REFERENCES

1 Wilska, A., (1953). *A New Method of Light Microscopy*. *Nature*, **171** 353.
2 Wilska, A., (1954). *Observations with the Anoptral Microscope*. *Mikroskopie*, **9** 1.
3 Françon, M., (1961). *Progress in Microscopy*, Pergamon Press, London, p.79.
4 Pluta, M., (1964). *Simple Method of Production of Phase Rings*, *Pomiary Automatyka Kontrola* **10** 226, in Polish.
5 Pluta, M., (1965). *Mikroskopia fazowo-kontrastowa i interferencyjna*, P. W. N. Warszawa, pp.116-129.
6 Pluta, M., (1958). *Microscope with Anoptral Contrast*, *Kosmos Ser. A, Biologia*, **7** 587, in Polish.
7 Pluta, M., (1963). *Studies of Microscopic Lenses by Means of a Polarizing interferometer with the Wollaston Prism*, *Pomiary Automatyka Kontrola*, **9** 292, in Polish.
8 Pluta, M., (1968). *Stereoscopic Phase Contrast Microscope*, *The Microscope*, **16** 32.

What Resolving Power Formula Do You Use?*

H. W. ZIELER
Atlantex and Zieler Corporation,
55 Eastern Avenue, P.O. Box 575,
Dedham, Mass. 02026, U.S.A.

The literature on microscopy includes publications in which resolving power and other physical optical aspects of image formation by the microscope are presented in great detail with much theory and various equations. These sources of information generally require so much basic knowledge in mathematics and optics that the average practising microscopist has neither the background nor the time to absorb this information and apply it to his practical problems.

There are other publications concerning the use of the microscope, usually in specific fields of science: biology, chemistry, mineralogy or metallography, in which the physical optical aspects of image formation are often presented too briefly. Statements to the effect that the resolving power of the objective depends on the wavelength of the light and the numerical aperture of the objective are supplemented by one or more of the following equations:

$$D = \frac{\lambda}{2NA} \quad \ldots (1)$$

$$D = \frac{0{\cdot}61\lambda}{NA} \quad \ldots (2)$$

$$\frac{\lambda}{NA} \geqslant D \geqslant \frac{\lambda}{2NA} \quad \ldots (3)$$

$$D = \frac{1{\cdot}22C\lambda}{NA} \quad \ldots (4)$$

*Presented at INTER/MICRO-69, 9-11 September, 1969, Imperial College, London, England.

In these equations, the factor D is the minimum distance between two points barely resolved by an objective of given numerical aperture (NA) when the image is formed by light of wavelength λ.

These equations, if accepted dogmatically without explanations as to their restricted validity, are contradictory and misleading. This fact, however, often escapes the attention of the microscopist because, in any given publication, only one of these equations (usually the first) is quoted.

Taken at face value, equation (1) states that the resolving power of an objective of a given NA, forming an image with light of wavelength λ, is *constant* and depends only on these *two* parameters regardless of the optical conditions prevailing in the formation of the image.

Equation (2) is essentially the same, though the minimum separation of two points just resolved, D is about 20% greater. For instance, if an objective of NA = 1·00 forms an image with light of wavelength 500 nm (nanometers), D according to equation (1) is: 500/2 = 250 nm but according to equation (2) it is: 0·61 (500)/1 = 305 nm.

Equation (3) states that the resolving power is *not* constant for a given wavelength and NA but *varies* from its highest value which is twice that of its lowest value. Accompanying remarks often state that there are three factors which influence the resolving power, the added one the NA of the illumination. Equation (3) is also quoted in the following form:

$$D = \frac{\lambda}{NA_{ill.} + NA_{obj.}} \qquad \ldots (3a)$$

For axial illumination, when the well-centered aperture iris of the condenser is closed as far as possible, the resolving power is at its worst ($NA_{ill.} \simeq 0$). When this iris is fully open and the NA of the illumination is equal to that of the objective, the resolving power is at its optimum.

Equation (4) contains a new factor C. According to the usual explanatory remarks, this factor varies from about 0·4 to about 1·0. This equation also indicates that the resolving power is *variable* for a given NA and λ. The lowest and highest values are about the same as those of equation (3). The factor, C, however, is not to be interpreted as the NA of the illumination. Its correct interpretation involves more than one optical factor and we will have to explain this later.

It has been my experience, that many microscopists accept equation (1) dogmatically and apply it to all conditions of image formation, although they must know from practical experience that the resolving power varies with the NA of the illuminating beam. Even the rankest beginner in microscopy can demonstrate how the resolving power varies by observing a suitable object with a selected objective; at first, with the aperture iris diaphragm of the condenser closed as far as possible and then under conditions of increasingly opened iris diaphragm with resulting continuously improving resolving power.

The apparently contradictory contents of these equations can be explained by considering the optical phenomena produced by various types of small objects or structures in the back focal plane of the objective and in the image plane under a variety of image-forming conditions.

By "optical phenomena", or optical manifestations, I mean the production of interference maxima and minima in the back focal plane of the objective and the production of diffraction discs in the image plane. Starting at the very beginning, it should be mentioned that no image can be formed without light. Unless the object is self-luminous, light must be conveyed to the object through an illumination system. The simplest optical conditions for image formation require a single, *self-luminous* point.

You may now skip to the conclusions, if you wish, to avoid the theoretical discussion. I have tried, however, to keep this section to a minimum of essential material, understandably written. I recommend reading this for a better understanding of which resolving power equation to use and why.

Physical Optical Aspects of the Propagation of Light

Light emitted from a single, self-luminous point travels in all directions with equal speed (in an isotropic medium) somewhat similar to waves spreading on the surface of water from a center, where a stone has been thrown into the water. Each circle, which can be seen on the surface of the water, is a "wavefront"—a front of equal phase of a wave. Light waves radiate in three dimensions rather than two and the "wave surfaces" are spheres of rapidly increasing diameters.

It is a characteristic property of light waves (as well as water waves) that each single point of a wavefront or surface becomes the center of a new wavefront if it is isolated (Figure 1).

Another characteristic property of ordinary light waves is, that the vibration directions vary throughout 360° in a plane normal to the direction of propagation (Figure 2).

The plane of vibration directions for a single bundle of light rays from a single self-luminous point creates a whole new set of radiating propagation directions each with its own set of perpendicular vibration directions when incident upon a small opening in a metal plate. The same bundle of light rays striking two separate holes will generate, by diffraction, two new families of propagation directions with their perpendicular planes of vibration directions (Figure 3).

As the lightpaths of these two new wave motions intersect, the waves interfere with each other constructively, or destructively, depending on the phase difference at that point. At any point, equidistant from the two holes, that a crest from one wave interferes with a crest from the other wave, an interference *maximum* is produced. The same occurs at any point for which the difference in the distances from the two holes is one full wavelength or a full multiple thereof. At any point, for which the difference between the distances from the two holes is one-half wavelength or an odd multiple thereof, there is an interference *minimum* because a crest of one wave interferes with a trough of the other wave.

Formation of the Image of a Single, Self-Luminous Point

As the spherical wave surfaces expand away from the self-luminous point, coherent portions of them enter the objective. Because of the higher refractive indices of the lens components, the speed of the lightwaves is reduced as they pass through the components. Further-

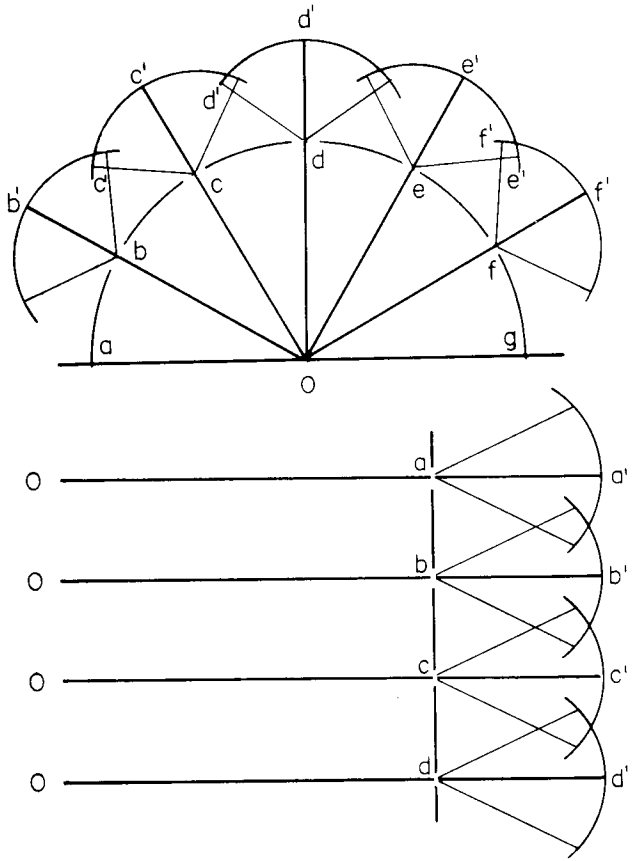

Figure 1. Original and diffracted wave motions.

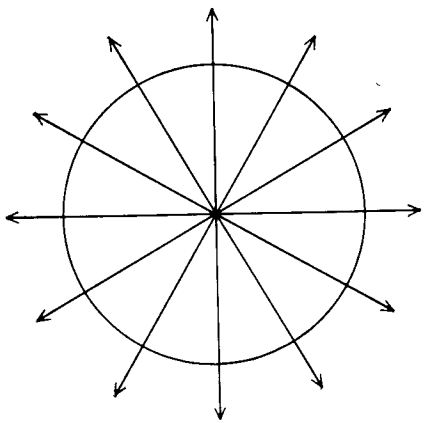

Figure 2. Azimuths of vibration directions perpendicular to the propagation direction.

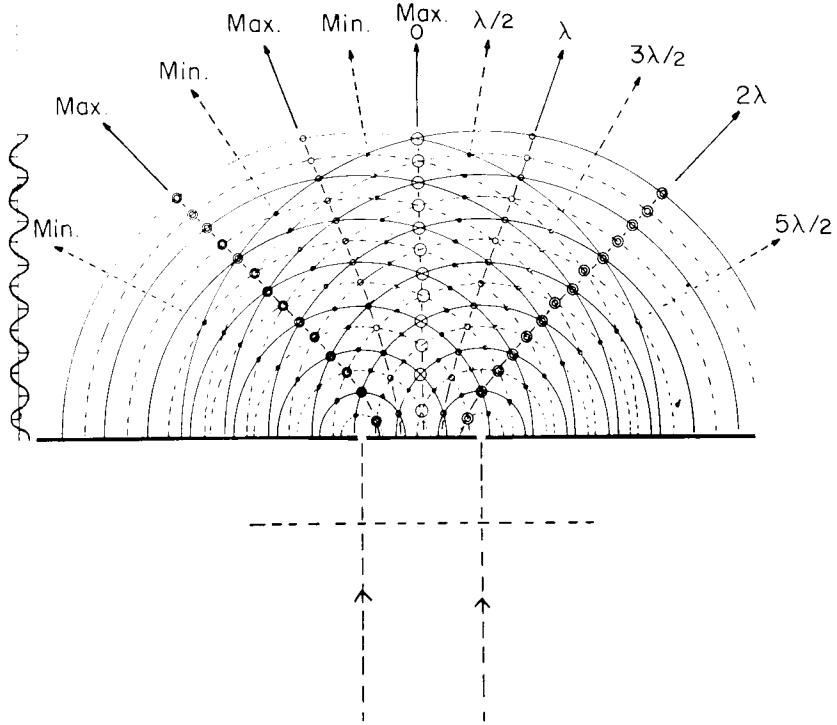

Figure 3. Light diffracted from two openings in otherwise opaque specimen.

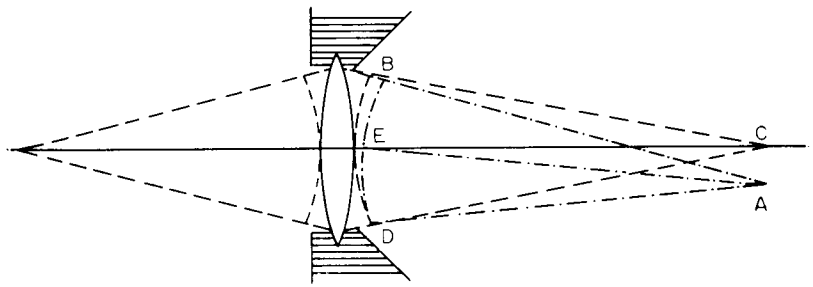

Figure 4. Image formation of single self-luminous point.

more, because these components have greater thicknesses in their central regions, those portions of the wave surfaces which pass through these regions undergo a reduction of speed for a greater optical pathlength than those passing through the outer regions. Thus, the shapes of the wave surfaces are changed. If the objective is of theoretical perfection—*and only in that case*—the emerging wave surfaces are again spheres which, however, are curved in opposite directions. They travel towards their common new center, the image point (Figure 4).

If the object point is on the optical axis of the objective, the image point at C to which the spherical wave surfaces converge, is also on the optical axis, equidistant from diametrically opposite points B and D of the lens mount. At point C there is a bright spot. To the right and left of this point, the difference in the distances from B and D increases. (AB–AD). In the image plane, normal to the optical axis, at equal distance from C in all directions, there is a ring-shaped area of substantially lower light intensity, followed by another ring-shaped area with slightly increased light intensity. An attempt to take a photograph of the image of a single self-luminous point, is represented in Figure 5.

This observable phenomenon can be interpreted as interference of diffracted light waves. At a point A (Figure 4) for which the distance to point B is one-half wavelength longer than that to point E, in the center of the emerging wave surface, a crest from E arrives at the same instant as a trough from B. Since the two waves vibrate in the same azimuth, they interfere with each other and produce at A a *minimum* of intensity. All of the other waves diffracted from adjacent points of the coherent wave surface portion between B and E, interfere at A with all of the diffracted waves from adjacent points of the other half of the emerging wave surface, between E and D. Since for all of these waves the pathlength difference is one-half wavelength, the total effect of these interferences is a minimum of intensity at A.

It appears as if an intensity minimum at A is to be expected when the total difference AB–AD is one wavelength.

$$(AB\text{–}AE=\frac{\lambda}{2},\ AE\text{–}AD=\frac{\lambda}{2};\ AB\text{–}AD=\lambda).$$

Actually, a minimum of intensity occurs when the pathlength difference is slightly more than one full wavelength. The reason for this is that the respective *areas* of the circular aperture of the rear lens (back focal plane) of the objective within which interference with one-half of a wavelength occurs are not equal (Figure 6).

Area III is larger than area I and area II is larger than area IV. The result of interferences with half a wavelength from the respective areas is incomplete cancellation of the light intensity.

Very precise measurements and calculations have shown that the first minimum of intensity in the image plane occurs in ring-shaped areas around point C for which the pathlength difference AB–AD is $1{\cdot}22\lambda$. In ring-shaped areas around C for which the pathlength differences are correspondingly greater, there are maxima and minima of higher "orders". These are much weaker and diminishing in intensity.

To summarize: the image of a single, self-luminous point is a "*diffraction* disc" of *finite* diameter, surrounded by ring-shaped minima and maxima, the latter of greatly reduced and rapidly diminishing intensities.

If the image is formed by white light, comprising wavelengths from about 400 to 700 nanometers, the central region of the diffraction disc is white with a colored border and colored ring-shaped maxima, because their distances from point C increase with increasing wavelength.

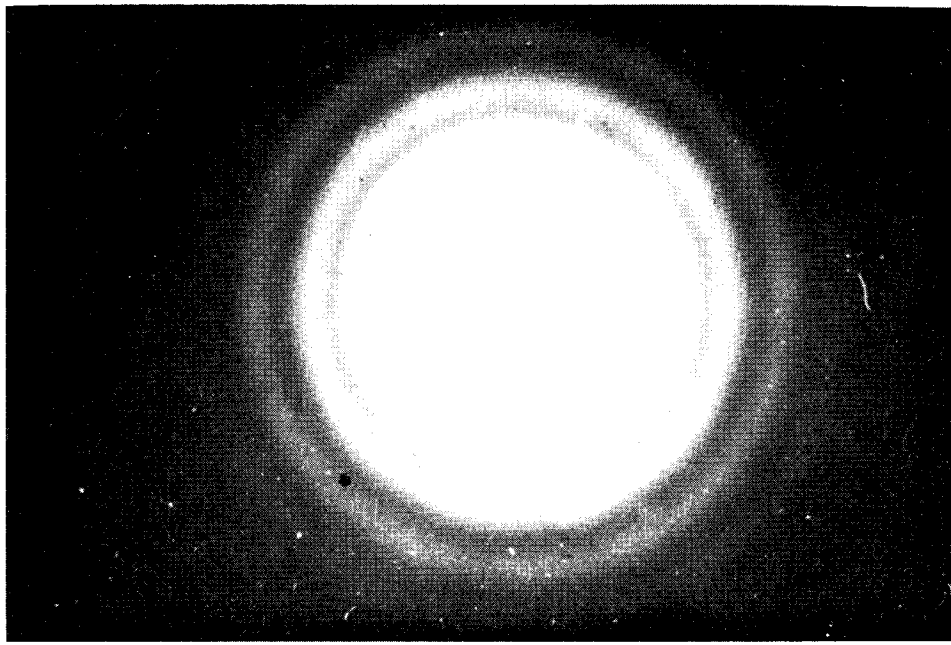

Figure 5. Image of a self-luminous point.

For image formation at a given distance EC (Figure 4) from the rear lens of the objective the *diameter of the diffraction disc* (from minimum to minimum) *decreases* for a given wavelength, as the diameter of the *rear lens increases.*

Formation of the Image of Two or more Self-Luminous Points

If there are two or more self-luminous points in the object plane, separated by small and equal distances D and if the magnification of the image formed by the objective is M, the centers of adjacent diffraction discs are separated by a distance DM. As the distance D decreases, a point is reached at which the interference minima of adjacent diffraction discs touch (Figure 7).

Geometrical interpretations of the relation between D and the numerical aperture (NA) of the objective reveal that this condition prevails when:

$$D=\frac{1{\cdot}22\lambda}{NA} \qquad \ldots (5)$$

Further decrease of the distance D causes adjacent diffraction discs to "overlap". In the overlapping areas, each one of the adjacent object points contributes to the intensity, because there is no interference between the light waves from adjacent object points. There is only a slight decrease of the intensity between the centers of the diffraction

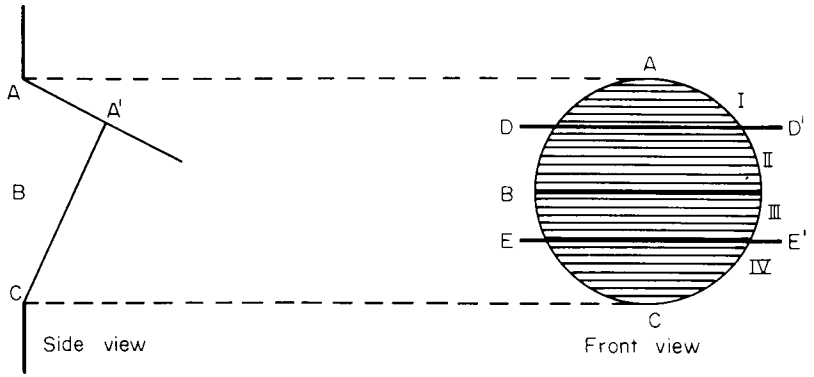

Figure 6. Cross section through the area of the back lens of the objective.

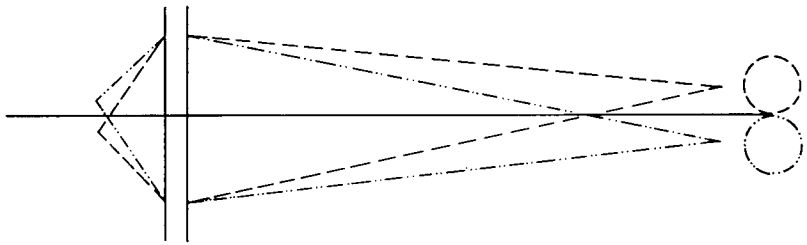

Figure 7. Adjacent interference minima of the image touch each other.

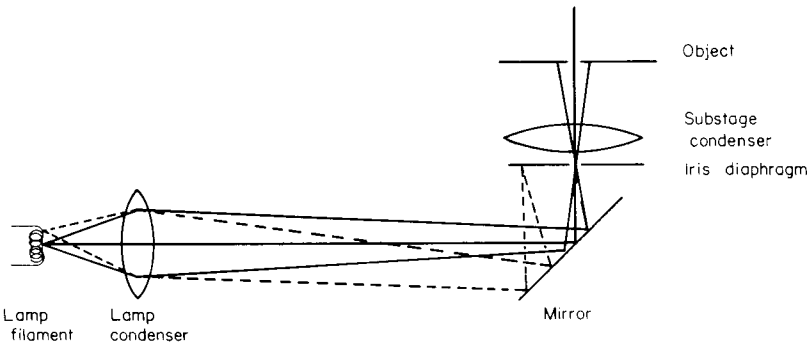

Figure 8. Axial illumination by the substage condenser.

discs. The closer the centers of the diffraction discs, the less is the decrease in light intensity in the overlapping area. Finally, there is no detectable difference and only *one* bright area is observable, although there are *two* self-luminous object points.

The conditions are similar, but in reverse, to those when we see the light from *two* headlights of an approaching automobile at night on a straight road. At a great distance, we see only one bright "glare" of light. As the car approaches, we can detect a *small* decrease of light between *two* centers of high light intensity. At still shorter distances to the approaching car, the area between the centers of light becomes completely dark and wider.

We can also observe thousands of stars at night, each one having the optical properties of a self-luminous point. The image of each star, formed by the lens in the human eye, is a small diffraction disc on the retina. There are areas of the sky in which the images of "adjacent" stars are separated by such small lateral distances that the diffraction discs overlap and we see only one fairly homogeneous area of weak light, known as "the milky way".

When the diffraction discs are separated by spaces, equal to their radii, an interference maximum of one disc, coincides with a minimum of the adjacent disc. Equation (2) is based on the assumption that this is the smallest distance at which it is still possible to detect the existence of *two* points of origin of light in the object plane. This assumption is somewhat arbitrary and is subject to confirmation by experiment and calculation. Experiments of this kind must be carried out with objectives of the highest obtainable perfection because the geometrical conditions on which the equation is based pre-suppose that the emerging wave surfaces are spheres. Aberrations inherent in the optical system cause deviations in shape of the emerging wave surfaces and increase the size of the diffraction discs.

Experiments have proved that with objectives of the highest obtainable perfection and under extremely favorable conditions, it is still possible to detect minute decreases of intensity between adjacent diffraction discs, even when the distance between their centers is smaller than the radius of a disc. To indicate that the resolving power is not a rigid magnitude when images of self-luminous points are formed, but varies within small limits, the factor C has been introduced in equation (4). It varies from a minimum of about 0·4 to about 1·0 depending on such factors as the correction of the objective and the individual capacity of observers to detect minute differences in intensity. In view of this, equation (4) should be used in preference to equation (2). Strictly speaking both equations are applicable to image formation of self-luminous points. Any object of complex structure can be considered as an infinite number of object points. Conditions under which equation (2) can be used for non-self-luminous objects will be covered below.

The images of self-luminous objects have one optical characteristic which is of practical significance. It concerns the relation between *variations* of the light intensity from one object point to another and the corresponding variations of the light intensity in the image. Since there is no interference between the light waves from adjacent object

points, the intensities of the diffraction discs depend only on the magnitudes of the amplitudes of the waves emitted by each object point. That means, that in addition to *morphological resemblance* between object and image, there is also another optical characteristic which may be described as *photometric fidelity of reproduction*. By this term I mean that the variations of light intensity in the image should correspond exactly to the variations in light intensity in the object.

If, for instance, there are two small self-luminous grains of the same size, about equal to the limit of the resolving power of the objective, separated sufficiently so that in the image plane their diffraction discs do not overlap—and the light intensity emitted by one grain is 20% greater than that by the other grain—there should be also a difference of 20% between the brightness of the two diffraction discs. If the higher intensity emitted by one grain—for instance, fluorescence—is caused by higher concentration of the fluorescent stain, it is possible to determine the relative concentrations by photometric measurements. The same can be done with non-self-luminous objects if they are illuminated under conditions of equivalence to self-luminosity. How non-self-luminous objects can be illuminated so that they assume the optical properties of self-luminosity will be explained later.

In view of the trend in microscopy to supplement qualitative observation by quantitative measurements, the fields of micro densitometry, micro photometry, micro spectrophotometry and micro spectrofluorometry gain steadily in popularity. For this type of photometric measurement, photometric fidelity of reproduction is an essential prerequisite.

Image Formation of a Single Non-Self-Luminous Point

In proceeding to descriptions of image formation of non-self-luminous objects, it is advisable to start again with the simplest optical conditions, selecting as an object a single non-self-luminous point.

For practical experiments, it is possible to select as an object a very small hole in the metallic coating of a plane glass plate, if the diameter of the hole is smaller than the limit of the resolving power of the objective. The light must now be traced to its source (Figure 8). This light source should be a single, self-luminous point. Experimentally, a light source of finite size can be used; for instance, a concentrated filament lamp. The light emitted by this source must pass through an illumination system with the following performance.

The first component of the optical system is a collector lens on the illuminator which forms an image of the light source in the lower focal plane of the microscope condenser. In this plane, there is an iris diaphragm, centered with respect to the lens system of the condenser. When this diaphragm is closed as far as possible, it may be assumed that its diameter is about the same as that of the diffraction disc which is the image of a single point of the light source, formed by the collector lens.

From this image point, isolated by the closed iris diaphragm, light waves proceed through the condenser. In so doing, the shapes of the wave surfaces are changed so that, on emergence, they travel in a single direction to form another image of the isolated point of the light source

at infinity. Therefore, a plane wave surface traverses the object plane and light proceeds only in the direction of the optical axis. The metallic coating of the object prevents light from further passage, except for the small portion of the wave surface which has passed through the hole. This hole becomes the origin of diffracted waves which travel towards the objective as spherical wave surfaces. In other words, there is no difference between the physical optical conditions of image formation of a single self-luminous or non-self-luminous point.

Image Formation of Two or more Non-Self-Luminous Object Points

A. *Illuminated with unidirectional light in the direction of the optical axis*

If there are two or more non-self-luminous points, separated by small and equal distances and illuminated by a coherent plane wave surface, parallel to the object plane, there is a radical change in the propagation of light from the object plane through the microscope.

Instead of coherent spherical wave surfaces proceeding towards the objective, light now proceeds only in the directions of the interference maxima. The interference maxima from two adjacent object points travel on a straight line, equidistant from the two points, in the direction of the optical axis, regardless of the wavelength. If there are many object points on a straight line, separated by equal small distances, D, all of the interference maxima without pathlength difference proceed parallel to each other in the direction of the optical axis. These maxima are also called "maxima of zero order". In the passage through the objective, they are refracted and emerge in convergent directions and intersect in the back focal plane of the objective. There, they form an image of the closed iris diaphragm of the substage and also the single point of the light source.

At all points for which the pathlength difference from adjacent object points is *one full wavelength*, another interference maximum occurs, the maximum of the first order. Actually two of these maxima occur, one to the right and one to the left of the maximum of zero order. These maxima of the first order travel on hyperbolic paths. For distances which are very long, their paths are straight lines and parallel to each other in the direction of the asymptotes. They are also refracted in their passage through the objective and intersect in single points of the back focal plane where they form additional images of the aperture iris diaphragm of the condenser and the light source equidistant from the maximum of zero order.

The angle α between the optical axis (the zero maximum) and the direction of the first order maximum depends on the wavelength λ of the image-forming light and the distance D between adjacent object points. This is shown in Figure 9.

The following equations apply:

$$\sin \alpha = \frac{\lambda}{D} \qquad \text{or} \qquad D = \frac{\lambda}{\sin \alpha} \qquad \ldots (6)$$

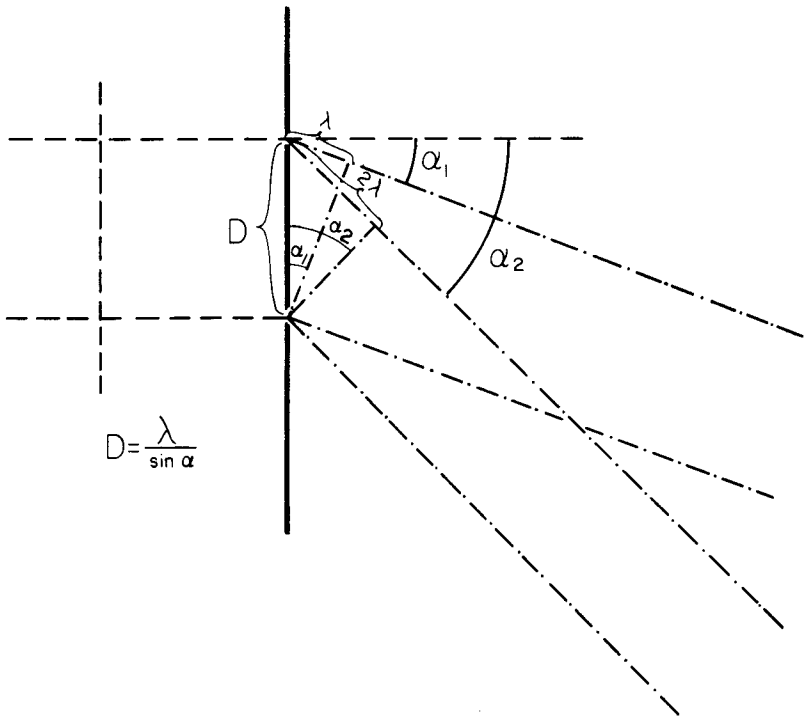

Figure 9. Directions of first and second order interference maxima for axial illumination.

The optical phenomena produced by small (periodic) object structures can be observed in the back focal plane of the objective after removal of the eyepiece (with a Bertrand lens, if present, or with a phase microscope "telescope"). A linear grating can be used for this demonstration. If the grating has 1,000 lines to the millimeter and the objective an NA of at least 0·65 and if a green filter, transmitting light of the wavelength of 500 nanometers, is interposed in the light path, the angle of inclination to the optical axis of the first diffraction maxima can be calculated with the aid of equation (6) as follows:

$$\sin \alpha = \frac{500}{1{,}000} = 0{\cdot}5;\ \alpha = 30^\circ$$

If white light is used for the demonstration, the maximum of zero order appears as a white spot but the first maxima appear as narrow spectra with violet closest to the center and red farther away. Remember the substage iris must be closed.

Light waves, proceeding from these maxima are still traceable to one common center of origin—a single, self-luminous point of the light

source. Recombined in the image plane, they once again interfere with each other. At any point of the image plane for which the pathlength difference to adjacent bright spots in the back focal plane is a full wavelength, there is a maximum of light intensity in the image plane. At points where the pathlength difference is one-half of a wavelength, there is a minimum. The total effect is a continuous pattern of light interference with rather *abrupt* changes from maximum brightness to maximum darkness in which the linear periodical structure of the grating is reproduced on a magnified scale. There is no photometric fidelity of reproduction in this image since the light intensity in each of its points is the result of interference of light waves from *two adjacent object points*.

In comparing the image formation of self-luminous and non-self-luminous objects, there is the following difference: interference of diffracted waves occur only once—in the image plane for self-luminous objects. Therefore, the name "*primary* image formation" is given to it. With non-self-luminous objects, however, interference between diffracted waves occurs *twice*. There are really two image formations: that of the light source in the back focal plane of the objective and that of the object in the image plane. Therefore, this is called "*secondary* image formation".

When the distance D between adjacent object points is so small that the first interference maximum proceeds under an angle of inclination which is higher than that which the objective can collect, only the maximum of zero order passes through the objective. There is only one bright spot in the back focal plane of the objective and no diffraction maxima with which the zero order maximum can interfere. Therefore, the image plane is uniformly illuminated and the structure of the grating is not "resolved".

Equation (6) is based on the assumption that light proceeds in a medium with the refractive index 1·00. If the light proceeds in a medium of higher refractive index, n, its speed is correspondingly reduced. The *frequency of vibration*, however, remains unchanged. This means that the wavelength of light of a given frequency of vibration is reduced from λ to λ/n. Equation (6) is therefore modified to read:

$$D = \frac{\lambda}{n \sin \alpha} \qquad \ldots (6a)$$

Since n sin α is defined as the numerical aperture of the objective, it can be modified again to read:

$$D = \frac{\lambda}{NA} \qquad \ldots (6b)$$

This equation expresses the limit of the resolving power of an objective for unidirectional (axial) illumination. It is the same as the first part of equation (3). Under these illumination conditions, the resolving power is at its worst.

The resolving power of an objective of given NA can be increased by changing the direction of the light illuminating the object.

B. *Unidirectional Oblique Illumination*

To illuminate the object with unidirectional *oblique* light, the aperture iris diaphragm of the condenser must be displaced laterally in its lower focal plane. The maximum of zero order proceeds again with the same angle of inclination as the illumination. The bright spot to which it converges in the back focal plane of the objective is no longer on the optical axis, it is displaced laterally by an equivalent angle. The first order maxima also undergo lateral displacement. One of them proceeds under an *increased* angle of inclination and may not be collected by the objective, the other, however, proceeds at a *lower* angle of inclination to the optical axis. Its image in the back focal plane of the objective is still at the same distance from that of the maximum of zero order and, therefore, it has moved closer to the center of the objective aperture stop.

If, for instance, an objective of NA = 0·50 ($\alpha = 30^\circ$) is used to form an image of a grating with 1,000 lines to the millimeter, illuminated with unidirectional axial light of the wavelength $\lambda = 500$ nanometers, the distance between the maximum of zero order in the focal point of the objective and the symmetrically placed maxima of the first order is equal to the *radius* of the objective aperture stop. Each one of these maxima is visible as a bright spot at the periphery. If a grating of *slightly* finer structure is used, the maxima are not collected by the objective.

If the light proceeds at an angle of 10° to the optical axis, the image formed by the maximum of zero order is laterally displaced in the back focal plane of the objective, about one-third of the radius of the aperture stop from the center. One of the two maxima of the first order now proceeds under an increased angle of 40° to the optical axis and is not collected by an objective of NA = 0·50. The other maximum of the first order proceeds at 20° inclination. This image is laterally displaced from the periphery to a point about two-thirds of the radius of the aperture stop from the center. There are still *two* light centers in the back focal plane of the objective and that is the minimum number of maxima to produce interference. Even if the structure of the grating is finer and the first interference maximum proceeds under an angle of inclination between 20° and 30°, it is still collected by the objective. Through further increase of the obliquity of the illumination, it is possible to resolve even finer structures. The limit is reached when the illumination proceeds at the highest angle of inclination to the optical axis, which the objective can collect. The finest structure which can be resolved under these conditions is one which produces the first maximum under an equally high angle of inclination. This is shown in Figure 10. The distance between adjacent maxima in the back focal plane of the objective is now equal to the *diameter* of the aperture stop. The resolving power has been doubled; equation (1) and the second half of equation (3) are now valid.

Unfortunately, we are not justified in assuming, on the basis of these explanations, the validity of the second part of equation (3) for *multidirectional* illumination even when the numerical aperture of the illumination is equal to that of the objective. Only with *unidirectional* extremely oblique illumination will the described interference phenomena occur.

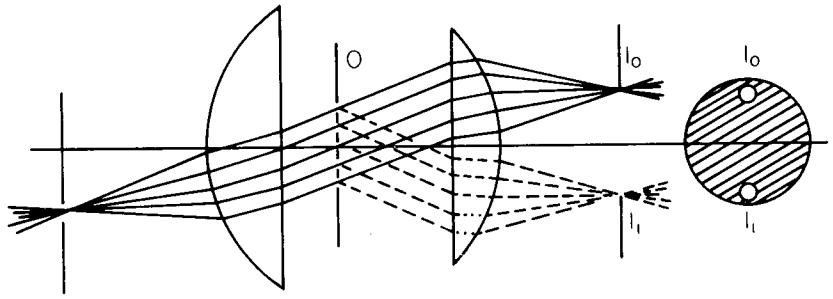

Figure 10. Increased resolving power with extremely oblique illumination of object. The circles at right and left show the zero and first order maxima in the back focal plane of the objective.

Although the magnitude of the smallest resolved structure according to the second part of equation (3) is about the same as that for equation (4) (assuming the minimum magnitude of factor C of slightly more than 0·4), there is a significant difference in the reproduction of object detail of the smallest resolvable magnitude.

When diffracted waves from two *continuous* halves of a wave surface interfere with each other and the total pathlength difference, AB–AD in Figure 4, is *about one full* wavelength an interference *minimum* occurs. When there are only two point-shaped light centers diametrically opposed, one at B and one at D, and the pathlength is exactly *one full* wavelength an interference *maximum* occurs. The *minimum* is at a point for which the pathlength difference, AB–AD, is only *one-half* of a wavelength. That is why, in the image of two self-luminous points, the light intensity *decreases only very slightly* between the centers of two overlapping diffraction discs. In the case of non-self-luminous objects, the intensity *decreases to zero* at the point of the interference minimum. Figure 11 is an attempt to show this sharp change of intensity. The limit of the resolving power is about the same for self- and non-self-luminous objects.

Theories Regarding Image Formation

There is a reason why the names of proponents of theories regarding image formation by the microscope and the resolving power of the objective have not been mentioned so far. It is well known that a theory regarding the image formation of non-self-luminous objects was developed by Professor E. Abbe. His theory is very ingenious, comprehensive and complete. A detailed description of it would go far beyond the scope of this presentation. He has also described a series of fascinating demonstrations in support of his theory. A study of the literature reveals that the Abbe theory did not remain unchallenged. Some of the objections were based on incorrect interpretations of the theory and, in substance, the Abbe theory is as valid today as it ever was.

To complete this presentation, it becomes necessary to describe image formation under conditions of equivalence to self-luminosity and with multidirectional illumination of reduced NA.

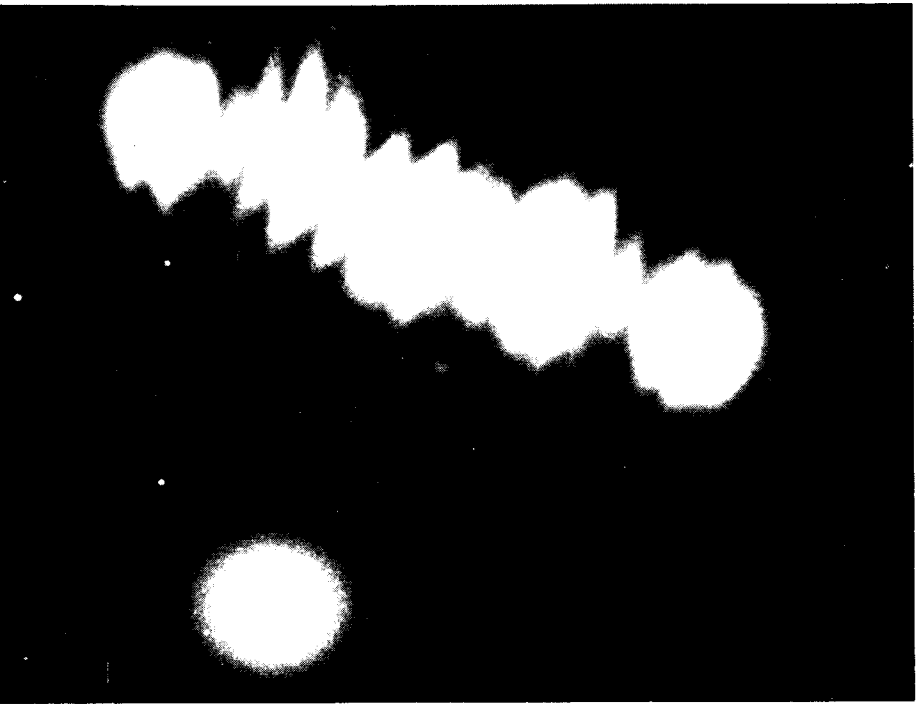

Figure 11. Image of non-self-luminous points showing abrupt changes from brightness to darkness.

C. *Image Formation under Conditions of Equivalence to Self-Luminosity*

If a homogeneous light source of adequate size is used for the illumination of the object and if the intensity of the light emitted throughout an area of adequate diameter is of the same value; if, furthermore, an illumination system is used which forms *an image of the light source in the object plane* and if, finally, the numerical aperture of the illumination is equal to that of the objective, the objects assume the properties of self-luminosity. Each object point is illuminated by light from a single point of the light source and the spherical wave surfaces from adjacent object points do not interfere with each other. This is truly primary image formation. Equation (4) is valid for these illumination conditions. One of the essential requirements for the condenser is, that its correction must be of a high degree of perfection so that the diameters of the diffraction discs of the images of individual self-luminous points of the light source do not exceed the limit of the resolving power of the objective.

This type of illumination became popular in England at a time when the Abbe theory was generally accepted in Germany. It seems that in England, there were more microscopists who observed objects for which photometric fidelity of reproduction was essential (stained objects, for

instance). Lord Rayleigh had strongly advocated image formation under conditions of equivalence to self-luminosity, although this expression was not used at that time. Image formation of the light source in the object plane became known as "critical illumination". It required light sources in which the intensity of the emitted light does not vary from one point to another, throughout an area of such size that its image illuminated the whole field of view. The ribbon filament lamp and the point-o-lite lamp are light sources of this type. About thirty to forty years ago, both types were quite popular. The ribbon filament lamp still enjoys popularity but the point-o-lite lamp has sunk into comparative oblivion.

Since critical illumination imposes restrictions on the selection of the light source and there are many heterogeneous light sources which have other desirable properties, it was only natural that attempts were made to create illumination conditions of equivalence to self-luminosity, using these other light sources.

When the illumination system is used which was described for unidirectional illumination (Figure 8) but with the aperture iris diaphragm of a suitable condenser fully open; when the image of the light source in the plane of this iris diaphragm is at least as large as the diameter of the lower condenser lens and when the NA of the illumination is equal to that of the objective, then the object plane is traversed by an infinite number of plane wave surfaces, each proceeding under a different angle of inclination and a different azimuth. The amplitudes of vibration are of equal magnitude along any single wave surface, but their magnitudes vary from one wave surface to another because of the heterogeneous structure of the light source. The question now arises: will equivalence to self-luminosity still prevail under these conditions?

The practising microscopist will be most interested in an answer based on practical experiments. Such experiments have shown that *the location of the image of the light source with respect to the object plane does not influence the resolving power*. This has been confirmed by theoretical reasoning and calculation. In view of this, there is no reason why heterogeneous light sources should not be used. Illumination of equal intensity throughout the field of view can be obtained with these light sources by adjusting the collector lens of the lamp housing to form an image of the light source in the lower focal plane of the condenser, where the aperture iris diaphragm is located. This method of illumination has become known as Köhler illumination. Practically every microscopist uses Köhler illumination. The centering and aligning of the illumination system has been greatly facilitated by the addition of another iris diaphragm, interposed in the light path at a suitably selected plane so that the condenser forms its image in the object plane. This is called the "field stop".

For critical illumination or Köhler illumination, when conditions of equivalence to self-luminosity prevail, equation (4) is valid for the determination of the limit of the resolving power. Actually, it yields practically the same value as the second part of equation (3), but the optical conditions of image formation are different.

This presentation has now covered two extremes of illumination conditions: unidirectional illumination with an NA which approaches

zero and multidirectional illumination with NA equal to that of the objective. There remain the illumination conditions most frequently used by the practising microscopist, multidirectional illumination, but with its NA smaller than that of the objective.

D. *Multidirectional Illumination with Reduced NA*

When critical illumination is used and the NA of the illumination is substantially lower than that of the objective (about $\frac{3}{4}$ or less) the diameters of the diffraction discs in the object plane (the images of single self-luminous points of the light source) increase to a size where they are larger than the limit of the resolving power of the objective. This means that adjacent object points, separated by a distance slightly larger than the limit of the resolving power of the objective, are illuminated by light from one common center of origin. Light from these points is subject to interference as described under secondary image formation. The result is increased contrast in the image which, however, is restricted to this very small object detail. In general, the optical character of the image retains photometric fidelity of reproduction.

The same effect is produced when Köhler illumination is used. There is also another detectable effect produced under conditions of reduced NA of the illumination. Any deviations from perfection of the correction of the objective, for instance, residual small amounts of spherical aberration, become more noticeable as the NA of the illumination increases. The more perfectly the objective is corrected, the less noticeable is the impairment of image quality under conditions of highest NA of illumination or, the less perfect the objective, the greater is the improvement of image quality when the NA of the illumination is reduced. Furthermore, there is also a slight increase in the depth of field when the NA of the illumination is reduced.

The increase in contrast on reducing the NA of the illumination also depends on the difference between the refractive indices of object and surrounding. If object and surroundings have the same refractive index and differ only with respect to variations of absorption of light from one object point to another, there is practically no noticeable increase of contrast.

There is no fixed ratio between the NA of illumination and objective, at which the quality of the image is at an optimum. The best way to find the optimum ratio is by keenly observing the image of the object while slowly closing the aperture stop of the condenser. This should be stated emphatically in view of the fact that in some publications on microscopy fixed ratios of the NA are recommended. Such recommendations are as unrealistic as the factor 0·61 in equation (2). The change of the optical character occurs rather abruptly and even a novice in microscopy can detect it, if his attention is drawn to it.

Reduced NA with multidirectional illumination is most necessary when the objects are stained preparations and differ from their surroundings primarily in regard to absorption variations. For these objects, an overall photometric fidelity of reproduction between object and image is desirable. When the optical differences between object and surroundings are primarily due to differences between refractive indices, conditions of secondary image formation will increase the contrast in

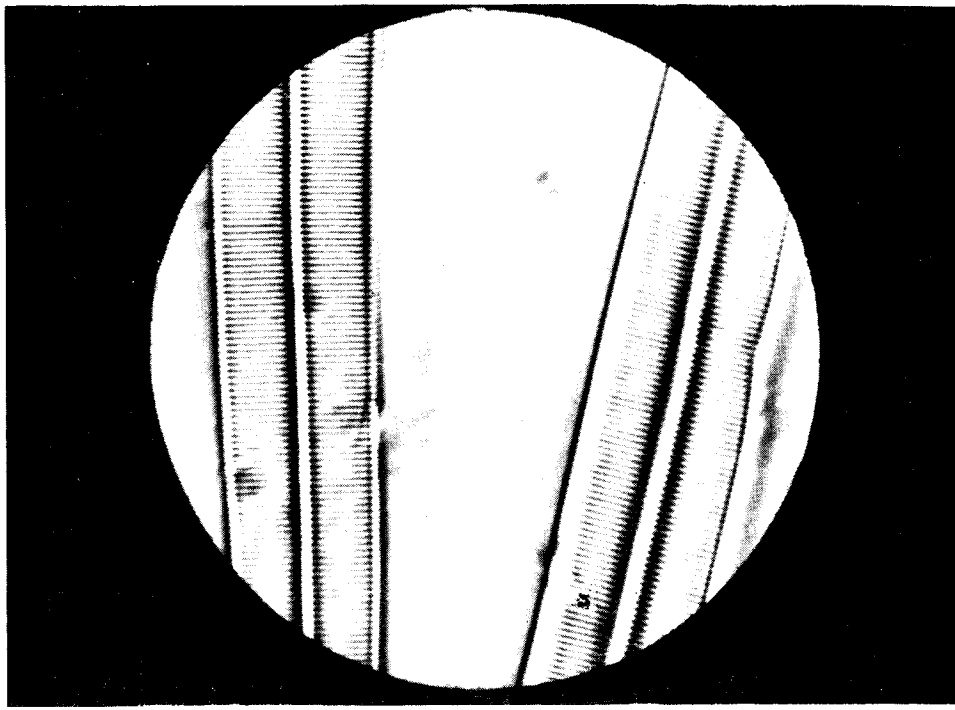

Figure 12. Periodical structures of amphipleura pellucida.

the image. It is unfortunate that, for optimum conditions of contrast in the image—unidirectional illumination—there is either a general reduction of resolving power (for axial illumination) or a non-uniform increase in resolving power (for oblique illumination) (depending on the azimuth of illumination and the direction of linear, periodic structures). Only when extremely small linear periodical structures are to be reproduced in the image is oblique illumination beneficial and then only if the azimuth of the illumination is correctly selected with respect to the alignment of the structures (Figure 12).

There are still many objects with irregular structures for which multidirectional illumination with conditions of approaching photometric fidelity of reproduction is not suitable. The limitations imposed on unidirectional illumination with regard to reduction of resolving power have been lifted by the creation of new illumination methods. The whole field of microscopy has been considerably widened with the creation of phase contrast, interference contrast and other illumination methods, which boost the contrast considerably and do not cause a reduction of the resolving power.

Conclusions

The name "emission image" can be given to that of a self-luminous object (fluorescent) because it reproduces the variations of the light

intensity *emitted* by each one of the infinite number of object points with photometric fidelity.

The name "absorption image" can be given to that of a non-self-luminous object, illuminated by multidirectional light under conditions of equivalence to self-luminosity, because it reproduces the variations of the light intensity caused by *absorption* in the passage through the object with photometric fidelity.

In each case, there is a point-for-point relation between the variations of the light intensities proceeding from object points and their images (down to the limit of the resolving power).

The name "diffraction image" can be given to that of a non-self-luminous object, illuminated by unidirectional light because variations of the light intensity in the image are due to interference of *diffracted* waves arriving at the image points with phase differences of half or full wavelengths or multiples of them as explained. There is no photometric fidelity of reproduction in diffraction images.

Although equation (1) is often considered to express the theoretical limit of the resolving power with the tacit inference that this limit can be approached, but never reached in practice, it has full validity for unidirectional illumination of extreme obliquity and linear periodical structures. This validity exists, even if the objectives and condensers are not of the highest possible perfection and have residual aberrations at the selected wavelength. As long as the magnitude of D is such that adjacent interference maxima are still collected and transmitted by the objective and the images of the light source in the rear focal plane are visible as diametrically opposed bright spots (see Figure 12), an interference pattern of maxima and minima of intensity is produced in the image plane in which the object structure is "resolved".

If, however, the same objectives are used for image formation of the same objects under conditions of equivalence to self-luminosity, the aberrations reduce the resolving power. Furthermore, the linear structures of objects like amphipleura pellucida will become all but invisible under these conditions because the optical differences between these structures and the surroundings are so small that detection of object detail becomes all but impossible, even if the magnitude of such detail is much greater than the limit of resolving power.

There is no optically justifiable reason for extending the validity of equation (1) beyond the conditions for which it was derived. These conditions are rarely, if ever, selected by practising microscopists observing objects with irregular structures and variations of absorption. Therefore, the most frequently quoted equation (1) has the lowest practical value, especially since other equations were derived for conditions most frequently used.

Equation (2) was derived for image formation of self-luminous objects and has equal validity for equivalence to self-luminosity. Its only drawbacks are that in the first place, it does not indicate that under these illumination conditions, not only the numerical aperture of the objective and the wavelength influence the resolving power, but also other factors as, for instance, the state of correction of the objective. In the second place, the factor 0·61 is based on an arbitrary assumption which is not confirmed by experiment. Under extremely favorable conditions, the

resolving power is higher than that expressed by equation (2); under less favorable conditions it is lower. In view of this, equation (2) should be disregarded in favor of equation (4).

The first part of equation (3) is valid for unidirectional axial illumination. Under these conditions, the resolving power is at its worst. Actually, it is impossible in practice to produce truly unidirectional illumination. The second part of equation (3) is identical to equation (1) and requires no additional comment.

The modification of equation (3) cited as (3a) has practical significance. Actually, it was also derived for unidirectional illumination of progressively increasing angle of obliquity. Therefore, the factor $NA_{ill.}$ should be replaced by n sin α to indicate that only light, progressing in the direction of the angle α, should illuminate the object. In practice, it is possibly to apply this equation to conditions of illumination with finite but low numerical aperture, when the aperture iris diaphragm is not completely closed.

Under conditions of lowest achievable NA of illumination, residual aberration of objective and condenser exert no detectable detrimental influence on the resolving power. With progressively increasing NA of the illumination, this influence becomes detectable. Furthermore, there is a gradual change of the optical character of the image from a true diffraction image to one with photometric fidelity of reproduction. This change becomes noticeable at first for the larger object detail and progresses to smaller detail. When the NA of illumination has reached the value at which only the smallest resolvable detail is still reproduced with slightly increased contrast, optimum image quality prevails. This is the condition under which the microscope is most frequently used.

Equation (4), derived for conditions of self-luminosity and also valid for equivalence to self-luminosity, has the greatest practical value because the factor C indicates that the resolving power *varies* within certain limits, depending on such additional optical factors as the state of correction of the objective. For objectives and condensers of highest perfection, the value of C can be assumed to be slightly more than 0·4 and the resolving power actually reaches the value expressed in equation (1), but with entirely different character of the image. For slightly reduced NA of the illumination, to produce optimum image quality, there is no significant increase in the value of C. The practising microscopist should always adjust the aperture stop of the condenser on the basis of the *observed* beneficial effect on the image quality and should not follow the arbitrary suggestion of adjusting this iris diaphragm to a fixed ratio of the NA of illumination to that of the objective.

The reason why multidirectional illumination with slightly reduced NA is used most frequently in practice is that the procedure of enhancing the contrast between object and surroundings most frequently used (at least, by biologists) is that of *staining* the object. This procedure produces variations of light *absorption* by the object and photometric fidelity of reproduction is essential to reproduce the slightest variations of this absorption. For observation of these objects, it is also desirable to use objectives and condensers of a high degree of perfection in their correction.

The disadvantages of diffraction images is that the enhancement of

contrast in the image can only be achieved by either sacrificing resolving power, using illumination of lowest NA or by using oblique illumination of a single azimuth which is advantageous for linear structures but more or less disastrous for objects with irregular structures.

It is still possible to produce images with greatly enhanced contrast, without sacrificing resolving power, by using special illumination methods, such as phase contrast, interference (Nomarsky) contrast and others. As Kipling would say, "This, my friends, is another story."

The Retention of Erythrocytes within the in vitro Coagulum by Submicroscopic Fibrils*

B. J. FORD

Science Unit Laboratory, Mill Park House, 57 Westville Road, Cardiff, South Wales

Introduction

Though the various cellular and biochemical constituents of coagulating blood have been subjected to vigorous and intensive examination for over a century, the erythrocyte is dismissed in relatively few words by the specialist works of reference on the subject[1-3]. This is anomalous, since the erythrocyte in many respects is the foremost constituent of the coagulation process, and it comprises the overwhelming bulk of the contracted clot itself[4]. It is important to remember that the erythrocyte may contain materials of significance in the biochemistry of the process[5]. Further, it is evident that the open network of a plasma clot cannot provide a fluid-tight haemostatic seal, and therefore the presence of the solid mass of erythrocytes is clearly the origin of the haemostatic properties of the clot. Thus there would appear to be functional, biochemical and structural reasons why the erythrocyte should be considered not merely as a histological constituent of the clot, but as an essential element of the coagulation mechanism and the key to its successful function. It is known that only 0·15% of the clot is composed, mass for mass, of fibrin. Thus for this structural integrity to be maintained, efficient deployment of the protein itself is paramount; the length of a fibrin thread of mean cross-sectional area 0·5 μm available on average for each erythrocyte (assuming a calculated volume of approximately 90 μm^3) may be shown by simple arithmetic to be approximately 0·26 μm. Much of the fibrin of a clot may be observed to be very much thicker than this, and so the allocation of fibrin to each constituent cell is meagre. Even on this scale, the length

*Presented at INTER/MICRO-69, 9-11 September, 1969, Imperial College, London, England,