensities of these intensity maxima allow one to deduce information about the relative amplitudes of the waves emanating from various atoms (i.e., about the relative number of electrons in these atoms), information sufficient to identify the atoms in which the electrons are located.

► *Interference from a crystal*

The preceding experiment is, in practice, impossible because the intensities produced by the electrons in a *single* molecule would be far too small to be observable. Hence one must use a sample consisting of many identical molecules, provided that these molecules are arranged in a regular way so that the waves from different molecules interfere constructively (thus enhancing, rather than destroying, the interference observed from a single molecule). Such a sample of regularly arranged molecules can be obtained if one can grow a single crystal of a material consisting of these molecules. Then the interference experiment becomes quite practical. Indeed, just as for the N coherent sources discussed in text section D of Unit 431, the intensity I of an interference maximum due to N regularly spaced molecules is N^2 times larger than the intensity I_1 due to a single molecule (since the amplitude of the interference maximum is N times larger than that due to a single molecule). If the number N of molecules in a sample is in the vicinity of Avogadro's number (e.g., if $N \approx 10^{23}$) the intensity I is thus about 10^{46} times larger than the intensity I_1 from a single molecule. This gigantic increase in intensity thus makes the X-ray

interference quite observable, despite the small intensity produced by the X-rays emanating from a single molecule.

► *Features of the X-ray method*

The preceding method of using X-ray interference for determining the arrangement of atoms in molecules and crystals is called "X-ray diffraction." The method is extremely important since it is essentially the only method capable of providing detailed information about the three-dimensional arrangement of atoms in complex molecules. Although the method has been extensively used, it is difficult for the following reasons:

- (1) It is necessary to grow a single crystal containing the molecules of interest. (Such crystal growing is fairly simple in the case of sodium chloride, but is much more difficult for complex molecules such as hemoglobin.)
- (2) Complicated calculations are required to use information from interference patterns (for various orientations of the crystalline sample with respect to the incident X-rays) in order to deduce information about the spatial arrangement of the atoms in a complex molecule. Modern electronic computers have, however, made the calculational tasks much more manageable.
- (3) An unambiguous determination of an atomic arrangement requires extensive data collection and some ingenuity, since the interference pattern provides only partial information about the waves emanating from the sample. (The information is partial since one can only measure the *intensities* of the waves in the interference pattern, but has no way of measuring their relative phases).

► *Applications*

Nevertheless, the method of X-ray diffraction has provided extremely important information about the structure of crystals and of complex molecules. For example, the method has provided information about the helical structure of the DNA molecule responsible for genetic replication, and also about the atomic arrangement of atoms in important proteins such as myoglobin, hemoglobin, lysozyme, Figure B-2a illustrates an X-ray interference pattern produced from a single crystal of sodium chloride. Figure B-2b shows a model of the crystal structure of sodium chloride, as deduced from such interference patterns. Figure B-3a illustrates an X-ray interference pattern produced from a single crystal of myoglobin (the protein responsible for storing oxygen in muscle tissues).

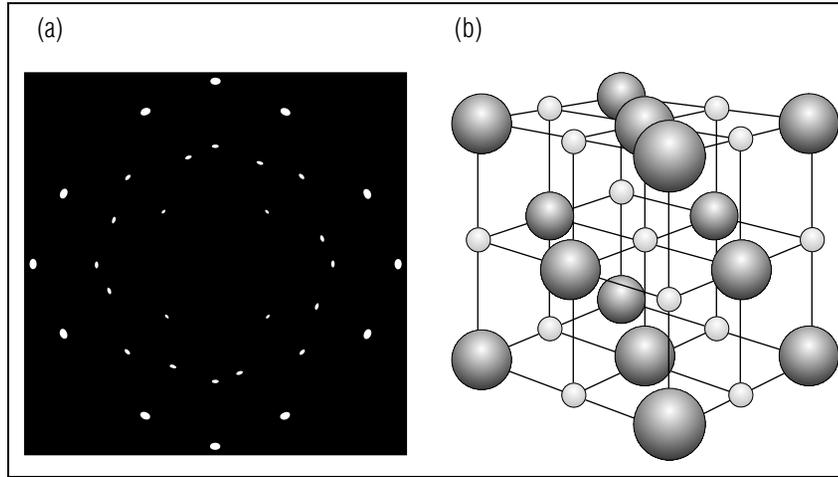


Fig. B-2: X-ray diffraction from a single crystal of sodium chloride. (a) Typical X-ray interference pattern. (b) Model showing the inferred crystal structure of sodium chloride.

Figure B-3b shows a model of the very complex structure of myoglobin, as deduced from such interference patterns.

B-1 *Illustration:* Fig. B-4 illustrates a beam of X-rays incident upon a single molecule, consisting of two atoms 1 and 2, oriented so that the line joining the atoms is perpendicular to the direction of the incident beam. Since the electric field of the incident X-rays causes the electrons in these atoms to oscillate with the frequency of the incident X-rays, these atoms then act as sources emitting in phase X-rays of the same wavelength as the incident X-rays.

The amplitude of the wave emitted by each atom is proportional to the number of electrons in that atom. As a result of interference, one then observes on a photographic plate, at a distance L behind the molecule, a series of interference maxima separated by a distance s much smaller than L . Furthermore, the maximum intensity I_{\max} observed on this plate is 4 times larger than the minimum intensity I_{\min} observed on this plate. (a) On the basis of this information, what is the separation between the two atoms in the molecule? (b) Is the ratio N_1/N_2 of the number N_1 of electrons in atom 1, compared to the number N_2 of electrons in atom 2, equal to 1,2,3, or 4? Justify your answer. (*Answer: 8*) (*Suggestion: [s-1]*)

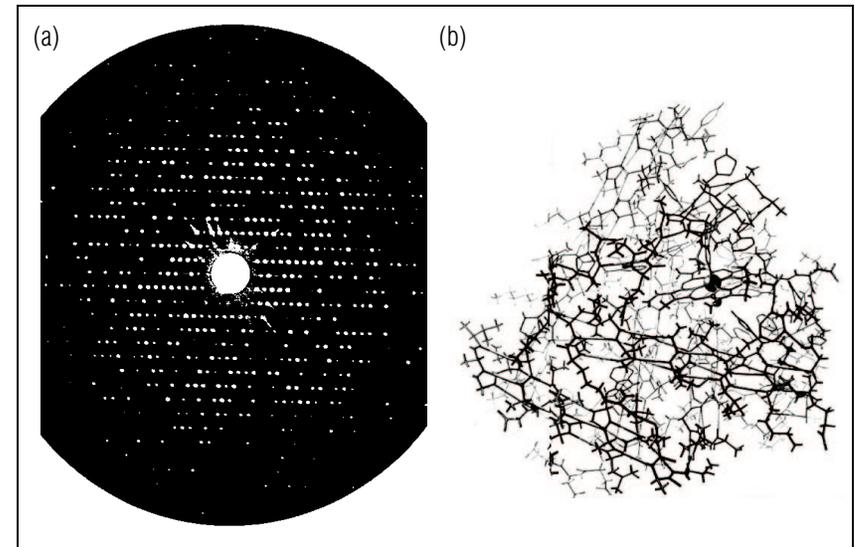


Fig. B-3: X-ray diffraction from a single crystal of myoglobin. (a) Typical X-ray interference pattern. (b) Model showing the inferred structure of a myoglobin molecule.

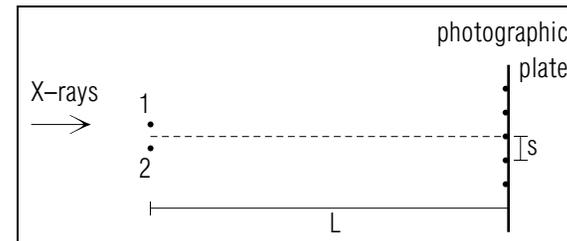


Fig. B-4.

SECT.

C DIFFRACTION AND SHADOW FORMATION

The word “diffraction” is used to describe this phenomenon:

Def.	Diffraction: The spreading of a wave in various directions when the wave emanates from an object.	(C-1)
------	--	-------

In particular, let us examine the common situation where a wave is incident upon an opaque screen containing a hole. How does such a wave spread out in various directions behind the hole? (Or, in fancy language, how is the wave “diffracted” by the hole in the screen?)

► *Very small hole*

Suppose that the hole is very small compared to the wavelength λ of the wave. Then (as already discussed in text section D of Unit 431) the disturbance arriving at the hole sets up a similar disturbance in the immediate vicinity, which sets up a similar disturbance in its vicinity, The result is then a wave which spreads out behind the hole in all directions. In other words, the hole acts like a “point source” (i.e., source small compared to λ) emitting a spherical wave. The frequency and phase of this source are just those of the original wave producing the disturbance at the hole.

► *Huygens’ principle*

How then can we understand the situation where the hole is not small, but may have any size whatever? We need only imagine that the hole is subdivided into many small parts, each part small compared to the wavelength. Then each such part should act as a point source of waves. Accordingly, the resulting wave at any point behind the screen should just be the superposition of the waves emanating from all parts of the hole. [The idea that any wave can be regarded as equivalent to a collection of point sources is known as “Huygens’ principle” since the idea was first formulated by the Dutch physicist and astronomer, Christian Huygens (1629-1695).]

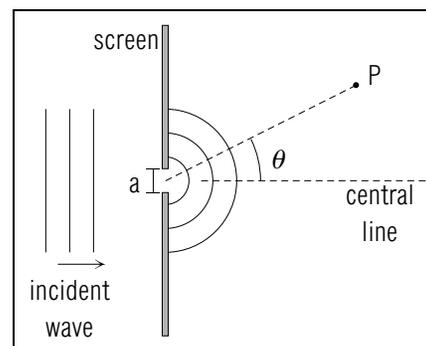


Fig. C-1: Wave incident on a hole in an opaque screen.

► *Description of situation*

Huygens’ principle can be used to discuss easily the waves emanating from any hole. In particular, we shall consider a plane opaque screen with a circular hole of diameter a . A wave of wavelength λ is incident on this hole so that the wavefronts of the wave are planes parallel to the hole (see Fig. C-1.) At any instant, the phase of the wave at all parts of the hole is then the same. If the hole is imagined to be subdivided into many *equal* parts, each small compared to λ , these parts then act as point sources with the same amplitude and phase. The resultant wave observed at any point P behind the screen is then the sum of the spherical waves emanating from all these parts of the hole. In general, we may thus expect to observe interference effects in the intensity observed at P . The nature of these interference effects depends on the relation between the size of the hole and the wavelength. Let us then consider qualitatively the results to be expected in various cases.

HOLES OF VARIOUS SIZES COMPARED TO THE WAVELENGTH

► $a \ll \lambda$

Consider first the case of a hole of diameter a much smaller than the wavelength λ . Then the path difference of any point P from any two parts of the hole is always much smaller than λ . Hence the waves, leaving any two parts of the hole in phase, arrive at any point P also in phase and interfere there constructively. The resultant wave at P is then just the sum of all the equal waves arriving from all the parts of the hole. Hence this resultant wave is merely a large spherical wave spreading out in all directions behind the hole.

► $a \gg \lambda$

By contrast, consider the case of a hole of diameter a much larger than the wavelength λ . We are again interested in the waves arriving from various parts of the hole at any point P (which we assume to be far from the hole, at a distance large compared to a .) Suppose that the point P is along the central line indicated in Fig. C-1. Then the path difference of P from any two parts of the hole is nearly zero (since P is nearly equidistant from all parts of the hole). Hence the waves from all parts of the hole arrive at P with the same phase and thus interfere there constructively to yield a large intensity. On the other hand, consider a point P appreciably away from the central line (so that the line from P to the hole makes an appreciable angle θ with respect to the central line). Since many parts of the hole are separated by a distance d large compared to λ the path difference $d \sin \theta$ of P from any two parts of the hole may easily be as large as $1/2 \lambda$. The waves from such parts of the hole then interfere destructively so as to cancel each other. Hence the sum of the waves arriving at P from all parts of the hole has a small magnitude since many of these waves have opposite signs and thus interfere destructively. Consequently, the intensity of the resultant wave at P is much smaller than that of the resultant wave at a point along the central line.

In sum, one should observe behind the screen a large intensity (e.g., bright light) along the central line directed along the velocity of the wave incident on the hole. On the other hand, one should observe a very small intensity (i.e., darkness or shadow) at points far from the central line. Thus the incident wave seems to travel through the hole predominantly along a straight line (directed along the incident velocity of the wave) and there is only a little spreading of the wave away from this direction.

► *Light vs. sound waves*

The preceding comments show that the spreading of a wave passing through a hole depends crucially on the size of the hole compared to the wavelength. The different behavior of sound and light waves is correspondingly merely a result of their very different wavelengths.

For example, the wavelength of sound waves in air is typically around 1 meter and is thus much larger than the size of many holes (such as a keyhole in a door). Hence a sound wave incident on such a hole spreads out behind the hole in all directions, so that the sound can be heard anywhere behind the hole.

On the other hand, the wavelength of light is smaller than 10^{-6} meter and is thus very much smaller than the size of most holes (such as the

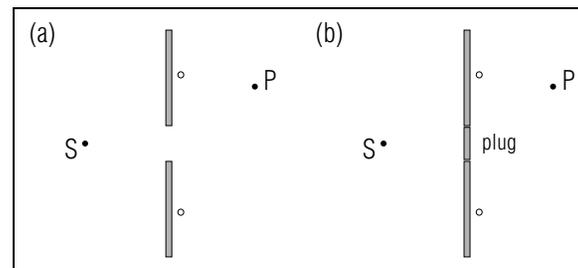


Fig. C-2.

same keyhole). Hence a light wave incident on such a hole travels through the hole, with negligible spreading, along a central straight line. Correspondingly, a dark region (or shadow) is produced behind the hole away from this central line so that no light can be seen there. (As we have discussed, the straight-line propagation and resulting shadow formation are ultimately due to the destructive interference of the light waves from different parts of the hole.)

Diffraction From a Hole Or Source (Cap. 2)

C-1 *Diffraction of sound waves passing through a hole:* The frequency of the sound wave of lowest pitch (the tone of C) produced by a cello is about 65 Hz. A sound wave of frequency $2^5 = 32$ times larger corresponds to one of the high tones produced by a violin (the tone of C which is 5 octaves above the cello C, or one octave above the “high C” produced by a soprano.) (a) If the speed of sound in air is 340 m/s, what are the wavelengths of the sound waves corresponding to this lowest cello tone and this high violin tone? (b) Suppose that such sound waves are incident upon a wall in which there is an open window about 1 meter on each side. Far behind the window, would the sound waves corresponding to the cello tone be heard almost equally well in all directions behind the window, or would they be heard predominantly only along the direction of the incident sound wave? (c) Answer the same question for the sound waves corresponding to the violin tone. (*Answer: 3*)

C-2 *Relation between holes and sources (alternative justification of Huygen’s principle):* Let us analyze in detail what happens when a wave, emanating from source S , arrives at an opaque screen containing a hole, as illustrated in Fig. C-2a.

Because of the forces exerted on every particle in the screen by the waves arriving at this particle from S (and from the other

particles), every such particle then oscillates and thus in turn emits a wave. Hence the total amplitude A of the resulting wave arriving at some point P behind the screen is equal to $A = A_S + A_0$, where A_S is the amplitude of the wave arriving at P from the source S and where A_0 is the total amplitude of the waves arriving at P from all the particles in the screen *outside* the hole. (a) Suppose that the hole is filled with a plug, as shown in Fig. C-2b. Then the waves arriving at P from the particles in the plug also produce at P a wave with a total amplitude A_p . What then is the total amplitude of the wave produced at P by the source S and all the particles in the entire plugged-up screen? Express your result in terms of A_S , A_0 , and A_p (assuming that the waves emitted from the parts of the screen outside the original hole remain essentially unaffected by the presence of the plug). (b) Since the entire plugged-up screen is *opaque*, what should be the total amplitude of the wave observed at P ? (c) Combine the answers for parts *a* and *b* to write an equation relating A_S , A_0 , and A_p . (d) By using this equation, express the original amplitude $A = A_S + A_0$, observed when the screen has a hole, in terms of A_p . (e) By using this result, express the intensity I , observed at P in the situation of Fig. C-2a where the screen has a hole, in terms of the intensity I_p which would be observed at P due to the waves emanating solely from the plug. (f) Hence compare the intensity produced by waves emanating through a hole in an opaque screen with the intensity produced by the waves emanating solely from a plug of the same size as the hole. (*Answer: 9*) (*Suggestion: [s-5]*)

C-3 *Diffraction of sound waves emitted by a loudspeaker:* When a diaphragm oscillates back and forth, the particles in the diaphragm emit sound waves in phase. The interference effects thus produced are then the same as those from sound waves emanating from a hole where, by Huygens' principle, waves emanate in phase from the different parts of the hole. (This conclusion was also explicitly established in the preceding problem.) Suppose that the diaphragm of a loudspeaker has a diameter of 10 cm and, when vibrating, emits sound waves having a speed of 340 m/s in air. (a) If the diaphragm vibrates at a frequency of 10^2 Hz so as to produce a tone of low pitch, what is the wavelength of the sound waves in air? Do these sound waves spread out uniformly in all directions from the speaker, or do they spread out predominantly along the forward direction perpendicular to the oscillating diaphragm? (b) Answer the same questions if the speaker diaphragm oscillates at a frequency of 2×10^4 Hz so as to produce a tone of the highest audible pitch. (*Answer: 7*)

SECT.

D DIFFRACTION PATTERN FROM A SLIT

► Description of situation

To gain a better understanding of what happens when a wave passes through a hole, let us look more closely at the transition between brightness and shadow produced behind the hole. For the sake of simplicity, we consider an opaque screen which contains a long *slit*, of width a perpendicular to the plane of the paper in Fig. D-1. A wave of wavelength λ is incident upon this screen, with a velocity perpendicular to the screen, so that the wave fronts are parallel to the screen. The wave at all parts of the slit has thus the same phase. We may then imagine that the slit is subdivided into many parts, each a thin strip parallel to the edges of the slit and of width much smaller than λ (see Fig. D-1).

In accordance with Huygens' principle, these strips can then be viewed as linear sources of waves with the same phase and the same amplitude.*

* The intensity produced by waves from a long slit is the same at all points of a line through P parallel to the slit. It is slightly easier to discuss a long slit of width a , rather than a circular hole of diameter a , because such a slit can be subdivided particularly simply into parts of equal area. However, the discussion of a circular hole is similar and leads to results which are qualitatively similar.

What then is the resultant intensity produced by these waves at any point P behind the screen? For simplicity, we shall only consider points P so

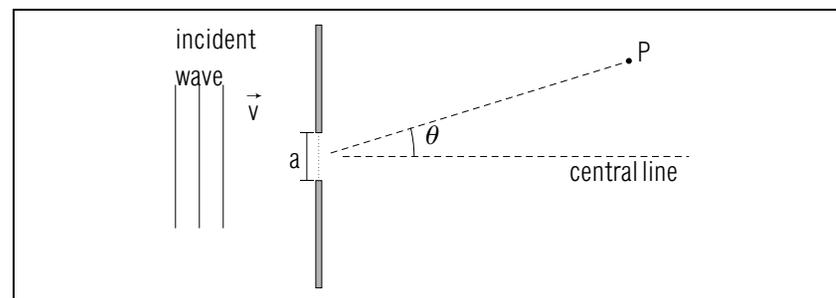


Fig. D-1: Wave incident on a long slit in an opaque screen. The slit can be imagined subdivided into many narrow strips, each of which acts as a source of waves.

far from the slit that their distance from the slit is large compared to the slit width a .

► *Interference from two strips*

As we discussed in the preceding section, the waves emanating from two parts of the hole (i.e., from two strips of the slit) may interfere with each other. Indeed, since the waves from any two strips have equal amplitudes, the waves from two strips cancel each other *completely* by destructive interference at any point P whose path difference from the two strips is equal to $\lambda/2$. But, by Relation (B-5) of Unit 431, this path difference is equal to $d \sin \theta$ where d is the separation between the two strips and θ is the angle between the central line and the line from the slit to P . If the waves from two strips are to cancel each other at a point P at an angle θ , the distance d_c between these strips must then be such that

$$d_c \sin \theta = \frac{\lambda}{2} \quad (\text{D-1})$$

We shall call d_c the “cancellation distance.”

Let us then systematically examine the intensity of the resultant wave produced behind the screen at various points P specified by successively larger angles θ . According to Eq. (D-1), these successively larger angles θ correspond then to successively smaller cancellation distances d_c for complete destructive interference of waves from two strips.

► $\theta \approx 0$

Suppose that $\theta \approx 0$ so that the point P is nearly along the central line. Then the cancellation distance d_c is much larger than the width of the slit, i.e., no destructive interference is ever observed from waves coming from any two strips within the slit. Indeed, since any point P along this central line is nearly equidistant from all strips of the slit, the waves from all these strips arrive at P in phase and thus interfere there *constructively* to produce a resultant wave of the largest possible amplitude A_0 and correspondingly largest possible intensity I_0 .

► θ for $d_c = a$

Consider now a point P at a larger angle $\theta = \theta_1$ such that the cancellation distance d_c is just equal to the width a of the slit. (See Fig. D-2a.) Then the waves arriving at P from the two strips adjacent to the edges of the slit cancel each other. But the waves arriving from strips nearer the center of the slit are less than $1/2$ cycle out of phase and thus add so as to produce a resultant intensity which is appreciable, although

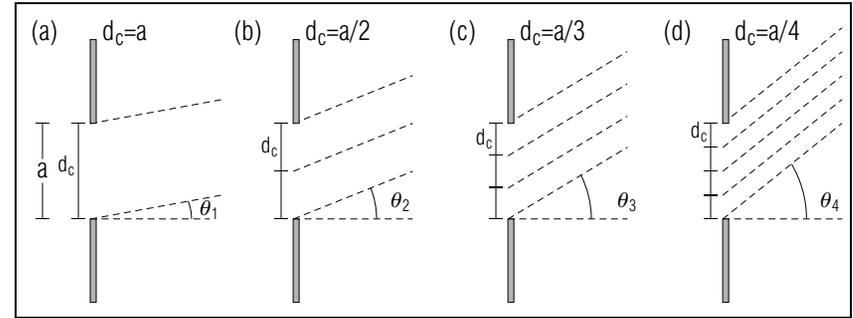


Fig. D-2: Observation points at various angles θ from the central line, corresponding to various values of the cancellation distance d_c .

smaller than the intensity I_0 along the central line.

► θ for $d_c = a/2$

Consider now a point P at a still larger angle $\theta = \theta_2$ such that the cancellation distance $d_c = a/2$, i.e., such that the slit width a is twice as large as the cancellation distance d_c . (See Fig. D-2b.) Then the wave from the strip near the top edge of the slit cancels the wave from the strip in the middle of the slit; the wave from the strip immediately below the top strip cancels the wave from the strip immediately below the middle strip (since these two strips are also separated by the distance d_c); and so forth. In other words, the waves from all such corresponding strips cancel each other. Hence all the waves coming from the first half of the slit cancel all the waves from the second half of the strip. Thus the resultant wave observed at P has *zero* intensity.

► θ for $d_c = a/3$

Consider now a point P at a still larger angle $\theta = \theta_3$ such that the cancellation distance $d_c = a/3$, i.e., such that the slit width a is 3 times as large as the cancellation distance d_c . (See Fig. D-2c.) Then an argument similar to the preceding one shows that the waves coming from the first third of the slit cancel the waves coming from the second third of the slit. Thus only the waves coming from the last third of the slit contribute to the resultant intensity observed at P . Although this intensity is not zero, it is considerably smaller than the intensity I_0 along the central line.

► θ for $d_c = a/4$

Consider now a point P at a still larger angle $\theta = \theta_4$ such that the cancellation distance $d_c = a/4$, i.e., such that the slit width a is 4 times

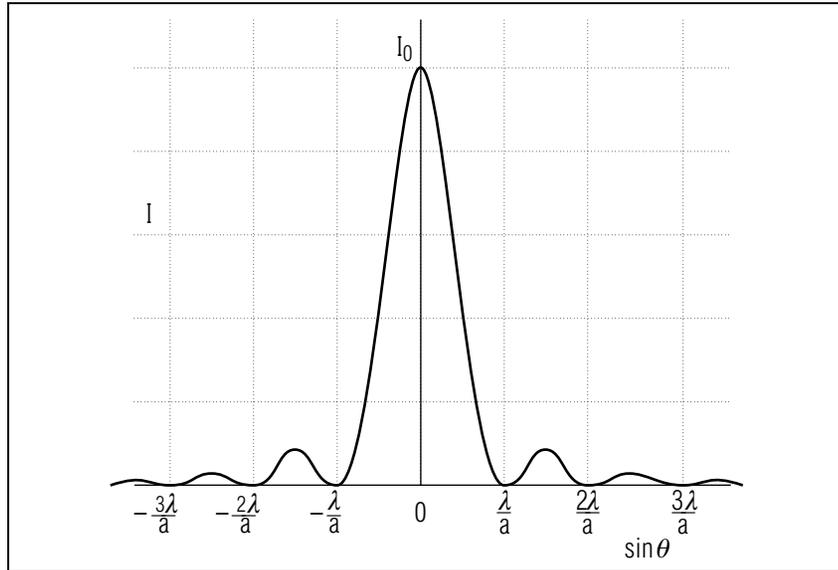


Fig. D-3: Intensity I , at various angles θ from the central line, observed behind a narrow slit of width a when a wave of wavelength λ is incident on this slit.

as large as the cancellation distance d_c . (See Fig. D-2d.) Then the waves from the first quarter of the slit cancel those from the second quarter of the slit; similarly, the waves from the third quarter of the slit cancel those from the last quarter of the slit. Thus the cancellation is complete and the resultant wave has again *zero* intensity.

► *Summary*

The preceding comments indicate that the intensity I observed at various angles θ varies in the manner indicated in Fig. D-3.

Thus the intensity is maximum, as a result of complete constructive interference, along the central line where $\theta = 0$. The intensity is zero as a result of complete destructive interference at angles θ such that $a = 2d_c$, $a = 4d_c$, $a = 6d_c$, *ldots* where waves from different parts of the slit cancel each other in pairs. Finally, between these angles the cancellation of the waves is not complete so that the intensity exhibits secondary maxima which decrease with increasing angle θ .

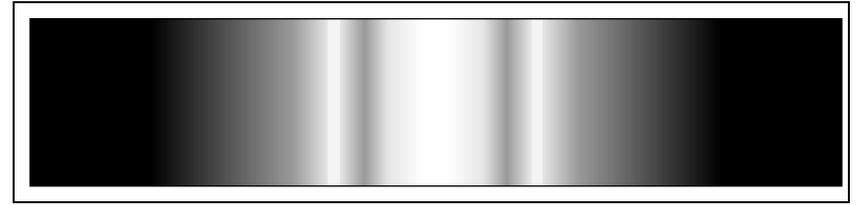


Fig. D-4: Photograph showing the observed intensity of light which has passed through a narrow slit.

DISCUSSION

► *Brightness and shadow*

Figure D-3 shows in detail how a wave spreads out behind a slit in an opaque screen. At angles close to the central line where $\theta = 0$, the intensity is large (i.e., there is a bright region). At angles far from the central line the intensity is negligible (i.e., there is a region of shadow). In the transition region between brightness and shadow the intensity decreases while fluctuating between zero and a series of small secondary maxima. Figure D-4 shows a photograph of the actual light observed on a photographic plate placed behind an opaque screen when light is incident on a narrow slit in this screen.

► *Width of diffraction pattern*

The width of the entire diffraction pattern illustrated in Fig. D-3 depends on the ratio λ/a of the wavelength λ compared to the width a of the slit. In particular, the central region of maximum intensity is an angular region (or “beam”) between the directions specified by the smallest angle θ_b at which the intensity is zero. (As indicated in Fig. D-5, the angle θ_b is thus half the angular width of the entire central beam.) But the angle θ_b corresponds to the situation of Fig. D-2b where the cancellation distance $d_c = a/2$ so that light from the first half of the slit just cancels light from the second half of the slit. Thus Eq. (D-1) implies that

$$\frac{a}{2} \sin \theta_b = \frac{\lambda}{2}$$

or

$$\boxed{\sin \theta_b = \frac{\lambda}{a}} \quad (\text{D-2})$$

The angular width of the central beam increases thus with the relative size of the wavelength λ compared to the width a of the slit. For example, if the width a of a slit is larger, the ratio λ/a is smaller so that the

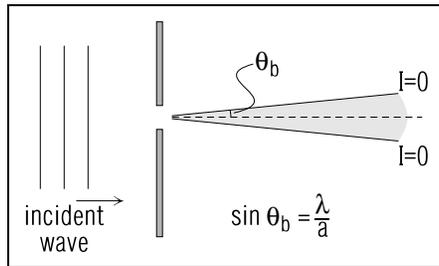


Fig. D-5: Central beam emerging from a narrow slit when a wave is incident perpendicularly upon this slit.

angular width of the central beam is correspondingly also smaller. (See Fig. D-6.) Interference effects are responsible for the secondary intensity maxima surrounding the central beam.

► $\lambda \gg a$

If the width of the slit is so small that $\lambda \gg a$, the angular width of the central beam is so large that it fills the entire region behind the slit. Thus the wave spreads out almost uniformly in all directions behind the slit and no destructive interference effects are observed. (This conclusion agrees with our previous comments in Sec. C about a wave passing through a hole of size small compared to the wavelength.)

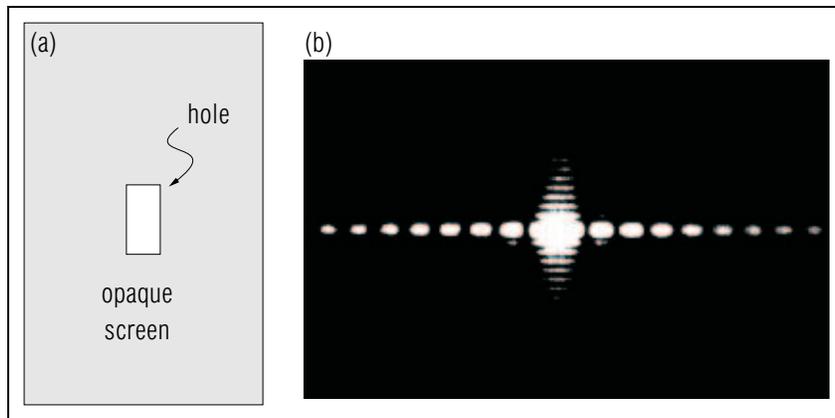


Fig. D-6: (a) Rectangular hole in an opaque screen. (b) Photograph showing the observed intensity of light passing through this hole. Note that the central beam of maximum intensity has a smaller width along the direction where the side of the hole is smaller.

► $\lambda \ll a$

If the width of the slit is so large that $\lambda \ll a$, the angular width of the central beam is very small. Hence the intensity behind the slit becomes negligible beyond a small angular range away from the central line. Thus the wave appears to travel through the slit nearly along the central straight line determined by the direction of the incident wave, while a shadow region of negligible intensity is produced along all directions appreciably different from this central direction.

► *Circular hole*

The diffraction of waves passing through a circular hole of diameter a is qualitatively similar to that discussed for the slit, with circularly symmetric intensity maxima and minima at angles approximately the same as those indicated in Fig. D-3. In particular, the half-angle θ_b of the central beam is still approximately given by Eq. (D-2) so that $\sin \theta_b \approx \lambda/a$.*

* A more precise calculation shows that $\sin \theta_b = 1.22(\lambda/a)$

Diffraction From a Hole Or Source (Cap. 2)

D-1 *Light passing through a slit:* Yellow light, having a wavelength of 6.0×10^{-7} m, passes through a narrow slit having a width of 6.0×10^{-6} m. (a) What is the value of the half-angle θ_b specifying the angular width of the light emerging through this slit? (b) If the light passing through the slit were blue light, which has a wavelength smaller than yellow light, would the angular width of the emerging central beam be larger or smaller? (c) If the yellow light passed through a slit having a width smaller than 6×10^{-6} m, would the angular width of the emerging central beam be larger or smaller? (d) For this yellow light, what should be the slit width so that the half-angle of the central beam is 90° (i.e., so that complete destructive interference, or shadow, is no longer observed behind the hole)? (*Answer: 12*)

D-2 *Spreading of laser light:* The coherent light produced by a laser travels parallel to the laser tube and then emerges through a hole, 2 mm in diameter, at the end of the tube. If the wavelength of this light is 6×10^{-7} m, estimate the radius of the beam of light produced by this laser at a distance of 1 km from the laser. (*Answer: 10*) (*Suggestion: [s-7]*)

D-3 *Measuring the width of a narrow slit:* To measure the width of a narrow slit without the use of a microscope, an experimenter

illuminates this slit with red light having a wavelength of 6.6×10^{-7} m. He then observes, on a piece of cardboard at a distance of 0.80 m from the slit, a diffraction pattern where there is a distance of 1.2 cm between the first intensity minimum to the left of the central maximum and the first intensity minimum to the right of the central maximum. What then is the width of the slit? (*Answer: 14*) (*Suggestion: [s-4]*)

More practice for this Capability: [p-4], [p-5]

SECT.

E RAY APPROXIMATION FOR WAVES

► *Decomposition of Waves*

To discuss the motion of a wave through space in complex situations (e.g., in the presence of various obstacles), it is easiest to regard the wave as decomposed into a sum of simple waves each of which can be easily studied. To achieve such a decomposition, imagine that an opaque screen, containing somewhere a small hole, is located along a wavefront of the wave (see Fig. E-1.) Then the wave beyond the screen can be regarded as the sum of waves emanating from the holes of all such screens, each with a hole at a different small part of the wavefront.

For a precise description, every such hole in an imaginary screen should be very small, i.e., small compared to the wavelength λ of the wave. Then the resultant wave can be regarded as the sum of the spherical waves emanating from all such holes. This point of view is, of course, just Huygens' principle which we exploited in the preceding sections to discuss the diffraction of a wave passing through slits or holes of any size.

► *Rays*

On the other hand, consider a situation where all objects encountered by a wave are much *larger* than the wavelength λ . [Such a situation is particularly common in the case of light since the wavelength of light (about 5×10^{-7} m) is much smaller than the size of most objects of usual interest.] In this case one may be satisfied with the following simplified description which, although approximate, may be sufficiently accurate for many purposes: The hole in every one of our imaginary screens may be chosen to be very small compared to the size of all objects, but nevertheless very *large* compared to the wavelength λ . Then the wave emanating from any such hole is (according to the discussion of Sec. D) a beam of negligibly small angular width. Such a beam is called a "ray" in accordance with this definition:

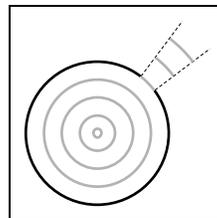


Fig. E-1: Wave emanating from a small hole in an imaginary opaque screen coinciding with the wavefront of an original wave. (In this diagram, the hole is large compared to the wavelength.)

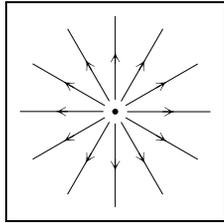


Fig. E-2: Rays emanating from a small source.

Def.

Ray: A wave in the form of a beam of negligible angular width, traveling along a specified direction.*

(E-1)

* If the hole in the imaginary screen has a diameter a , we know from Eq. (D-2) that the angular width of the beam is approximately λ/a (measured in radians). The angular width of the beam is thus negligible if the diameter a of the hole, although small compared to all relevant objects, can be chosen so large that the ratio λ/a is negligibly small.

By considering all the waves emanating from different holes in all such imaginary screens, the original wave can then be regarded as a collection of rays. (See Fig. E-2.)

► *Simplicity of rays*

Such a description in terms of rays has the virtue of great simplicity. Indeed, in a homogeneous medium each ray travels simply along a straight path while its angular width remains negligible. (Thus the path of such a ray resembles the path of particles moving through space without interaction.) By tracing the paths of individual rays as they encounter various objects, it is then quite easy to discuss the motion of waves through complex surroundings.

For example, since the wavelength of light is much smaller than the size of common objects, the motion of light waves encountering lenses or mirrors can be discussed to good approximation by regarding light as a collection of light rays. As we shall see in Unit 436, one can thus easily analyze and design important optical instruments (such as microscopes or telescopes) by simply tracing the paths of various light rays through the lenses of these instruments.

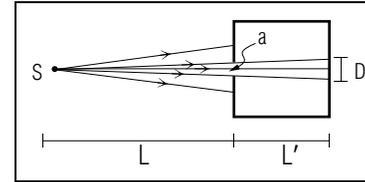


Fig. E-3.

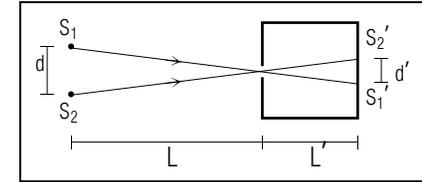


Fig. E-4.

Knowing About Rays

E-1 *Light rays passing through a hole:* Fig. E-3 shows a box, made of opaque material, which has a length L' . In the front side of the box there is a circular hole of diameter a which is small compared to L' , although much larger than the wavelength of light. Light emanates from a small source S at a large distance L in front of the hole in the box. Since the wavelength is very small compared to all relevant dimensions, the light emanating from the source can then be considered to consist of a bundle of light rays traveling outward from the source S along straight lines, as indicated in Fig. E-3. The light rays incident on the hole in the box then enter the box, but the rays incident on the opaque parts of the box do not pass through to enter the box. (a) On the basis of purely geometric arguments, what then is the diameter D of the illuminated disk formed on the back wall of the box by the light coming from S and passing through the hole? (b) What is the approximate value of D if the length L' of the box is very much smaller than the distance L of the source S from the box? (*Answer: 11*) (*Suggestion: [s-8]*)

E-2 *Pinhole camera:* Suppose that the hole in the preceding problem is very small (i.e., of pinhole size), although it is still much larger than the wavelength of light. Then the angle subtended at the source by the bundle of rays entering the hole is so small that the whole bundle of rays can be approximated by a single ray. Figure E-4 shows then two such rays (i.e., two such very narrow bundles of rays) emanating from two sources S_1 and S_2 separated by a distance d and located at a distance L in front of the hole. Light from each of these sources then produces on the back wall of the box a very small illuminated disk corresponding to this source. (Each such disk is called an “image” of the corresponding source.) (a) What then is the separation d' between the images S_1' and S_2' corresponding to the two sources? (b) As discussed in the preceding problem, if $L \gg L'$, the size of the illuminated disk produced by each source is approximately equal to the diameter a of the hole. If the disks

corresponding to the two sources are not to overlap (i.e., if the images produced by the two sources are to be distinct), what must then be the minimum separation between the sources? [If the hole is small enough, the images from different sources are distinct and the box acts as a “pinhole camera”. For example, light emanating from different points of an object then forms corresponding images of these points on the back wall of the box. Hence a photographic film placed on this back wall produces a good photographic picture of the original picture.] (*Answer: 15*) (*Suggestion: [s-6]*)

SECT.

F SUMMARY

DEFINITIONS

resolution; Def. (A-1)

diffraction; Def. (C-1)

ray; Def. (E-1)

IMPORTANT RESULTS

Minimum resolvable distance: Eq. (A-3), Eq. (A-5)

$d_{\min} \approx \lambda/(2 \sin \theta)$, where θ is half-angle subtended at sources by a detector aperture

if $L \gg a$, $d_{\min} \approx \lambda(L/a) = L(\lambda/a)$

Diffraction by a hole or slit: Sects. C and D, Eq. (D-2)

Most intensity concentrated in central beam of half-angle θ_b such that $\sin \theta_b \approx \lambda/a$.

USEFUL KNOWLEDGE

X-ray methods for determining the structure of matter (Sec. B)

Huygens' principle (Sec. C, Sec. E)

Interference effects responsible for brightness and shadow (Sec. B)

Validity of ray approximation (Sec. E)

NEW CAPABILITIES

- (1) Relate the approximate minimum resolvable distance between two points to the wavelength of the waves used for the observations, to the detector size, and to the detector position relative to the points. (Sec. A; [p-1], [p-2], [p-3])
- (2) For waves emanating in phase from a hole or source of any size, (a) use interference arguments to describe qualitatively the resultant intensity at large distances from the hole or source; (b) relate the angular width of the emerging central beam to the wavelength and the size of the hole (exactly for a long slit, and approximately for a circular hole.) (Sects. C and D, [p-4], [p-5])

F-1 *Sound-wave radar used by bats (Cap. 1):* Bats have poor eyesight, but navigate and detect their insect prey by emitting sound waves with frequencies as large as 1.2×10^5 Hz. The detection of these sound waves, reflected off obstacles and insects, then provides a bat with most of its information about its environment. (a) What is the shortest wavelength of the sound waves emitted by a bat if the speed of sound in air is 340 m/s? (b) What is approximately the shortest distance (in millimeters) between two points which the bat can recognize as distinct as a result of such sound waves reflected from these points? (c) Is this distance small enough so that the bat can identify individual insects? (*Answer: 13*)

F-2 *Sound emitted by a loudspeaker horn (Cap. 2):* The horn of a loudspeaker has a rectangular aperture which is 3 feet high and 1 foot wide. Is the beam of sound emerging from this horn broader in the vertical or in the horizontal plane? (*Answer: 17*)

SECT.

G PROBLEMS

G-1 *Useful magnification of a microscope:* An optical microscope is designed so that light, emanating from two points separated by a distance d in a sample, gives rise to two corresponding image points separated by a much larger distance d' which can be readily observed with the eye. The ratio d'/d is then called the “magnification” provided by the microscope. The “*useful* magnification” of the microscope is then the particular magnification such that two points, separated by the minimum possible resolvable distance, give rise to image points separated by 0.1 mm (the minimum distance between points which can comfortably be recognized as distinct by the naked eye). (a) What then is approximately the useful magnification of a microscope for a sample in air, if one uses violet light of the shortest possible wavelength (4×10^{-7} m) visible to the eye and if one uses an objective lens under optimum conditions (i.e., placed very close to the sample)? (b) Could one obtain more detailed information about a sample if one observed this sample with a microscope having a magnification larger than its useful magnification? (*Answer: 19*)

G-2 *Location of interference minima in a diffraction pattern:* Consider the diffraction pattern produced when waves, of wavelength λ , pass through a long slit of width a . At large distances behind the slit, the intensity is then zero at a series of angles $\theta_1, \theta_2, \theta_3, \dots$ measured from the central direction along which the intensity is maximum. What are the magnitudes of these angles, expressed in terms of λ and a ? (*Answer: 16*) (*Suggestion: [s-10]*)

G-3 *Intensities of subsidiary maxima in a diffraction pattern:* In the diffraction pattern described in the preceding problem and illustrated in Fig. D-3, there are a series of subsidiary maxima (other than the central maximum) located between the angles where the intensity is zero. Estimate the ratio I_2/I_1 of the intensity I_2 of the second such subsidiary maximum compared to the intensity I_1 of the first such subsidiary maximum. (*Answer: 20*) (*Suggestion: [s-9]*)

G-4 *Diffraction with different wavelengths:* When a long slit is illuminated with light of various wavelengths, it is observed that the first intensity minimum produced by light with a wavelength λ_1 coincides with the second intensity minimum produced by light with a wavelength λ_2 . What then is the relation between the wavelengths λ_1 and λ_2 ? Use

the results of problem G-2. (*Answer: 22*)

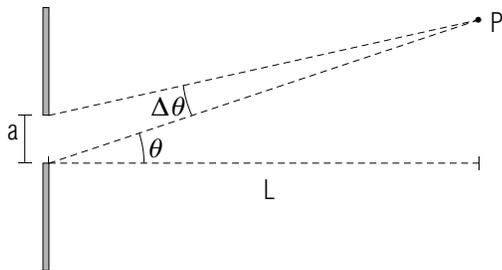
G-5 *Optimum hole size for a pinhole camera:* To obtain a clear unblurred image with the pinhole camera described in P14, the diameter of the illuminated disk produced by light from any small source should be as small as possible. As discussed in problem E-1, this can be achieved by making the pinhole smaller, so that the bundle of rays passing through the hole is narrower. Thus the diameter of the illuminated disk is approximately equal to the pinhole diameter a (if the distance L' of the back of the camera from the pinhole is very small compared to the distance L of the pinhole from the source). (a) The preceding comments, based on the approximation of light rays, neglect diffraction effects. Suppose, however, that the diameter of the pinhole is so small that diffraction effects are important. If the wavelength of the light is λ what then is the diameter D_d of the illuminated disk produced on the back of the camera because of the diffraction of light passing through the pinhole? (b) If the diameter of the pinhole is decreased, do diffraction effects lead to a decrease or increase of the diameter of this disk? (c) As the pinhole diameter is decreased from a size much larger to one smaller than λ the approximate diameter of the illuminated disk decreases then until this diameter becomes equal to the disk diameter D_d caused by diffraction. What then is approximately the pinhole diameter which leads to the minimum size of the illuminated disk? (*Answer: 18*)

Note: Tutorial section G contains additional problems.

TUTORIAL FOR G

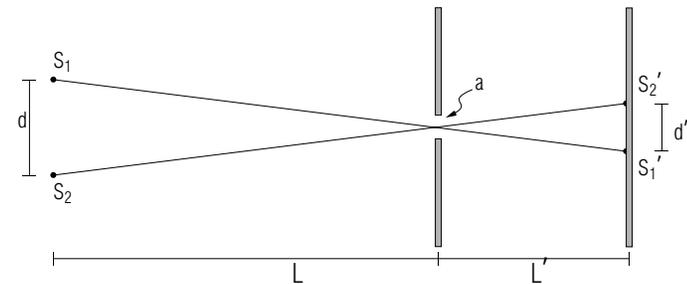
Additional Problems

g-1 * *VALIDITY CONDITION FOR DIFFRACTION AT LARGE DISTANCES*: In Sec. D we considered waves, of wavelength θ , passing through a long slit of width a . We then found the intensity at a distance L much larger than the width of the slit. Under these conditions, lines drawn from any part of the slit to a distance point P nearly parallel so that one can neglect any difference in the angle θ (relative to the central direction) of a line drawn to P from one edge of the slit or from the other edge of the slit.



(a) In the diagram, what is approximately the difference $\Delta\theta$ in the angles between these two lines? Express your answer in terms of a and L . (b) To make our calculation valid, this angular distance $\Delta\theta$ should be much smaller than the angle θ_1 determining the position of the first interference minimum of the diffraction pattern. What is this angle θ_1 (in radians), assuming the λ is much smaller than a ? (c) If $\Delta\theta$ is to be much smaller than θ_1 , what is the quantitative condition specifying how large L must be so that the arguments used in Sec. D of the text are valid? (*Answer: 56*)

g-2 *RELATION BETWEEN DIFFRACTION AND RESOLUTION*: The diagram shows two distant sources S_1 and S_2 , separated by a distance d , at a large distance L from an opaque screen containing a hole of diameter a much smaller than L .



The light from the sources then passes through a hole and arrives on an observation screen at a distance L' behind the hole. (a) What is the separation d' between the centers of the diffraction patterns formed on the observation screen by the light from the two sources? (b) For any one of these diffraction patterns, what is the approximate distance R between the central maximum of the diffraction pattern and the first interference minimum of zero intensity? Assume the the wavelength λ is small compared to a . (c) To recognize that the observed diffraction patterns are due to two distinct sources, the diffraction patterns must be sufficiently distinct, i.e., the separation d' between the centers of these patterns must be larger than the width R of the central maximum of each pattern. To satisfy this condition, what must be the minimum separation between the two sources? (d) Compare this condition with the minimum resolvable distance between the sources, as predicted on the basis of the arguments presented in Sec. A of the text. (*Answer: 59*)

PRACTICE PROBLEMS

p-1 *RESOLUTION AND WAVELENGTH (CAP. 1): Minimum resolvable distance for the eye:* The pupil of the human eye, when dilated, is about 0.5 cm in diameter. Furthermore, the eye can only accommodate so as to observe objects no closer than about 25 cm from the eye. (This is the usual at which people with normal eyesight read a printed page.) On the basis of this information, estimate the minimum distance (in millimeters) by which two points must be separated so that they can barely be recognized as distinct by the unaided eye under optimal conditions? Assume that the wavelength of light is 5.5×10^{-7} m and neglect effects due to the structure of the retina. (*Answer: 53*) (*Suggestion: See [s-2] and text problems A-4 and A-5.*)

p-2 *RESOLUTION AND WAVELENGTH (CAP. 1): Resolving the headlights of an approaching car:* The two headlights of an automobile are 4 feet (i.e., 1.2 m) apart. Estimate the distance to which such an automobile much approach an observer before he can be sure that he sees two headlights rather than one (and can thus distinguish the automobile from a motorcycle.) Assume that the pupil of the eye has a diameter of 5 mm and that the wavelength of light is 5.5×10^{-7} m. Neglect effects due to the structure of the retina and express your answer in terms of miles. (1 mile = 1.6 km.) (*Answer: 55*) (*Suggestion: Review text problems A-4 and A-5.*)

p-3 *RESOLUTION AND WAVELENGTH (CAP. 1): Distinguishing holes in acoustic tile:* The walls of a large room are covered with acoustic tile containing many small holes whose centers are separated from each other by a distance of 5.0 mm. Approximately how far can a person stand from the wall of such a room and still distinguish the individual holes? Assume that the pupil of the eye has a diameter of 5 mm and that the wavelength of light is 5.5×10^{-7} m. Neglect the structure of the retina. (*Answer: 51*) (*Suggestion: Review text problems A-4 and A-5.*)

p-4 *DIFFRACTION FROM A HOLE OR SOURCE (CAP. 2): Sound wave passing through a window:* Sound waves corresponding to the tone of “high C” (the highest tone produced by a soprano voice) have a frequency of 1.05×10^3 Hz. (a) What is the wavelength of such sound waves in air if the speed of sound is 340 m/s? (b) Suppose that such sound waves incident perpendicularly upon a wall containing a long window 1.0 m wide. Is the

intensity observed at appreciable distances behind the window fairly much the same in all directions, or are there certain directions along which the intensity detected by a microphone is zero? If the latter is true, what is the magnitude of the smallest angle from the central direction (perpendicular to the window) at which the detected intensity is zero? (*Answer: 57*) (*Suggestion: Review text problem D-1.*)

p-5 *DIFFRACTION FROM A HOLE OF SOURCE (CAP. 2): Sound emitted by a loudspeaker:* Text problem C-3 discussed the nature of the sound emanating from a loudspeaker having a moving diaphragm 10 cm in diameter. Suppose that this diaphragm vibrates at a frequency of 2×10^4 Hz, thus emitting sound of the highest audible pitch. If the speed of sound in air is 340 m/s, find approximately the half-angle of the central beam along which the sound is emitted by the loudspeaker. (*Answer: 52*) (*Suggestion: Review text problems C-3 and D-1.*)

SUGGESTIONS

s-1 (*Text problem B-1*): (a) If the separation between the atoms is d , relate λ to the angle θ between the central direction and the position of the first interference maximum. (If necessary, review Unit 431.) (b) Since $s \ll L$, the angle θ is very small. To very good approximation, what then is the relation between $\sin \theta$ and the distances s and L ? (c) According to the information provided, the ratio A_1/A_2 of the amplitudes of the waves, arriving from the two atoms at the photographic plate, is equal to either 1, 2, 3, or 4. Thus $A_1 = A_2$, $A_1 = 2A_2$, $A_1 = 3A_2$, or $A_1 = 4A_2$. What then would be the ratio I_{\max}/I_{\min} in each of these cases? (*Answer: 54*)

s-2 (*Practice problem [p-1]*): We know from Sec. A of the text that optimum conditions, for resolving points the smallest possible distance apart, are realized if the size of the pupil is as large as possible (i.e., about 0.5 cm) and if the eye is as close as possible to the points (i.e., about 25 cm.) Note that this distance is much larger than the pupil of the eye so that the relation (A-5) of the text is applicable.

s-3 (*Text problem A-2*): The answers to all questions are implied by the relation (A-3) of the text. *Part e*: What is the radius of the detector? In Fig. A-1 of the text, what is then the distance from the sources to the edge of the detector? What then is the angle θ subtended at the sources by the radius of the detector?

s-4 (*Text problem D-3*): On the basis of the given information, what is $\sin \theta_b$, where θ_b is the half-angle of the central beam of light emerging from the slit? (Note that this angle is very small so that $\sin \theta_b \approx \tan \theta_b \approx \theta_b$, measured in radians.) What then must be the width of the slit?

s-5 (*Text problem C-2*): *Part e*: Remember the $I = \gamma A^2$ and that $I_p = \gamma A_p^2$. If A and A_p differ only in sign, what is the relation between I and I_p ?

s-6 (*Text problem E-2*): Use similar triangles, comparing the triangle of height L with that of height L' .

s-7 (*Text problem D-2*): What is approximately the half-angle θ_b of the central beam emerging from the laser? What then is the radius of the beam at a distance of $1 \text{ km} = 10^3 \text{ m}$? Note that, since θ_b is small, $\sin \theta_b \approx \tan \theta_b \approx \theta_b$ (measured in radians).

s-8 (*Text problem E-1*): Consider the bundle of rays which leave S and then enter the box after passing through the hole of diameter a . To compare D with a , it is then only necessary to compare these two similar triangles: (1) The triangle, of height L , formed by the outermost rays in the bundle and the diameter a of the hole. (2) The triangle, of height $L + L'$, formed by these outermost rays and the diameter D of the illuminated disk.

s-9 (*Text problem G-3*): The first subsidiary maximum occurs when the slit width is approximately 3 times as large as the cancellation distance. Then the waves from the two strips in the slit cancel each other, and only the waves from the remaining one third of the slit contribute to the resultant intensity. Similarly, the second subsidiary maximum occurs when the slit width is approximately 5 times as large as the cancellation distance. Then the waves from four strips in the slit cancel each other in pairs, and only the waves from the remaining one fifth of the slit contribute to the resultant intensity. (a) What then is the ratio A_2/A_1 of the resultant amplitude A_2 due to one fifth of the slit compared to the resultant amplitude A_1 due to one third of the slit? (b) What then is the ratio I_2/I_1 of the resultant intensity due to one fifth of the slit, compared to the resultant intensity due to one third of the slit? (*Answer: 58*)

s-10 (*Text problem G-2*): Zero intensity results at all those angles at which the slit width a is equal to an even number $2n$ (where $n = 1, 2, 3, \dots$) of cancellation distances d_c , since waves coming from different parts of the slit then interfere so as to cancel each other in pairs. But the cancellation distance is related to the angle so that $d_c \sin \theta = \lambda/2$. These comments suffice to determine the angles at which the resultant intensity is zero.

ANSWERS TO PROBLEMS

1. a. smaller
b. smaller
c. $d'_{\min} = 0.6d_{\min}$
2. a. larger
b. closer
c. blue-green
d. 0.71
e. 3.1×10^{-7} m
3. a. 5.2 m, 0.16 m
b. all directions
c. predominantly along incident direction
4. 3.3 cm
5. a. 2×10^{-6} m
b. 7.5×10^8 Hz
6. a. 4×10^4 m
b. 40 m
7. a. 3.4 m, uniformly in all directions
b. 0.017 m, predominantly in forward direction
8. a. $\lambda(L/s)$
b. 3
9. a. $A_S + A_o + A_p$
b. 0
c. $A_S + A_o + A_p = 0$
d. $A = -A_p$
e. $I = I_p$
f. same
10. 0.3 m
11. a. $D = a[(L + L')/L] = a(l + L'/L)$
b. $D \approx a$

12. a. 5.7°
b. smaller
c. larger
d. width = $\lambda = 6.0 \times 10^{-7}$ m
13. a. 2.8 mm
b. 1.4 mm
c. yes
14. 0.088 mm (NOT 0.044 mm)
15. a. $d' = d(L'/L)$
b. $a(L/L')$
16. $\theta_n = \sin^{-1}(n\lambda/a)$ where $n = 1, 2, 3, \dots$
17. horizontal
18. a. $D_d = 2(\lambda/a)L'$
b. increase
c. $a = \sqrt{2\lambda L'}$
19. a. 5×10^2
b. no
20. 0.36
21. BLK
22. $\lambda_1 = 2\lambda_2$
23. BLK
24. BLK
51. 45 m
52. 10°
53. 0.03 mm
54. a. $d \sin \theta = \lambda$
b. $\sin \theta \approx s/L$
c. $\infty, 9, 4, 25/9$

55. 7 mile
56. a. $\Delta\theta \approx a/L$
 b. $\theta_1 \approx \lambda/a$
 c. $L \gg a^2/\lambda$
57. a. 0.325 m
 b. Intensity is not uniform; 19°
58. a. $(1/5)/(1/3) = 3/5$
 b. $9/25 = 0.36$
59. a. $d' = d(L'/L)$
 b. $R = (\lambda/a)L'$
 c. $d_{\min} = L(\lambda/a)$
 d. same
60. BLK

MODEL EXAM

- Resolution when viewing the earth from an airplane.** Imagine you are flying above the earth 8×10^3 meter from the ground. The pupil of your eye has a diameter of 5×10^{-3} meter, and the wavelength of the light is approximately 6×10^{-7} meter.
 If diffraction at the pupil of the eye is the only factor limiting resolution, what is the minimum resolvable distance on the earth's surface under the conditions described?
- Passage of ultrasonic waves through a small hole.** Dolphins emit ultrasonic waves with frequencies as high as 2.5×10^5 hertz. Consider what happens if waves of this frequency are perpendicularly incident on a hole of diameter 1.6×10^{-2} meter. The speed of sound in seawater is 1.5×10^3 meter/second.
 - Give a brief, qualitative description of the intensity of sound at points far from the hole.
 - Approximately what is the half-angle of the sound wave after it has passed through the hole?
- Diffraction pattern of a heated crystal.** A crystal is arranged so that X-rays pass through it and form a characteristic diffraction pattern on a photographic plate some distance beyond. If the crystal is then heated, so that it expands (and the distances between atoms increase) are the spots on the resulting photograph closer together or further apart than those on the original photograph?

Brief Answers:

- 0.96 meter (1 meter OK)
- The waves spread out some, rather than just going straight through.
 - 22°
- closer together

