The Fundamental Idea of Wave Mechanics[1]

When a ray of light passes through an optical instrument, such as a telescope or a photographic lens, it undergoes a change of direction as it strikes each refractive or reflective surface. We can describe the path of the light ray once we know the two simple laws which govern the change of direction. One of these is the law of refraction, which was discovered by Snell about three hundred years ago; and the other is the law of reflection, which was known to Archimedes nearly two thousand years before. *Figure 1* gives a simple example of a ray, A–B, passing through two lenses and undergoing a change of direction at each of the four surfaces in accordance with Snell's law.

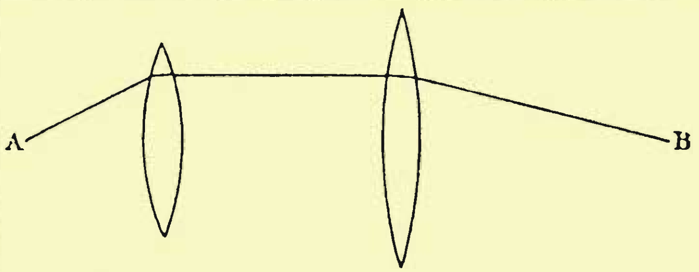


Figure 1.

From a much more general point of view, Fermat summed up the whole career of a light ray. In passing through media of varying optical densities light is propagated at correspondingly varying speeds, and the path which it follows is such as would have to be chosen by the light if it had the purpose of arriving *within the shortest possible time* at the destination which it actually reaches. (Here it may be remarked, in parenthesis, that any two points along the path of the light ray can be chosen as the points of departure and arrival respectively). Any deviation from the path which the ray has actually chosen would mean a delay. This is Fermat's famous *Principle of Least Time*. In one admirably concise statement it defines the whole career of a ray of light, including also the more general case where the nature of the medium does not change suddenly but alters gradually from point to point. The atmosphere surrounding our earth is an example of this. When a ray of light, coming from outside, enters the earth's atmosphere, the ray travels more slowly as it penetrates into deeper and increasingly denser layers. And although the difference in the speed of propagation is extremely small, yet under these circumstances Fermat's Principle demands that the ray of light must bend earthwards (see *Figure 2*), because by doing so it travels for a somewhat longer time in the higher "speedier" layers and comes sooner to its destination than if it were to choose the straight and shorter way (the dotted line in *Figure 2*, the small quadrangle WW W^1W^1 to be ignored for the present). Most people will have noticed how the sun no longer presents the shape

---

[1] Nobel Address delivered at Stockholm on December 12th, 1933.

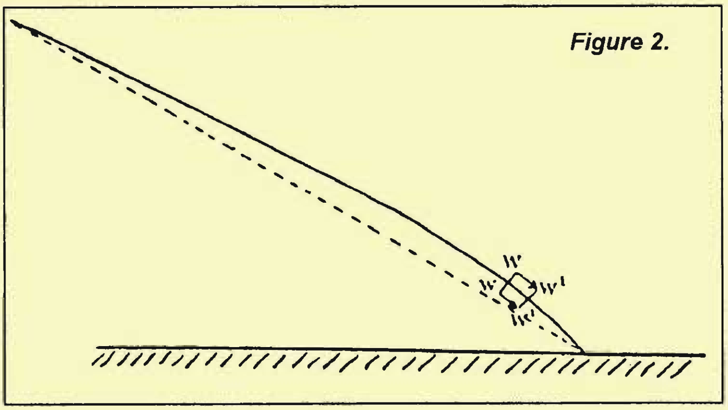


Figure 2.

of a circular disk when it is low on the horizon, but is somewhat flattened, its vertical diameter appearing shortened. That phenomenon is caused by the bending of the light rays as they traverse the earth's atmosphere.

According to the wave theory of light, what we call light rays have, correctly speaking, only a fictitious meaning. They are not the physical tracks of any particles of light, but a purely mathematical construction. The mathematician calls them "orthogonal trajectories" of the wave-fronts, that is, lines which at every point run at right angles to the wave-surface. Hence they point in the direction in which the light is propagated and, as it were, guide the light's propagation. (See *Figure 3*, which represents the simplest case of concentric spherical wave-fronts and the corresponding rectilinear rays, while *Figure 4* illustrates the case of bent rays.) It seems strange that a general principle of such great importance as that of Fermat should be stated directly in reference to these mathematical lines, which are only a mental construction, and not in reference to the wave-fronts themselves. One might therefore be inclined to take it merely for a mathematical curiosity. But that would be a serious mistake. For only from the viewpoint of the wave theory does this principle become directly and immediately intelligible and cease to be a miracle. What we called *bending* of the light ray presents itself to the wave theory as a *turning* of the wave-front, and is much more readily understood. For that is just what we must expect in consequence of the fact that neighbouring portions of the wave-front advance at various speeds; the turning is effected in the same

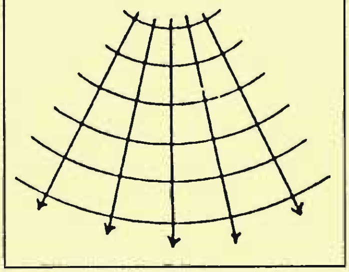


Figure 3.

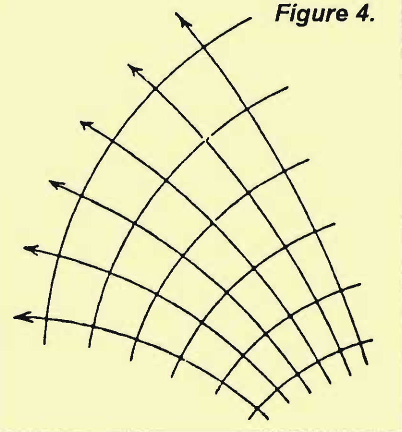


Figure 4.

way as with a company of soldiers marching in line, who are ordered to "right wheel." Here the soldiers in each rank take steps of varying lengths, the man on the right wing taking the shortest steps and the man on the left taking the longest. In the case of atmospheric refraction (*Figure 2*) consider a small portion WW of the wave-surface. This portion must necessarily perform a "right wheel" towards W^1W^1, because its left part is in the somewhat higher and rarer air and therefore is moving forward faster than the right, which is in the deeper layer[2]. Now in examining the case more closely it is found that the statement made in Fermat's Principle is virtually identical with the trivial and obvious assertion that, because the velocity of light varies from point to point, the wave-front must turn, as in the instance I have referred to. I cannot prove that here; but I shall try to show that it is quite reasonable.

Let us revert to the row of soldiers marching in line. To prevent the front rank losing its perfect alignment, let us suppose that a long pole is placed abreast of the men and that each man holds it firmly with his hand against his chest. No word of command as to direction is given, but simply the order that each man must march or run as fast as he can. If the condition of the ground slowly changes from place to place, then either the left or the right section of the line advances more quickly than the other and this inevitably produces quite spontaneously a wheeling of the whole line to the right or left respectively. After a time it will be noticed that the line of advance, when looked upon as a whole, is not straight, but shows a definite curvature. Now this curved route is precisely the one along which the soldiers reach any place on their way *in the shortest possible time,* taking into account the nature of the ground. Although this may seem remarkable there is actually nothing strange about it, for after all, by hypothesis, each soldier has done his best to travel as quickly as possible. And it may be further noticed that the bending will always have taken place in the direction towards which the condition of the ground underfoot is less favorable; so that finally it will appear as if the marchers had purposely avoided unfavorable conditions by making a detour around those regions where they would have found their forward pace slackened.

Thus Fermat's Principle directly appears as the *trivial quintessence* of the wave theory. Hence it was a very remarkable event when Hamilton one day made the theoretical discovery that the orbit of a mass point moving in a field of force (for instance, of a stone

[2] In passing, I may call attention to a point in which Snell's concept fails. A ray of light emitted horizontally ought to remain horizontal, because in the horizontal direction the index of refraction does not vary. But, as a matter of fact, a horizontal ray is deflected to a greater degree than any other. According to the concept of the "wheeling" wave-front, this is obvious.

thrown in the gravitational field of the earth or of some planet in its course around the sun) is governed by a very similar general principle, which thenceforth bore the name of the discoverer and made him famous. Although Hamilton's Principle does not precisely consist in the statement that the mass point chooses the quickest way, yet it states something *so* similar – that is to say, it is *so closely* analogous to the principle of minimum light time – that one is faced with a puzzle. It seemed as if Nature had effected exactly the same thing twice, but in two very different ways – once, in the case of light, through a fairly transparent wave-mechanism, and on the other occasion, in the case of mass points, by methods which were utterly mysterious, unless one was prepared to believe in some underlying undulatory character in the second case also. But at first sight this idea seemed impossible. For the laws of mechanics had at that time only been established and confirmed experimentally on a large scale for bodies of visible and (in the case of the planets) even huge dimensions which played the role of "mass points," so that something like an "undulatory nature" here appeared to be inconceivable.

The smallest and ultimate constructive elements in the constitution of matter, which we now call "mass points" in a much more particular sense, were at that time purely hypothetical. It was not until the discovery of radio-activity that the process of steadily refining our methods of measurement inaugurated a more detailed investigation of these corpuscles or particles; the development was crowned by C.T.R. Wilson's highly ingenious method, which succeeded in taking snap-shots of the track of a single particle and measuring it very accurately by means of stereometric photographs. As far as the measurements go they confirm, in the case of corpuscles, the validity of the same mechanical laws that hold on a large scale, as with planets, etc. Moreover, it was found that neither the molecules nor the atoms are to be considered as the ultimate building stones of matter, but that the atom itself is an extremely complicated composite system. Definite ideas were formed of the way in which atoms are composed of corpuscles, leading to models that closely resembled the celestial planetary system. And it was natural that in the theoretical construction of these tiny systems the attempt was at first made to use the same laws of motion as had been so successfully proved to hold good on a large scale. In other words, we endeavoured to conceive the "inner" life of the atoms in terms of Hamiltonian mechanics, which, as I have said, have their culmination in the Hamiltonian principle. Meanwhile the very close analogy between the latter and Fermat's optical principle had been almost entirely forgotten. Or if any thought were given to this at all, the analogy was looked upon as merely a curious feature of the mathematical theory of the subject.

Now it is very difficult, without going closely into details, to give a correct notion of the success or failure encountered in the attempt to explain the structure of matter by this picture of the atom which was based on classical mechanics. On the one hand the Hamiltonian principle directly proved itself to be the truest and most reliable guide; so much so as to be considered absolutely indispensable. On the other hand, in order to account for certain facts, one had to tolerate the "rude intrusion" (*groben Eingriff*) of quite new and incomprehensible postulates, which were called quantum conditions and quantum postulates. These were gross dissonances in the symphony of classical mechanics—and yet they were curiously chiming in with it, as if they were being played on the same instrument. In mathematical language, the situation may be stated thus: The Hamiltonian principle demands only that a certain integral must be a minimum, without laying down the numerical value of the minimum in this demand; the new postulates require that the numerical value of the minimum must be a whole multiple of a universal constant, which is Planck's Quantum of Action. But this, only in parenthesis. The situation was rather hopeless. If the old mechanics had failed entirely, that would have been tolerable, for thus the ground would have been cleared for a new theory. But as it was, we were faced with the difficult problem of saving its *soul*, whose breath could be palpably detected in this microcosm, and at the same time persuading it, so to speak, not to consider the quantum conditions "rude intruders" but something arising out of the inner nature of the situation itself.

The way out of the difficulty was actually (though unexpectedly) found in the possibility I have already mentioned, namely, that in the Hamiltonian Principle we might also assume the manifestation of a "wave-mechanism," which we supposed to lie at the basis of events in point mechanics, just as we have been long accustomed to acknowledge it in the phenomena of light and in the governing principle enunciated by Fermat. By this, of course, the individual "path" of a mass point absolutely loses its inherent physical significance and becomes something fictitious, just as the individual light ray. Yet the "soul" of the theory, the minimum principle, not only remains inviolate but we could even never reveal its true and simple meaning, as was stated above, *without* introducing the wave theory. The new theory is in reality no *new* theory, but is a thorough organic expansion and development, one might almost say merely a re-statement of the old theory in more subtle terms.

But how could this new and more "subtle" interpretation lead to results that are appreciably different? When applied to the atom, how could it solve any difficulty which the old interpretation could not cope with? How can this new standpoint make the "rude intruder" not merely tolerable but even a welcome guest and part of the household, as it were?

These questions, too, can best be elucidated by reference to the analogy with optics. Although I have asserted, and with good reason, that Fermat's principle is the quintessence of the wave theory of light, yet that principle is not such as to render superfluous a more detailed study of wave processes. The optical phenomena of *diffraction* and *interference* can be understood only when we follow up the particulars of the wave process; because these phenomena depend not merely upon where the wave finally arrives but also on whether at a given moment it arrives there as a wave-crest or a wave-trough. To the older and cruder methods of investigation interference phenomena appeared as only small details and escaped observation. But as soon as they were observed and properly accounted for by means of the undulatory theory, quite a number of experimental devices could be easily arranged in which the undulatory character of light was prominently displayed, not only in the finer details but also in the general character of the experiment.

To explain this I shall bring forward two examples: the first is that of an optical instrument, such as a telescope or a microscope. With such an instrument we aim at obtaining a sharp image. This means that we endeavor to focus all the rays emitted from an object point and re-unite them at what is called the image point (see *Figure 1a*). Formerly it was thought that the difficulties which stood in the way were only those of geometrical optics, which are actually very considerable. Later it turned out that even in the best constructed instruments lack of precise focusing was considerably greater than might have been expected if in reality each ray, independently of its neighboring ray, followed Fermat's principle exactly. The light which is emitted from a luminous point and received by an instrument does not focus at an exact point after it has passed the instrument. Instead of this, it covers a small circular area, which is called the diffraction image and which is mostly circular only because the diaphragms and the circumference of the lenses are usually circular. For diffraction results from the fact that the instrument cannot possibly receive the whole of the spherical waves which are emitted from a luminous point. The borders of the lenses, and sometimes the diaphragms, cut off a part of the wave surface (*Figure 1b*) and – if I may use a somewhat crude expression – the torn edges of the wound prevent an exact focus at a point and bring

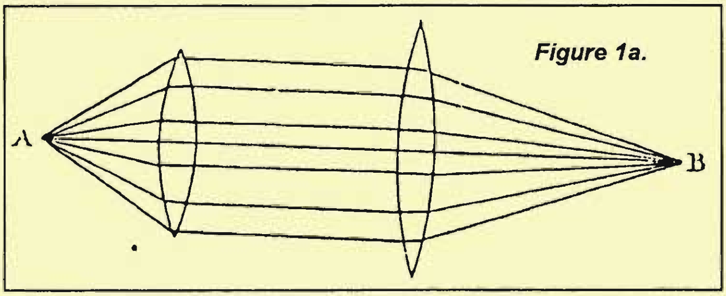


Figure 1a.

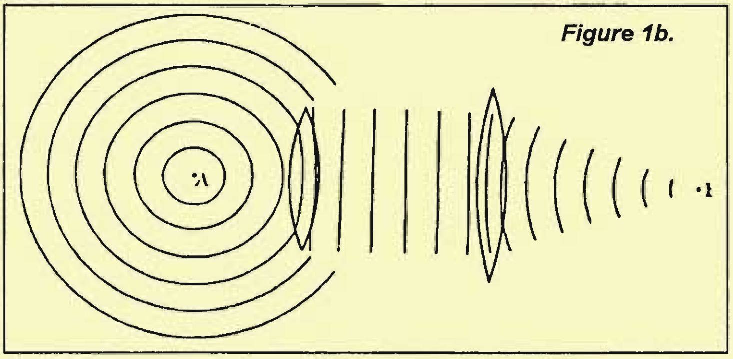


Figure 1b.

about the indistinctness or blurring of the image. This blurring is closely connected with the *wave-length* of the light and is absolutely unavoidable, owing to this deep seated theoretical connection. This phenomenon, originally scarcely noticed, now completely governs and inescapably limits the efficiency of the modern microscope, all the other causes of the lack of distinctness in the image having been successfully overcome. With respect to details, which are not much more coarse-grained than the wave-length of light, the optical image can only reach a distant similarity to the original, and none at all whenever the structural details in the object are *finer* than the wave-length.

The second example is of a simpler nature. Let us take a tiny source of light, just a point only. If we place an opaque body between it and a screen we find a shadow thrown on the screen. To construct the shadow theoretically we should follow each ray of light emitted from the point and should ascertain whether the opaque body prevents it from reaching the screen. The *rim* of the shadow is formed by those light rays which just graze and pass by the outline of the opaque body. But it can be shown by experiment that even where the light source is made as minute as possible, and the outline of the opaque body as sharp as possible, the outer rim of the shadow cast by the opaque body on the screen is not really sharp. The cause of this is again the same as in the former example. The wave-front is split, as it were (*Figure 5*), by the outline of the opaque body; and the traces of this lesion blur the rim of the shadow. This would be inexplicable if the individual light rays were independent in themselves and traveled independently with no reference to one another.

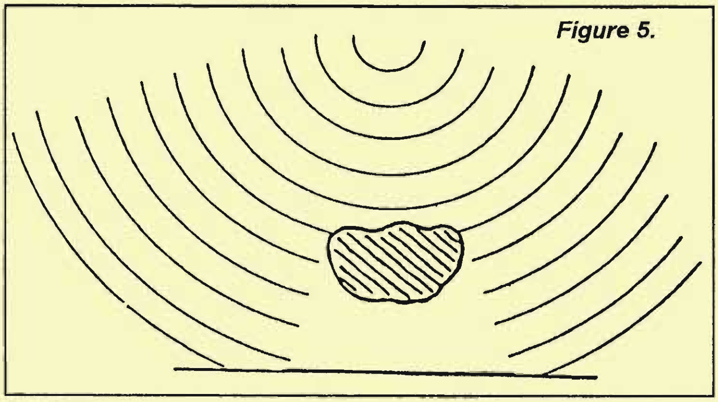


Figure 5.

This phenomenon, which is also called *diffraction,* is generally speaking not very noticeable where larger bodies are concerned. But if the opaque body which throws the shadow be very small, at least in one dimension, then the diffraction has two effects, first, nothing like a true shadow is produced and, secondly – which is far more striking – the tiny body seems to be glowing with its own light and emitting rays in all directions (predominantly, however, at very narrow angles with the incoming rays). Everybody is familiar with the so-called "motes" that appear in the track of a sunbeam entering a dark room. In the same way the filigree of tiny strands and cobwebs that appear around the brow of a hill behind which the sun is hidden, or even the hair of a person standing against the sun, sometimes glows marvelously with diffracted light. The visibility of smoke and fog is due to the same phenomenon. In all these cases the light does not really issue from the opaque body itself but from its immediate surroundings, that is to say, from the area in which the body produces a considerable perturbation of the incident wave-fronts. It is interesting, and for what follows very important, to note that the area of perturbation is always and in every direction at least as large as one or a few wave-lengths, no matter how small the opaque body may be. Here again, therefore, we see the close relation between wave-length and the phenomenon of diffraction. Perhaps this can be more palpably illustrated by reference to another wave process, namely, that of sound. Here, on account of the much longer wave-length, which extends into centimeters and meters, the shadow loses all distinctness and the diffraction predominates to a degree that is of practical importance. We can distinctly *hear* a call from behind a high wall or around the corner of a solid building, although we cannot *see* the person who calls.

Let us now return from optics to mechanics and try to develop the analogy fully. The optical parallel of the *old* mechanics is the method of dealing with isolated rays of light, which are supposed not to influence one another. The new wave mechanics has its parallel in the undulatory theory of light. The advantage of changing from the old concept to the new must obviously consist in clearer insight into diffraction phenomena, or rather into something that is strictly analogous to the diffraction of light, although ordinarily even less significant; for otherwise the old mechanics could not have been accepted as satisfactory for so long a time. But it is not difficult to conjecture the conditions in which the neglected phenomenon must become very prominent, entirely dominate the mechanical process and present problems that are insoluble under the old concept. This occurs inevitably *whenever the entire mechanical system is comparable in its extension with the wave-lengths of "material waves,"* which play the same role in mechanical processes as light waves do in optics.

That is the reason why, in the tiny system of the atom, the old concept is bound to fail. In mechanical phenomena on a large scale it will retain its validity as an excellent approximation, but it must be replaced by the new concept if we wish to deal with the fine interplay which takes place within regions of the order of magnitude of only one or a few wave-lengths. It was amazing to see all the strange additional postulates, which I have mentioned, arising quite automatically from the new undulatory concept, whereas they had to be artificially grafted onto the old one in order to make it fit in with the internal processes of the atom and yield a tolerable explanation of its actually observed manifestations.

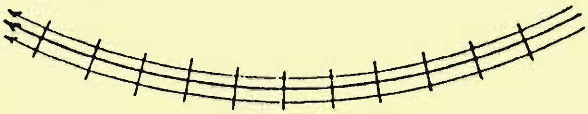
In this connection it is, of course, of outstanding importance that the diameter of the atom and the wave-length of these hypothetical "material" waves should be very nearly of the same order of magnitude. And you will undoubtedly ask whether we are to consider it as purely an accident that in the progressive analysis of the structure of matter we should just here encounter the wave-length order of magnitude, or whether this can be explained. Is there any further evidence of the equality in question? Since the material waves are an entirely new requisite of this theory, which had not been hitherto discerned elsewhere, one might suspect that it is merely a question of suitable *assumption* as to their wave-length, an assumption forced upon us in order to support the preceding arguments.

Well, the coincidence between the two orders of magnitude is by no means a mere accident, and there is no necessity to make any particular assumption in this regard; the coincidence follows naturally from the theory, on account of the following remarkable circumstances. Let us begin by stating that Rutherford's and Chadwick's experiments on the dispersion of Alpha rays have firmly established the fact that the heavy *nucleus* of the atom is very much smaller than the atom, which justifies us in treating it as a point-like center of attraction in all the argument which follows. Instead of the *electron* we introduce hypothetical waves, the wave-length of which is left an open question as yet, because we do not know anything about it. It is true that this introduces into our calculations a symbol, say *a*, which represents a number as yet undefined. But in such calculations we are accustomed to that sort of thing and it does not hinder us from inferring that the nucleus of the atom will inevitably produce a sort of diffraction phenomenon of these waves, just as a minute mote does with light waves. Precisely as with light waves, here too the extension of the perturbed area surrounding the nucleus turns out to bear a close relation to the wave-length and to be of the same order of magnitude. Remember that the latter had to be left an open question! But now comes the most important step: *we identify the perturbed area, the diffraction halo, with the atom; the atom being thus regarded as really nothing more than the diffraction phenomena arising from an electron wave that has been* *intercepted by the nucleus of the atom.* Thus it is no longer an accident that the size of the atom is of the same order of magnitude as the wave-length. It is in the nature of the case itself. Of course numerically we know neither the one nor the other; because in our calculation there always remains the *one* undefined constant which we have called *a*. It can, however, be determined in two ways, which control one another reciprocally. Either we can choose for *a* that value which will quantitatively account for the observable effects produced by the atom, especially for the emitted spectral lines, which can be measured with extreme accuracy; or, in the second place, the value of *a* can be adapted in order to give to the diffraction halo the right size, which from other evidence is to be expected for the atom. These two ways of defining *a* (of which the second is, of course, much less definite, because the phrase "size of the atom" is somewhat indefinite) are *in perfect accord with one another.* Thirdly, and finally, it may be remarked that the constant which has remained indeterminate has not really the physical dimension of Length, but of Action, that is, energy multiplied by time. It is, then, very tempting to assign to it the numerical value of Planck's universal Quantum of Action, which is known with fair accuracy from the laws of heat radiation. The result is that with all desirable exactitude, *we now fall back upon the first (the most exact) method of determining a.*

Thus, from the quantitative point of view, the theory answers its purpose with a minimum of new assumptions. It involves a single available constant, to which we only have to assign a numerical value that is already quite familiar to us in the earlier Quantum Theory, in order, first, to give the proper magnitude to the diffraction halos and therewith render possible their identification with the atoms; and, secondly, to calculate with quantitative exactitude all the observable effects produced by the atoms, their radiation of light, the energy required for ionisation, etc., etc.

I have tried to explain to you in the simplest possible manner the fundamental concept on which this wave theory of matter is based. Let me confess that, in order to avoid bringing the subject before you in an abstruse form at the very outset, I have embellished it somewhat. Not indeed as regards the thoroughness with which conclusions properly deduced from the theory have been corroborated by experiment, but rather as regards the conceptual simplicity and absence of difficulty in the chain of reasoning which lead to these conclusions. In saying this I do not refer to the mathematical difficulties, which eventually are always trivial, but rather to the conceptual difficulties. Naturally it does not call for a great mental effort to pass from the idea of a path to a system of wave-fronts perpendicular to the path (see *Figure* 6). But the wave-surfaces, even when we restrict them to small elements of surface, still involve at least a slender *bundle* of possible paths,

Figure 6.



to all of which they stand in the same relation. According to the traditional idea, in each concrete case one of these paths is singled out as the one "really traveled," in contradistinction to all the other "merely possible" paths. According to the new concept the case is quite different. We are confronted with the profound logical antithesis between

Either this or that (Particle Mechanics)
(*aut — aut*)
and
This as well as that (Wave Mechanics)
(*et — et*)

Now this would not be so perplexing if it were really a question of abandoning the old concept and *substituting* the new one for it. But unfortunately that is not the state of affairs. From the standpoint of wave mechanics the innumerable multitude of possible particle paths would be only fictitious and no single one would have the special prerogative of being that actually traveled in the individual case. But, as I have already remarked, we have in some cases actually observed such individual tracks of a particle. The wave theory cannot meet this case, except in a very unsatisfactory way. We find it extraordinarily difficult to regard the track, whose trace we actually *see*, only as a slender bundle of equally possible (*Gleichberechtigten*) tracks between which the wave-fronts form a lateral connection. And yet these lateral connections are necessary to the understanding of diffraction and interference phenomena, which the very same particles produce before our eyes with equal obviousness – that is to say, produce experimentally on a large scale and not only in those concepts of the interior of the atom discussed previously. It is true that we can deal with every concrete individual case without the two contrasted aspects leading to different expectations as to the result of any given experiment. But with the old and cherished and apparently indispensable concepts, such as "really" and "merely possible," we cannot advance. We can never say what really *is* or what really *happens*, but only what is *observable*, in each concrete case. Shall we content ourselves with this as a permanent feature? In principle, yes. It is by no means a new demand to claim that, in principle, the ultimate aim of exact science must be restricted to the description of what is really observable. The question is only whether we must henceforth forgo connecting the description, as we did hitherto, with a definite hypothesis as to the real structure of the Universe. To-day there is a wide-spread tendency to insist on this renunciation. But I think that this is taking the matter somewhat too lightly.

I would describe the present state of our knowledge as follows: The light ray, or track of the particle, corresponds to a *longitudinal* continuity of the propagation process (that is to say, *in the* direction of the spreading); the wave-front, on the other hand, to a *transverse* one, that is to say, perpendicular to the direction of spreading. *Both* continuities are undoubtedly real. The one has been proved by photographing the particle tracks, and the other by interference experiments. As yet we have not been able to bring the two together into a uniform scheme. It is only in extreme cases that the transverse – the spherical – continuity or the longitudinal – the ray-continuity shows itself so predominantly that we *believe* we can avail ourselves either of the wave scheme or of the particle scheme alone.

Erwin Schrödinger

Nobelists and family members at the Stockholm train station, 1933, From right: Schrödinger, Heisenberg, Dirac, Dirac's mother, Schrödinger's wife and Heisenberg's mother.