How I Discovered Phase Contrast*

F. Zernike

Department of Physics, University of Groningen, Netherlands

PHASE CONTRAST was not discovered while I was working with a microscope, but while I was working in a different part of optics. It originated in my interest in diffraction gratings, which began about 1920. A diffraction grating consists of a plane or concave mirror with a large number of equidistant grooves ruled on its surface. Small nearly unavoidable imperfections in the location of the grooves show clearly in the optical behavior of the grating. The most conspicuous error is a periodic one that repeats itself after each revolution of the screw of the ruling engine. The regularly recurring displacement of the grooves causes corresponding changes of the optical path, just as if the mirror surface were wavy. Consequently the instrument behaves as if a coarse grating, with a constant of about 2 mm, were superimposed on it, with the result that each strong spectral line is accompanied to its right and left by a number of weak spurious lines, the so-called "Rowland ghosts." These lines have a remarkable effect if one looks down at the surface of the grating, placing the eye at the position of a spectral line. A perfect grating would in this case show an evenly illuminated surface in the color of the spectral line. In reality, however, one sees a strongly striped surface.

At the end of a 1902 paper H. S. Allen remarked that these stripes were nothing real but were simply the effect of the interference between the principal line and its ghosts. Indeed the stripes disappear when the ghosts are covered up. I remember strongly objecting to his conclusion of unreality. On the contrary, I was convinced that the striped surface gave more information about the periodic ruling errors than that obtainable by photographing the ghosts, because in the first case, the relative phases of the ghosts come into play, whereas these are lost in the second case. I kept the question in mind, planning to look further into it as soon as an opportunity arrived.

About 1930 our laboratory obtained a large concave grating ruled by Wood and set it up in a Runge-Paschen mounting. The striped appearance of the surface was soon found, but because the grating was 6 m from the eye, I tried pointing a small telescope at it. Then the unexpected happened. The stripes could be seen very clearly, but they disappeared when the telescope was exactly focused on the surface of the grating! By a succession of experiments and calculations I soon succeeded in explaining this.

On looking back to this event, I am impressed by the great limitations of the human mind. How quick we are to learn—that is, to imitate what others have done or thought before—and how slow to understand —that is, to see the deeper connections. Slowest of all, however, are we in inventing new connections or even in applying old ideas in a new field. In my case the really new point was that the ghosts differed in phase from the principal line. Now it is common knowledge that in all interference phenomena differences of phase are all-important. Why then had phases never been considered before in this case or in the corresponding one in the microscope?

Some excuse may be found in the difficulty to define them exactly. Let me explain this for a simpler case, the diffraction image of a slit. The way to observe this may be as follows. A telescope is pointed at a vertical line-source of light, such as the filament of an incandescent lamp. A vertical slit of, say, 2 mm width is placed close behind the objective of the telescope. This causes the image of the source to be broadened out into a diffraction pattern: a bright central stripe (order zero) is accompanied on both sides by weaker and weaker secondary maxima (orders one, two, and so forth). The formula for this diffraction pattern is given in the textbooks, the amplitude being determined by the function sin x/x. In the few cases where the phases are mentioned in the literature, on the other hand, there is no consensus. Some say that the phases are equal over the whole pattern—except for the obvious reversal of the odd orders—whereas others make them change proportional to x^2. I find that it all depends on the surface, often tacitly assumed, to which the phases are referred. If this reference surface is the focal plane of the telescope objective, one comes to the second statement, if it is a cylindrical surface with the center line of the slit as its axis, the equality of phases results.

You may want to ask whether these phases can be observed. I find they can. All one has to do is to throw the diffraction image on a *coherent background* obtained in the following way. The slit is covered by a glass plate with a thin metallic layer that transmits a small percentage of the light. A fine scratch is made in this layer, forming a narrow slit that is adjusted until it lies in the center of the broad slit. The light through the scratch is broadened out by diffraction and thus forms the desired background, which interferes with the diffraction pattern. The phases of this pattern are thus compared with those of the auxiliary wave forming the background. In the experiment the auxiliary wavefront therefore plays the role of the cylindrical reference surface in the theoretical treatment.

It is only by the introduction of an adequate refer-

* This article is, with some minor alterations, Dr. Zernike's Nobel prize address delivered in Stockholm, Sweden, when the 1953 award in physics was made. It is published with permission of the Nobel Foundation and will also appear in *Les Prix Nobel 1953*.

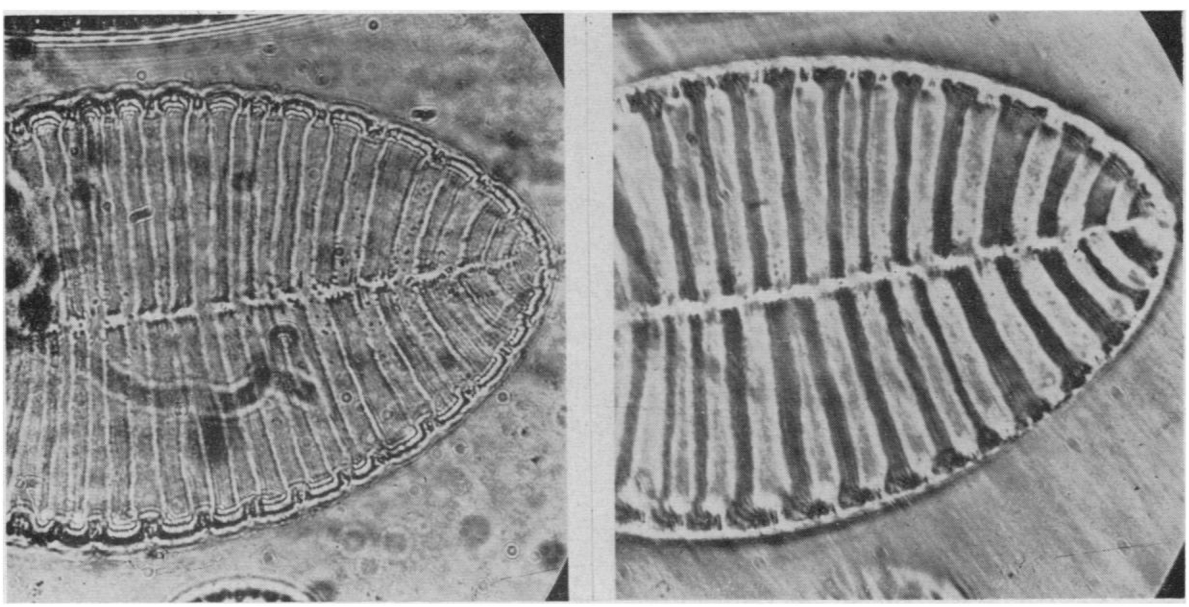
ence surface that a definite statement about the phase differences involved can be made. In the case of the Rowland ghosts the result was that their phases differ by 90° from the principal line. Now I happened to know of a simple method to change this. Lord Rayleigh described in 1900 how to make very shallow etchings in glass surfaces without spoiling their optical quality, by the slow action of very dilute hydrofluoric acid. By this process I made what I called *phase strips*: glass plates with a straight groove, 1 mm or less wide and of a uniform depth of half a wavelength. Such a phase plate was placed in the spectrum so that a bright spectral line fell on the strip, whereas its ghosts passed through the glass beside it. In a telescope behind the phase plate the stripes on the grating surface then stood out clearly.

For a physicist interested in optics it was not a great step to change over from this subject to the microscope. Remember that in Ernst Abbe's remarkable theory of the microscope image the transparent object under the microscope is compared with a grating. To be precise a transmission grating is considered as the test object and the diffraction by this grating as the primary phenomenon. At first sight this has nothing to do with the magnified image of the object formed by the microscope objective. Instead, the objective forms an image of the light source, practically in its back focal plane, consisting of a central direct image accompanied by diffracted images on both sides. This, although on a very much smaller scale, is the analog of the grating line with its ghosts. The light issuing from these images overlaps in the eyepiece of the microscope and by interference gives rise to stripes which, curiously enough, resemble a magnified image of the object! Abbe's theory has been summarized in this sentence: "The microscope image is the interference effect of a diffraction phenomenon."

It is easy to see that, acquainted with this theory, I soon tried my phase strip in a microscope, throwing the direct image of a linear light source on the strip placed close above a low-power objective.

I must now explain why the unexpected discovery of the 90° phase shift applies to the microscope image as well. It all depends on the nature of the object under the microscope. In his theory Abbe and his followers always considered an object of alternate opaque and transparent strips. The diffraction images for such a grating, calculated in the well-known way, are in phase with the central image. On the other hand, if the object consists of alternate thicker and thinner transparent strips, then the phase difference of 90° is found. In the first case, the diffraction is caused by the unequal *amplitudes* of the light passing the strips; in the second case, it is caused by the unequal light paths, that is, by the unequal *phases*. I therefore distinguish the two by calling the first kind an *amplitude grating,* the second a *phase grating,* or in the general case of an irregular structure, an *amplitude object* and a *phase object,* respectively. Nearly all objects of biological or medical interest belong naturally in the second group. The highly developed staining techniques evidently aim at changing them, or the special details one wants to see, into amplitude objects.

It will now be seen that for a phase object my phase strip in the focal plane of the microscope objective brought the direct image of the light source into phase with the diffracted images, making the whole comparable to the images caused by an amplitude object. Therefore the image in the eyepiece ap-



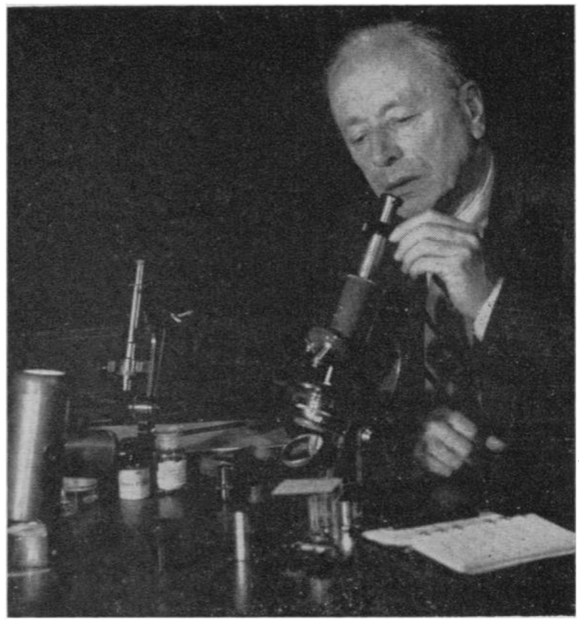
(Left) A diatom with brightfield (traditional narrow iris diaphragm). (Right) The same with phase contrast. Oldest photomicrograph by the author, 1932.

pears as that of an absorbing object—that is, with black and white contrast, just as if the object had been stained. The full name of the new method of microscopy might be something like "phase-strip method for observing phase objects in good contrast." I shortened this into ***phase contrast method.*** Before going into further practical details about the development of the method, a few general remarks should be made.

In a treatise on the Abbe theory, Otto Lummer comes to the conclusion that "in the ideal case the microscope image is exactly similar to the object in structure and phase." Now the absolutely transparent details of a phase object leave the intensity of the passing light unchanged. All they do is impress phase differences on it. According to Lummer, then, the image will show the same phase differences, which however are invisible, and an equal intensity everywhere. In other words, the phase object is absolutely invisible "in the ideal case." Of course the practical microscopist has never been content with this; as a matter of fact, he has never found it out! Without realizing it, he has always turned the fine adjustment —that is, put the object a little out of focus—in order to see the tricky transparent details. Only a somewhat diffuse and watery image is obtained in this way. This also could be exactly explained by the wave theory.

With the phase contrast method still in the first somewhat primitive stage, I went in 1932 to the Zeiss works in Jena to demonstrate it. It was not received with as much enthusiasm as I had expected. This may be explained by the following facts. The great achievements of the firm in practical and theoretical microscopy were all the result of the work of their famous leader Ernst Abbe and dated from before 1890, the year in which Abbe became sole proprietor of the Zeiss works. After 1890 Abbe was absorbed in administrative and social problems, and partly also in other fields of optics. Indeed his last work on microscopy dates from that same year. In it he gave a simple reason for the difficulties with transparent objects, which we now see was insufficient. His increasing staff of scientific collaborators, evidently under the influence of his inspiring personality, formed the tradition that everything worth knowing or trying in microscopy had been already achieved.

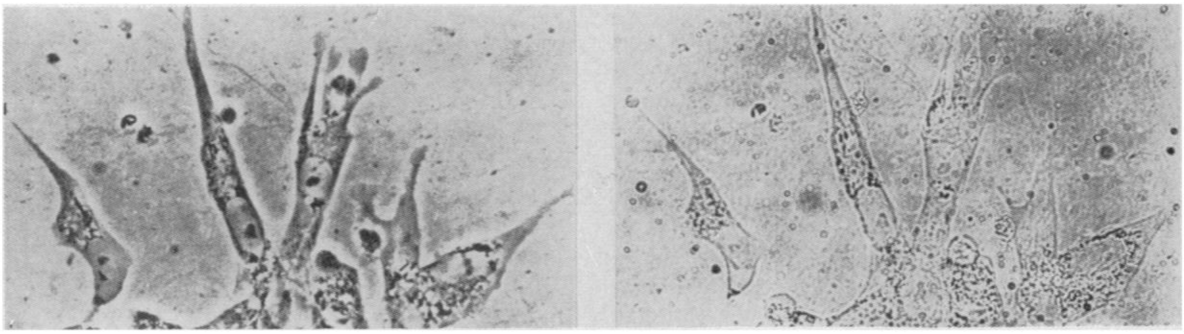
Here is one more remarkable historical point. Whereas all the other achievements of Abbe's were greatly appreciated by all practical microscope users, his theory of image formation was firmly rejected by most of them. To the physicist this may seem incredible, especially when he remembers Abbe's experiments, which in his opinion confirm the theory in a convincing way. The opposing microscopists, however, said these experiments showed only how the microscope may be used, or rather misused, by the physicist for interference experiments that have nothing to do with the ordinary proper use of the instrument. A long story could be told about the violent controversies of this kind that occurred time and again through half a century. This can now be understood because the



Dr. Zernike in his laboratory, November 1953.

theory of Abbe and his followers was too abstract and had been applied only to the oversimplified cases of a point source of light and an object of regular structure. But then it was also incomplete, for it did not explain the peculiarities in the imaging of transparent objects; what is worse, its defenders never recognized this incompleteness. Small wonder therefore that the microscopists rejected the theory as useless in practice.

Returning to the phase contrast method, I will now give a consistent account of its working principle. Let the incident light for simplicity be a plane wave. Without an object—that is, if there is only a clear glass plate under the microscope—this wave passes unchanged, is brought to a focus closely above the objective (in its back focal plane), and spreads out again to an evenly illuminated field in the eyepiece. If there is a real object, every small detail of it will give rise to a slight perturbation of the wave. One may always consider this as resulting from a perturbed wave to be superimposed—in amplitude, not in energy —on the unchanged wave. This last one shall be called the *direct light*; it will clearly give the even background. The perturbed wave will spread out from the detail in all directions, will fill the whole aperture of the objective, and will reunite in the corresponding image point in the eyepiece. The perturbed waves from all the object points together will be called the *diffracted light.* The microscope image in the eyepiece now results from the interference of the diffracted light with the direct light. In order to obtain phase contrast the two must be treated differently, in order to change their relative phases. This is possible because they are spatially separated in the back focal plane of the objective. The interplay of phases in this



(Left) Living tissue culture with phase contrast. (Right) The same with brightfield.

decomposing and reuniting of vibrations can best be visualized in a vector diagram (Fig. 1, left). As is well known, a harmonic vibration is obtained from a vector MV rotating uniformly around M. The projection P on the horizontal axis performs the vibration. The point P', obtained by projection of MV' which remains always perpendicular to MV, performs a similar vibration, one quarter period in advance of P. In accordance with general usage, the projecting is understood, and we speak of the vibrations MV, MV', and so forth.

Now consider a microscopic object with slightly absorbing details on a transparent background (stained preparation). The incident vibration may be represented by MA (Fig. 1, center). An absorbing detail weakens the light, and it gets a smaller amplitude, such as MD. The vector MD results also from compounding MA with MD', with the result that MD' represents the change caused by the detail, that is, the perturbed vibration. Now according to a well-known theorem the optical paths along all rays from an object point to its image are equal. Therefore the direct and the diffracted vibrations arrive at the image point in the same relative phases they had in the object, and the center diagram in Fig. 1 may thus serve for the reuniting of these vibrations. As a result the absorbing detail is seen darker than the background. Now compare this with the case of a transparent object (unstained preparation). Its details will ordinarily be somewhat stronger refracting than the imbedding medium. This means that the light is propagated with less speed and therefore that the emerging vibration MD (Fig. 1, right) will be retarded in phase compared with MA but equal in amplitude. The change caused by the detail is now represented by MD', nearly perpendicular to MA. The compounding of these in the image again gives MD, equal in intensity to the background MA, and the detail remains invisible. It will appear, however, on slightly defocusing, because the light paths are no longer equal in that case, resulting in some change of respective phases. At the same time the image becomes blurred, and the observer has to find a compromise between its disappearance from the first cause, exact focus, and from the other, fading out by lack of focus. In the phase contrast method, however, the direct light has to pass the phase strip, which is thinner than its surroundings, through which the diffracted light passes. The direct light is thus advanced by 90°, being then represented by MPh. This causes the detail to be represented by the vector sum of MPh and MD', making it darker than the background. Clearly the relationships are about the same as they are in the center diagram (Fig. 1), and the transparent detail may be said to be "optically stained."

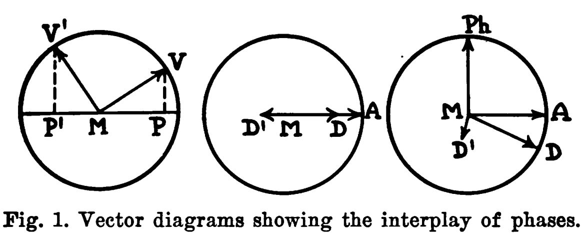


Fig. 1. Vector diagrams showing the interplay of phases.

Two further improvements of phase contrast, which I made in the first years, can now be explained. One is the absorbing phase strip. Details in the object that are very thin will cause only very small phase differences. This corresponds (Fig. 1, right) to a very short vector MD', to be compounded with MPh. The thin detail therefore appears only very little darker than its surroundings, that is, with very little contrast. Now there is no simple way of increasing the amplitude MD' of the diffracted light, but the same result, increased contrast, may be attained by diminishing the amplitude of the direct light MPh. To accomplish this the phase strip must not only accelerate the direct light but also partly absorb it. This is obtained, for instance, by a thin metallic deposit on the strip. An absorption of 75 percent is often used; the strip then transmits 25 percent of the energy, or one-half of the amplitude of the direct light. The contrast is thus doubled, a quite marked effect. In my own experiments I could go down to 4-percent transmission, that is, a 5-times enhanced contrast, the limit being set by the unavoidable stray light. It is only under especially favorable circumstances that a higher increase has been attained by the French astronomer Lyot. In his study of the minute ripples of polished lens surfaces he had independently rediscovered phase contrast and could use strips that diminished the amplitude to one-thirtieth, so that ripples only one one-thousandth of a wavelength high showed in good contrast.

A last point to explain is the halo that is always observed surrounding objects that show great contrast. This must be ascribed to the action of the phase strip on the diffracted light. As we saw before, the phase strip is meant to act only on the direct light. However, the diffracted light, which fills the whole aperture of the objective, will for a small part be intercepted by the phase strip, and this part remains inactive. To find the effect of this missing part, we consider the reverse case, that it would be the only active part. Because of the narrow strip, it would form an image of much less resolving power, that is, blurred by diffraction. Because this part is missing, the "strip image" must be subtracted, in amplitude, from the full image formed by the whole aperture. The interference with the direct light then results in a very diffuse and weak negative image, appearing as a bright halo around dark details and as a dark halo around bright details.

With the straight phase strips used in the beginning, the halo may be disturbing, because the strip image of a small detail is by diffraction spread out in only one direction, namely, perpendicular to the strip. This makes small bright spots in the image appear as if they were marked by short, crossing pencil streaks. To remedy this I soon introduced *annular strips,* which make the halo spread out in all directions, so that it is much fainter and indeed quite harmless.

Zeiss in Jena slowly continued with the design of instruments. After several more of my visits, after some years of development work, and after further delay by the war, they brought out phase contrast objectives and accessories in 1942.

George James Peirce, Pioneer American Plant Physiologist

THE death of George James Peirce on 15 October 1954 marks the passing of a man whose scientific career spanned the entire development of plant physiology in the United States. He was born in Manila on 13 March 1868; when he was 6 years old he returned to the United States with his widowed mother, who established a home in Cambridge, Massachusetts. And after receiving his secondary education in the public schools of Cambridge, Peirce entered Harvard University and graduated in 1890.

Peirce majored in botany at Harvard, and the teacher most influential in directing him toward a particular discipline of botany was George L. Goodale. Goodale's special field of interest was what was called "physiological botany," which placed greater emphasis on structure than on function. Two years after graduating from Harvard, Peirce went to Germany for graduate study. It was natural that, as a man trained under Goodale, he should study both plant anatomy and plant physiology. The first semester in Germany was spent at Bonn in the laboratory of Strasburger, the great plant morphologist. The remainder of his time abroad was spent at Leipzig, primarily in the laboratory of the plant physiologist Pfeffer. In addition, Peirce received extensive training from Fischer in the infant science of bacteriology. His dissertation for the doctorate, which was granted in 1894, was prepared under the guidance of Pfeffer and was entitled "A contribution to the physiology of the genus *Cuscuta.*"

Although Peirce did little original research in bacteriology, he remained interested in its development for many years. However, it is of interest to note that he was the first to offer a course in bacteriology both at Indiana University and at Stanford University. He was among the first in the United States to trace the source of epidemics of typhoid. At Bloomington, Indiana, the source of an epidemic was found to be contamination of the water supply; at Palo Alto, California, it was traced to the milk supplied by a local dairyman.

Upon returning to the United States in 1895, Peirce was appointed assistant professor of botany at Indiana. Two years later he joined the faculty at Stanford, an institution with which he remained associated for the next 59 years. From his first year at Stanford and until he became emeritus in 1933, his primary teaching activity was in the field of plant physiology. In his course on experimental physiology, offered during his first year at Stanford, the emphasis was on function instead of on structure, as in the "physiological botany" he had been taught the decade before.

To Peirce plant physiology was not exclusively a laboratory science but rather was one where illustrative material should be drawn from the outdoors whenever possible. His two books on plant physiology, *Plant Physiology* (1903) and *The Physiology of Plants* (1926), mention numerous examples of the physiology of plants growing in the open. When the weather was favorable, he often took his class in plant physiology outdoors for the lecture. The lectures were presented in a small garden near his laboratory, where he could emphasize a point by directing the students' attention to a nearby plant. Emphasis in the plant physiology that was taught 50 years ago was quite different from that of today. This is well illustrated by the space devoted to different topics in his *Plant Physiology.* At that time the subject of irritability occupied the attention of many plant physiologists, and so it is not surprising to find that nearly a quarter of the book is taken up by the chapter entitled "Irritability." This is in contrast with present-day treatises on plant physiology, in which no author