THE GREAT PROBLEMS OF SCIENCE

WORKS COLLECTED BY MME. P. FÉVRIER

VII

THE THEORY OF MEASUREMENT IN WAVE MECHANICS

(USUAL INTERPRETATION AND CAUSAL INTERPRETATION)

BY

Louis de BROGLIE

of the French Academy
Perpetual Secretary of the Academy of Sciences
Professor at the Sorbonne

Translated by

D. H. Delphenich

PARIS

GAUTHIER-VILLARS, ÉDITEUR-IMPRIMEUR-LIBRAIRE

Quai des Grands-Augustins, 55

—

1957

PREFACE.

The present volume defines a sort of complement to the book that I recently published on the interpretation of wave mechanics by the theory of the double solution ([1]). I recall, in more detail, certain questions that seem to me to necessitate a new examination of the role of measurement in quantum physics, but developed in a more concrete fashion that is closer to experimental reality than what has been done up to now.

The plan of this book is the following: After recalling some well-known principles of wave mechanics in the first chapter, I will present the theory of measurement that is due to J. von Neumann in chapters II and III, while presenting some arguments that were developed by Einstein and Schrödinger not long ago, and I will show that this theory, despite its elegant character and the perfectly satisfying appearance of its formalism, nonetheless leads to some consequences that are very difficult to accept. The difficulties that it raises derive, on the one hand, from the fact that, in accord with presently dominant ideas, it does not allow for the permanent localization of corpuscles in space and, on the other, that it visualizes the processes of measurement in a very abstract manner.

After summarizing the fundamental concepts of the theory of the double solution in chapters IV and V, while adding some complementary notions that did not find their places in my previous treatises, I recall the study of the processes of measurement in chapter VI and VII from a concrete viewpoint. I will introduce the essential ideas that wave trains are always bounded and that we can make observations or measurements on microphysical reality only by the intermediary of observable, macroscopic phenomena that are triggered by the local action of a corpuscle. Upon adding to these fundamental remarks the idea of the permanent localization of corpuscles in space such as would result from the theory of the double solution, I will show that one thus obtains a clear image of the processes of measurement that do not raise the same objections are the theory of von Neumann and his heirs.

A last chapter is dedicated to a very rapid examination of von Neumann's thermodynamics and its interpretation with the aid of the ideas that were discussed previously.

The goal of the present book is, in summation, to exhibit the reasons by which it seems to me to re-establish the notion of a permanent localization of microphysical corpuscles, and why, once I again became aware of that necessity, I sought in recent years to resume the attempt to interpret wave mechanics that I sketched out in 1927.

September 1956

Louis DE BROGLIE

([1]) Bibliography [**3**].

TABLE OF CONTENTS

Chapter IV.

Causal interpretation of wave mechanics (theory of the double solution).

CHAPTER V.

Some complementary notions to the theory of the double solution and guidance.

CHAPTER VI.

Position of the causal interpretation in regard the problem of measurement in microphysics.

CHAPTER VII.

Measurement of quantities by the interaction of two corpuscles.

CHAPTER VIII.

Glimpse of von Neumann's thermodynamics.

APPENDIX.

_____

CHAPTER I

REVIEW OF SOME GENERALITIES ON WAVE MECHANICS AND MEASUREMENT

---

1. Some known principles of wave mechanics. – The present interpretation that wave mechanics allows supposes that one can describe a corpuscle or a system of corpuscles in as complete a fashion as possible with the aid of a wave function Ψ that is, moreover, capable of having several components, as in the Dirac theory of the electron or in that of corpuscles with higher spin. The function Ψ is always assumed to be "normalized" by the formula:

$$\int |\Psi|^2 \, d\tau = 1. \tag{1}$$

The evolution of the wave function in the course of time is governed by a partial differential equation – viz., the wave equation – which is the well-known Schrödinger equation in the simplest case of a corpuscle without spin in the non-relativistic approximation. It will take on a more complicated form for particles with spin (the Dirac electron, for example), because in these cases it will become, in reality, a system of partial differential equations that couple the various components of Ψ. In a general fashion, the wave equations, along with the initial conditions and boundary conditions, will determine the evolution of the function Ψ completely.

If one completely forgets the origins of wave mechanics and the physical intuitions upon which it is founded then most authors will consider the function Ψ to be a simple mathematical instrument that serves to predict the probabilities of the various results of measurements that are performed in the corpuscle or the system, since that function (by chance?) will have the same form as the waves of classical physics.

Now, here are – briefly summarized – the postulates that constitute, in a way, the "recipes" that permit one to utilize the function Ψ – which is assumed to be known – for the calculation of the probabilities of the measurements that one can perform on corpuscular quantities. One assumes that each of these quantities will correspond to a linear, Hermitian operator A whose proper-value equation:

$$A\varphi = \alpha\varphi \tag{2}$$

will permit one to define a continuous or discontinuous (or even partially continuous and partially discontinuous) set of proper values α and corresponding proper functions $\varphi(\alpha)$. The proper functions φ will form a complete system of functions with an orthonormal basis, in such a way that one can always write:

$$\Psi = \int c(\alpha)\, \varphi(\alpha)\, d\alpha, \tag{3}$$

or, more simply, in the case of a discontinuous spectrum:

(4) $$\Psi = \sum_i c_i \varphi_i ,$$

upon enumerating the proper values and the proper functions by an index. Moreover, a mathematical formalism like the Stieltjes integral will permit one to combine the two cases of continuous spectrum and discontinuous spectrum into just one formula. The set of proper values of A define the "spectrum" of that operator.

The fundamental principle that one takes to be the basis is then the following one: Let Ψ be the wave function of a corpuscle (or a system), upon which one must perform the measurement of a quantity A, with the aid of an appropriate device. One will develop Ψ in proper functions φ of the corresponding operator A, and one can assert that the probability for the measurement to give a value that belongs to an interval $d\alpha$ is $| c(\alpha) |^2 d\alpha$. In the case of a discontinuous spectrum, one will say, more simply, that the probability of the value of given by $| c_i |^2$.

The mathematical expectation of the value α_i – or, if one prefers, the mean value of the result of the measurement of A that is performed upon a very large number of corpuscles that have the same function Ψ – will be:

(5) $$A = \sum_i | c_i |^2 \alpha_i = \int \Psi^* A \Psi \, d\tau.$$

When these general principles are applied to the measurement of the position of a corpuscle, that will give the following result: The probability for the coordinates of a corpuscle to be found inside the interval $x \to x + dx$, $y \to y + dy$, $z \to z + dz$ – i.e., in order for the corpuscle to be found in the volume element $d\tau = dx\, dy\, dz$ – will be $| \Psi |^2 dx\, dy\, dz$. An analogous statement will be valid for the probability of the presence of the figurative point of a system in the configuration space to which it corresponds.

The statements that relate to $| \Psi |^2$ (e.g., the principle of interference or its localization) can be deduced from the general formalism in such a way that, from the viewpoint of that formalism, the probability of presence $| \Psi |^2$ will be seen to have the same status as any other probability $| c_i |^2$. The set of all possible developments of Ψ in the different systems of proper functions φ_i that correspond to the various measurable quantities will thus appear to be entirely equivalent from the formula standpoint. That idea, which serves as the basis for the "theory of transformations," gives rise to some elegant mathematical developments, although we shall discuss only its physical significance.

The general postulate that was assumed above in relation to the statistical significance of the $| c_i |^2$ will imply, by an argument that I will not reproduce, the following consequence: The same experimental device can permit one to measure two quantities A and B simultaneously with any precision only if the corresponding operators commute; i.e., if one has $AB\varphi = BA\varphi$ for any φ. If that were not true – i.e., if $AB\varphi \neq BA\varphi$, in general – then any experimental device that would permit one to attribute a value to A that is affected with a certain uncertainty and would leave behind an uncertainty in the value of B that is greater than the measurement of A would have to be more precise, and conversely. The typical example of two quantities that are not simultaneously

measurable with precision is provided by any pair of quantities that are "canonically" conjugate, in the sense of analytical mechanics, such as, for example, the coordinate x of a corpuscle and the corresponding component p_x of the quantity of motion. In the latter case, the corresponding operators (which are x and $-\frac{h}{2\pi i}\frac{\partial}{\partial x}$) are such that $AB - BA = \frac{h}{2\pi i}$, and in turn, will not commute. One then shows that the uncertainties that exist in the values of x and p_x will always satisfy the Heisenberg inequalities:

$$\delta x \, \delta p_x \geq h, \tag{6}$$

and, in turn, can never be zero simultaneously.

Moreover, there exist quantities that, without being canonically conjugate, nonetheless, do not commute; for example, the three rectangular components M_X, M_Y, M_Z of the moment of the quantity of motion, for which one will find:

$$M_X M_Y - M_Y M_X = \frac{h}{2\pi i} M_Z \, , \ldots$$

and one will then show that the uncertainties in the values of two of these components cannot be zero simultaneously, in general.

One can translate these results into a somewhat different language by saying that our general principle will make the value of any measurable physical measurement correspond to a probability distribution that has the form of Ψ. In the discontinuous case, the probabilities of the values α_i will be $P_i = | c_i |^2$, and in the continuous case, the probability density will be $\rho(\alpha) = | c_i |^2$. Since the state of a corpuscle (or a system) is defined by a certain function Ψ, the set of measurable physical quantities will correspond to a set of probability distributions that the theory presently considers (perhaps mistakenly, as well will see) to be intervening with exactly the same status for the corpuscle (or system) in the state Ψ.

One can then define a "dispersion" for every probability distribution that is equal to the square root of the mean square of the distance from the mean value. One thus sets this distribution to be:

$$\sigma(A) = \sqrt{\overline{(\alpha - \bar{\alpha})^2}} = \sqrt{\overline{\alpha^2} - \bar{\alpha}^2} \, . \tag{7}$$

On can then prove that one will have:

$$\sigma(A) \, \sigma(B) \geq \tfrac{1}{2} \left| \overline{AB - BA} \right| \tag{8}$$

for two quantities A and B.

If the operators A and B commute then the right-hand side of (8) will be zero, which one can interpret by saying that one can get precise values (so the dispersions will be zero) of the quantities A and B by the same measuring device. If the operators A and B do not commute then the right-hand side of (8) will give a non-zero lower limit for the product of the dispersions, in such a way that no measurement operation can provide

precise values for A and B simultaneously. For two canonically-conjugate quantities, one will have $AB - BA = \frac{h}{2\pi i}$, and one will find that:

(8 *bis*) $$\sigma(A)\ \sigma(B) = \frac{h}{4\pi},$$

which will constitute a way of stating the uncertainty relations (6) that is more precise.

Before pursuing the study of the consequences of this formalism, I would like to insist upon something that is extremely abstract: The wave function Ψ shall be considered to be a simple mathematical function that is a complex solution to a partial differential equation that will have – speaking casually – the form of an equation of wave propagation. While casting a pall on the physical considerations that guided me in the beginning of my research and on the ones that were then developed by Schrödinger, one will no longer seek to give any physical picture for the relationships between the wave and the corpuscle. We do not even know whether the wave Ψ is anything but a mathematical expression that will permit the calculation of probabilities, and whether it will remain somewhat obscured from physical reality. On the other hand, simultaneously considering all of the developments of the wave Ψ and giving the same status to all of the probability distributions that one can deduce from it is somewhat strange, since one knows that each of these distributions will be physically significant only *after* one performs the corresponding measurement, and that the measurement will, as we shall see, completely modify the initial state of things. Obviously, one can always say that any physicist that knows Ψ will have the right to appeal to it in order to calculate the values of a physical quantity that represent the possible results of a measurement of that quantity and the corresponding probabilities. However, the probability distributions thus obtained will have only a subjective value, and can take on an objective value only after the effective performance of the measurement, which implies the intervention of an appropriate device. Later on, we will return to these questions, which will remain very obscure in the formalism that we are presently using, and we shall pursue the study of the consequences of that formalism.

2. Reduction of the probability packet. – Measurement plays an essential role in the interpretation of the formalism that was presented above, and which we are presently assuming, even if it does seem a little mysterious. It is what changes the state of our knowledge of the system under study while giving us new information, and as a result, we are obliged to modify briefly the form of the wave Ψ that represents our knowledge of the corpuscle (or the system). For example, if the measurement is a measurement of position that is more or less precise then the wave train Ψ that is initially associated with the corpuscle will be found to be "reduced" to a less-extended wave train, which can even be almost point-like if the measurement is precise, since the region where the probability of presence $|\ \Psi\ |^2$ is non-zero will have diminished in extent. One thus gets the term "reduction of the probability packet" that Heisenberg recently gave to that modification of Ψ. On the contrary, if the measurement consists of the determination of one of the

components of the quantity of motion p_x then it will be in the space of momenta that reduction of the probability packet will take place, since that will then be the extent of the values of p_x that effectively appear in the Fourier representation of Ψ that will be diminished.

The question of the reduction of the probability packet will then pose a difficult problem in the present interpretation, namely: Is it the action of the measuring device that modifies the wave Ψ or is it the knowledge we acquire from the results of the measurement that implies that modification? I do not know if all of the authors who have adopted the present probability interpretation are in accord on the answer to that question.

Some of them (and that will probably be the case for Bohr) will anxious to preserve a certain character of physical reality for the wave Ψ, and to say that it is the action of the measuring apparatus on the wave Ψ that provokes the reduction of the probability packet. Others, who are perhaps being more logical, will say that it is the knowledge of the result of the measurement that necessitates the modification of the wave, since, while the result of the measurement is not known to us, it will be the old predictions of the probabilities that correspond to the original form of Ψ that will remain valid in order for us to make those predictions. However, if one adopts the second opinion then the wave Ψ will only be a *purely subjective* representation of the probabilities, and cannot be a representation of objective reality to any degree. How then can it obey an equation of wave propagation and, despite everything, provide us with a statistical representation that is probably exact of phenomena whose physical reality is not in doubt? This question remains truly obscure; we shall return to it.

The reduction of the wave train Ψ will give rise to a new situation that is characterized by a new form of Ψ, which is a situation that is unpredictable in advance, since only the probabilities of the various possible measurements can be calculated before making an effective measurement. We shall have to demand whether that unpredictability results from a real indeterminacy, as one presently assumes, or, on the contrary, on the value of certain hidden variables, as is suggested by the theory of the double solution, which is a question that has a close relationship with a theory that was stated by von Neumann in his theory of measurement in wave mechanics.

The Heisenberg uncertainty relations show that a device that permits one to simultaneously perform various measurements on a corpuscle cannot simultaneously tell us precisely the values of all of the quantities that characterize the corpuscle. There will therefore be an incomplete maximum knowledge of these quantities that is compatible with the uncertainty relations. Once we have acquired this maximum knowledge, we can construct the wave function that serves to represent our knowledge immediately after the measurement, and upon starting with the initial form of Ψ, we can follow its ultimate evolution in the course of time with the aid of the wave equation. At any instant, we can then calculate the probabilities of the results of various measurements that one can perform at that instant. That will be true up to the point at which we know the result of the new measurements, which will modify the state of our knowledge and briefly interrupt the regular evolution of the wave Ψ. The regular evolution of that wave between two measurements – which is an evolution that is ruled by the wave equation – is itself determined entirely by the initial form of Ψ (and possibly by the boundary conditions), since the wave equation is of first order in time. The evolution of Ψ between two measurements will be determined, but not the observable phenomena, since the

knowledge of the wave function will give only probabilities for them. If the description of physical reality by the function Ψ is a *complete* description – i.e., if there exist no description that is more complete, for example, by the introduction of hidden variables – then physical phenomena will be undetermined.

3. Destruction of phases by measurements. Interference of probabilities. – The A measurement will introduce a discontinuity into the evolution of the wave function. The knowledge of it *after* the measurement does not allow one to reconstruct the form that it had *before* the measurement.

Consider a large number of corpuscles (or systems) that are initially found in the same state that is represented by Ψ. Measure a certain quantity A for each of them that has proper functions φ_j and proper values α_j . After these measurements, the proportion of corpuscles (or systems) for which one will have found the various values α_i for A will give us the squares of the moduli of the coefficients c_i in the development $\Psi = \sum_i c_i \varphi_i$ of the wave function *before* the measurement. The knowledge of Ψ for all of the corpuscles (or systems) after the measurement will then provide us with the values of the $| c_i |$, but in order to know the c_i themselves, we would need to know their arguments, and thus, the relative phases of the components $c_i \varphi_i$ of the initial wave function.

It was that remark that led Bohr to emphasize that any measurement must have the effect of completely destroying the phases. It is this destruction of the phases by the act of measurement that brings about what constitutes a break in the evolution of Ψ. Indeed, the differences in phase between the components of the development $\sum_i c_i \varphi_i$ are of paramount importance, and any knowledge that relates to the wave function that does not involve knowledge of these phase differences will be radically incomplete. The importance of these phases is clearly manifested to us in the study of interference phenomena for the probabilities.

Consider two quantities A and B whose operators do not commute, and which, in turn, are not simultaneously measurable. The proper values and proper functions of A are α_i and φ_i, while those of B are β_k and χ_k . One easily proves that since A and B do not commute, the system of the φ_i cannot coincide with that of the χ_k . Meanwhile, since the χ_k define a complete system, each φ_i can be expressed with the aid of the χ_k in the form:

(9) $$\varphi_i = \sum_k s_{ik} \chi_k \ ,$$

in which the s_{ik} are elements of a unitary matrix $\mathcal{S}$. More than one term in the right-hand side will appear in this development, since the system of φ_i and that of χ_k do not coincide. Suppose then that the state of the corpuscle (or system) being examined is represented by the wave function:

(10) $$\Psi = \sum_i c_i \varphi_i = \sum_{i,k} c_i s_{ik} \chi_k \ .$$

If one then measures the quantity A then one will find one of the proper values α_i, where the probability of finding α_j will be $|\, c_j \,|^2$ *a priori.* After the measurement, the corpuscle (or system) will be found in the state φ_j, and in that new state a measurement of B will lead to the value β_k with the probability $|\, s_{ik} \,|^2$. Therefore, the probability of finding the value β_k for B by first measuring A and then B will be equal to $\sum_i |\, c_i \,|^2 |\, s_{ik} \,|^2$.

However, now suppose that we have performed the measurement of B directly on the initial state. Then, from the form of the last expression in (10), the general principle that relates to the probabilities of the results of measurement will tell us that the probability of finding β_k will be equal to $\left| \sum_i c_i s_{ik} \right|^2$. That expression will be entirely different from the preceding one, because it will depend upon the phases (or arguments) of the c_i and s_{ik}, while $\sum_i |\, c_i \,|^2 |\, s_{ik} \,|^2$ obviously does not. That is what one calls the "interference of probabilities."

We illustrate this with a simple example: Take a one-dimensional domain with length L. The normalized proper functions of the quantity of motion will be $\varphi_i = \frac{1}{\sqrt{L}} e^{-\frac{2\pi i}{h} p_i x}$ in that domain. Then, let:

(11) $$\Psi = \sum_i \frac{c_i}{\sqrt{L}} e^{-\frac{2\pi i}{h} p_i x} \qquad \left(\sum_i |\, c_i \,|^2 = 1 \right)$$

be the wave function of the corpuscle in its initial state. If one first measures p, and then x then the probability of the position $x = x_0$ will be:

$$\sum_i |\, c_i \,|^2 \left| \frac{1}{\sqrt{L}} e^{-\frac{2\pi i}{h} p_i x_0} \right|^2,$$

or simply $1 / L$, which will imply the equal probability of all positions on the segment of length L.

However, if, on the contrary, one measures the coordinate x in the initial state directly then the probability of the value $x = x_0$ will be $|\, \Psi(x_0) \,|^2$, and this will involve the interference of the plane waves whose superposition will constitute the Ψ, a result that is necessary in order to account for interference in optics and the diffraction of electrons. One will then see that the interference of probabilities, whose existence is necessary for the interpretation of experimental facts, will depend essentially upon phases, whose role is then seen to be paramount.

The fact that the probability of the value β_k of B, when measured directly in the initial state, will be $\left| \sum_i c_i s_{ik} \right|^2$, and not $\sum_i |\, c_i \,|^2 |\, s_{ik} \,|^2$, can seem, on first glance, to be contrary to the theorem of composed probabilities, but in reality, that is not so: The probability

$\sum_i |c_i|^2 |s_{ik}|^2$ is indeed the one that one must choose when one first makes a determination of *A*, and then *B*, since it is equal to the sum of the products of the probilities for *first* getting a value α_i for *A* times the probability of *then* getting the value β_k for *B*. The theorem of composed probabilities is then safe, and if one envisions the probabilities from a purely subjective viewpoint then one can say that there is no reason for the probability $\sum_i |c_i|^2 |s_{ik}|^2$ to be equal to that of directly obtaining the value β_k of *B* by a measurement of that quantity in the initial state. However, if one analyzes that idea closely then one will see that all of the probability distributions that are introduced in the usual theory (except, without a doubt, $|\Psi|^2$) will exist in the initial state only subjectively for the physicist who must make the predictions on the result of possible measurements. These distributions will exist objectively only after the corresponding measurement has been performed when one further ignores the result of that measurement. It is that situation that will explain, later on, why the schema of the usual probabilistic interpretation of wave mechanics is not in agreement with the usual schema that is assumed by statisticians.

4. Divergence between the statistical schema of wave mechanics and the usual schema of statisticians. – In the usual schema of statisticians (which we will present by assuming that one is dealing with continuous variables), one defines a probability density $\rho_X(x)$ for every random variable *X* such that $\rho_X(x)\,dx$ will be the probability for *X* to have a value between *x* and $x + dx$. One will likewise define $\rho_Y(y)$ for another continuous random variable *Y*.

One then defines a density $\rho(x, y)$ such that $\rho(x, y)\,dx\,dy$ is the probability of obtaining values for *X* and *Y* by the same measurement operation (the statisticians often say "by the same *proof*") that are contained in the intervals $x \to x + dx$ and $y \to y + dy$, respectively. That definition will seem quite natural if one adopts a concrete image of the probability in which "individuals" appear, for each of which the quantities *X* and *Y* will have a well-defined value, so statistics will be introduced by the simultaneous consideration of a very large number of individuals for which *X* and *Y* have different values.

Outside of $\rho_X(x)$, $\rho_Y(y)$, and $\rho(x, y)$, statisticians will also consider the probability density $\rho_Y^{(X)}(x, y)$ of *Y*, when coupled to *X*, which will correspond to the probability of obtaining the value *y* of *Y when one knows that X has the value x*, and one likewise defines the probability of *X*, when coupled to *Y*, with the aid of $\rho_X^{(Y)}(x, y)$.

One must now have the following relations, which one can consider to be obvious, between the five probability densities that we just defined:

$$(12)\qquad \begin{cases} \rho_X(x) = \int \rho(x, y)\,dy, \quad \rho_Y(y) = \int \rho(x, y)\,dx, \\ \rho_X^{(Y)}(x, y) = \dfrac{\rho(x, y)}{\rho_Y(y)}, \quad \rho_Y^{(X)}(x, y) = \dfrac{\rho(x, y)}{\rho_X(x)}, \end{cases}$$

from which, one will infer that:

$$(13) \qquad \rho_X(x) = \int \rho_X^{(Y)}(x,y)\rho_Y(y)\,dy\,, \qquad \rho_Y(y) = \int \rho_Y^{(X)}(x,y)\rho_X(x)\,dx\,.$$

Now, the essential fact is that the preceding schema, which is usually taken for granted by statisticians, is not applicable to the probability distributions that are envisioned in the present interpretation of wave mechanics. Indeed, it is, in general, impossible to define the density $\rho(x, y)$ for two measurable quantities, since it is, in general, impossible to simultaneously measure the values of the quantities *X* and *Y*. Formulas (12) no longer make sense then. Without a doubt, it is always possible to define the densities $\rho_X(x)$, $\rho_Y(y)$, $\rho_X^{(Y)}(x,y)$, and $\rho_Y^{(X)}(x,y)$, but they will no longer be related by formulas (12) and (13).

As an example of this, recall the preceding case that was examined of two measurable quantities *A* and *B* that are not commutative, and rewrite formulas (9) and (10) by passing from the discontinuous case to the continuous case. We will have:

$$(14) \qquad \varphi(\alpha) = \int s(\alpha,\beta)\,\chi(\beta)\,d\beta, \qquad \chi(\beta) = \int s^{-1}(\alpha,\beta)\,\varphi(\alpha)\,d\alpha\,.$$

If Ψ is of the form:

$$(15) \qquad \Psi = \int c(\alpha)\,s(\alpha,\beta)\,d\alpha = \iint c(\alpha)\,s(\alpha,\beta)\,\chi(\beta)\,d\alpha$$

then one will find that:

$$(16) \qquad \rho_A(\alpha) = |\,c(\alpha)\,|^2, \qquad \rho_B(\beta) = \left|\int c(\alpha)\,s(\alpha,\beta)\,d\alpha\right|^2,$$

where the second formula expresses the interference of probabilities, so:

$$(17) \qquad \rho_B^{(A)}(\alpha,\beta) = |\,c(\alpha)\,|^2\,|\,s(\alpha,\beta)\,|^2, \qquad \rho_A^{(B)}(\alpha,\beta) = \left|\int c(\gamma)\,s(\gamma,\beta)\,d\gamma\right|^2 \left|s^{-1}(\alpha,\beta)\right|^2,$$

but here the products $\rho_B(\beta)\,\rho_A^{(B)}(\alpha,\beta)$, and $\rho_A(\alpha)\,\rho_B^{(A)}(\alpha,\beta)$ have no reason to be equal, which will indeed show that the non-existence of the density $\rho(\alpha, \beta)$, which must be equal to their common value.

Where does this very strange special character of the statistical distributions of modern quantum mechanics come from? The answer seems to be contained in the essential role that is played by measurement. Since the probability distributions of modern quantum mechanics (with the possible exception of some of them) do not constitute objective probabilities, they can be regarded as all corresponding to a collection of individuals at the same instant for which the quantities will have well-defined values. The implicit hypothesis that makes the relations (12) and (13) "obvious" for the statistician is not realized here.

It is only *after* the action of the measuring device of a quantity for the corpuscle (or system) that the probability distribution can be considered to be realized objectively. To speak more precisely, if one imagines that the measurement of a certain quantity is

performed simultaneously on an infinitude of corpuscles (or systems) that initially have the same function Ψ then it will be only after performing the measurement on all of these corpuscles (or systems) that one will really have a collection of individuals that each possess a precise value of the measured value such that these values will be distributed according to the law of probability in $| c_k |^2$, and it can be further remarked that the law of probability in $| c_k |^2$ will thus not be found to be realized objectively by a collection, so much as for the measured quantity and the ones that commute with it, to the exclusion of the other ones. If the physicist knows the wave function in the initial state, when no measured has been performed, then he can calculate the various probability distributions that he can subsequently decide to measure. However, each of these distributions can be found to be thus realized, and thus correspond to a collection, only after performing the corresponding measurement. The distributions can never be all found to be realized simultaneously, since one cannot simultaneously measure all of the quantities, and one must employ two incompatible measuring devices in order to measure two non-commutating quantities.

Certainly, the physicist always has the right to simultaneously consider the set of probability distributions before any measurement that can be deduced from the various developments of the initial Ψ, but these probabilities will then have a subjective character, and are not objective probabilities that are statistically realized by the same collection of individuals. As we have seen, that is what prevents us from attributing the properties (12) and (13) – which will be obvious for objective distributions that refer to a collection of individuals with well-defined characteristics – to the probability distributions of conventional wave mechanics. We think that it is for the same reason that the celebrated theorem of von Neumann, which we will discuss soon, is basically only a truism, and does not at all prove the impossibility of re-establishing determinism in wave mechanics by the introduction of hidden variables.

———————

CHAPTER II.

THE THEORY OF MEASUREMENT, ACCORDING TO VON NEUMANN ([1]).

1. Pure case and mixture. – First, recall some considerations regarding the interference of probabilities. Let there be a very large number $\mathcal{N}$ of corpuscles (or systems) that all have the same wave function Ψ. If A is a measureable physical quantity with proper values α_k and proper functions φ_k then if one has $\Psi = \sum_k c_k \varphi_k$ then the measurement of A must lead one to find the value α_1 for $| c_1 |^2 \mathcal{N}$ systems, the value α_2 for $| c_2 |^2 \mathcal{N}$ systems, etc. The mean value of A will be $\sum_k | c_k |^2 \alpha_k$.

Now, imagine that instead of having $\mathcal{N}$ systems in the same state, we have $| c_1 |^2 \mathcal{N}$ systems in the state φ_1, $| c_2 |^2 \mathcal{N}$ systems in the state φ_2 , etc. The measurement of A will then give us the same statistical results as in the former case. One might then believe that the two cases are equivalent, but we shall see that this is not true.

Indeed, consider a measurable physical quantity B that does not commute with A. The proper functions of B will not coincide with those of A, and if β_k and χ_k are the proper values and proper functions of B, resp., then one will have $\varphi_k = \sum_l d_{kl} \chi_l$, in which the development will generally contain several terms. First, imagine the previous case, in which we had $\mathcal{N}$ systems that were all in the same state:

$$\Psi = \sum_k c_k \varphi_k = \sum_{k,l} c_k d_{kl} \chi_l \, .$$

The measurement of B for all of these systems will then give $\mathcal{N} \left| \sum_{k,l} c_k d_{kl} \right|^2$ times the value β_l, and the mean value of B will be:

(1) $$B = \sum_l \left| \sum_k c_k d_{kl} \right|^2 \beta_l = \int \Psi^* B \Psi \, d\tau = \sum_{k,l} c_k^* c_l B_{kl}^{(\varphi)} \, ,$$

with:

$$B_{kl}^{(\varphi)} = \int \varphi_k^* B \varphi_l \, d\tau \, .$$

([1]) *See* bibliography [**1**], [**2**].

We then place ourselves in the second case, where we had $|c_1|^2\mathcal{N}$ systems in the state φ_1. The measurement of A on the first $|c_1|^2\mathcal{N}$ systems will give the value β_l for a proportion of these systems that is equal to $|d_{kl}|^2$, etc. In total, the value of β_l of B will be obtained:

$$\mathcal{N}\sum_k |c_k|^2 |d_{kl}|^2$$

times, and in turn, the mean value of B will be:

$$B = \sum_{k,l} |c_k|^2 |d_{kl}|^2 \beta_l = \sum_k |c_k|^2 B_{kl}^{(\varphi)}, \tag{2}$$

with:

$$B_{kl}^{(\varphi)} = \int \varphi_k^* B \varphi_l \, d\tau .$$

One then sees that the two cases that we envisioned are completely different for any quantity that does not commute with A. In the first one there is interference of probabilities, while in the second one that interference is not present. One cannot therefore consider the $\mathcal{N}$ systems as defining a collective system that is composed of $\mathcal{N}|c_1|^2$ individual systems that have the value α_1 for A, etc. Moreover, it is obvious that it will therefore be entirely legitimate to consider the $\mathcal{N}$ systems as defining a collective system that is composed of $\mathcal{N}|d_1|^2$ systems that have the value β_1 for B, etc., with $d_1 = \sum_k c_k d_{k1}$, and this second collective system will not coincide with the first one. We can thus not consider the set of $\mathcal{N}$ systems as defining a well-defined collective system, since that collective system will vary according to the quantity that is envisioned. We then recover the idea that we previously brought to light: The probabilities that are envisioned in conventional wave mechanics correspond to a unique collective system that is realized in the state Ψ. In order to distinguish the case where the probability distribution for a quantity A has only a subjective value before the measurement from the one where that distribution is realized after the measurement, von Neumann said that the former case constitutes a "pure" case, while the latter one constitutes a "mixture."

Without making any act of measurement intervene, one can imagine $\mathcal{N}_1$ systems that have a wave function $\Psi^{(1)}$, $\mathcal{N}_2$ systems that have a wave function $\Psi^{(1)}$, etc. The set of all $\mathcal{N}$ systems will then define a "mixture" of $\mathcal{N}_1$ pure cases that correspond to $\Psi^{(1)}$, $\mathcal{N}_2$ pure cases that correspond to $\Psi^{(2)}$, etc. We recover the second case that was studied at the beginning of this paragraph by taking $\mathcal{N}_1 = \mathcal{N}|c_1|^2$, … If we set $\mathcal{N}_i / \mathcal{N} = p_i$ then we will have a "mixture" that is defined by the set of "statistical weights" p_k with $\sum_k p_k = 1$.

If we set $c_k = \sqrt{p_k} e^{i\alpha_k}$ then we will see that the $p_k = |c_k|^2$ are the statistical weights of the mixture that is equivalent to the pure case Ψ, *as far as the measurement of A is concerned*. However, this mixture is realized only after the measurement that

transformed the initial pure case into this mixture. The mixture that is equivalent to the pure case Ψ for the measurement of a quantity B that does not commute with A will involve statistical weights that are different from the preceding ones, and will be realized only by a measurement that involves a device of a different type. That is why one cannot reduce a pure case to a well-defined mixture.

We have seen that the mean value of B is given by formula (1) for the pure case Ψ. If one replaces this pure case with a mixture that is found to be realized by the measurement of A then the mean value of B will be given by formula (2). It is easy to specify the manner in which the two expressions (1) and (2) differ. Formula (1) can be written:

$$\overline{B} = \sum_{k,l} |c_k| \cdot |c_l| e^{i(\alpha_l - \alpha_k)} B_{kl}^{(\varphi)} . \tag{3}$$

If one assumes that the phases α_k (i.e., the arguments of the α_k) are known completely with equal probabilities for their possible values then the mean value of the expression (3) will be obtained by taking a mean over the values of the α_k, which are all assumed to be equally probable. The terms where $k \neq l$ will give zero, and we will recover the expression (2). In other words, one passes from the pure case Ψ to the mixture that is realized by the measurement of A by assuming that this measurement has made one lose all knowledge of the phases α_k . Here, we indeed recover the conclusion that the measurement of A that is performed on the initial state that is represented by $\Psi = \sum_k c_k \varphi_k$ will have the effect of completely destroying the phase difference that exist between the components φ_k of the initial Ψ.

Finally, we have obtained a neat idea of the difference between a "pure case" that is defined by a wave function Ψ and a "mixture" that is defined by a set of pure cases with wave functions Ψ_1, Ψ_2, … that are affected with statistical weights p_1, p_2, …

2. The statistical matrix of J. von Neumann for the pure case. – First, envision a pure case that is defined by a wave function of a given form. That function can be considered to be a vector in a Hilbert space. If φ_1, φ_2, …, φ_n, … is a complete, orthonormal system of basis functions (for example, the proper functions of a Hermitian, linear operator A) then the φ_i can be considered to define a complete system of unitary vectors in Hilbert space, and the expression $\Psi = \sum_k c_k \varphi_k$ will be analogous to the expression of a vector with the aid of its components along orthogonal directions that are defined by the unitary vectors. One can say that the c_k are the components of Ψ in the basis system of the φ_k . The Hilbert space that we consider will be a complex space, and the components c_k will be complex, in general.

Now, let:

$$\Psi = \sum_k c_k \varphi_k \quad \text{and} \quad \chi = \sum_k d_k \varphi_k$$

be two vectors in Hilbert space. By definition, their scalar product is (D being the domain variation of the variables in the φ):

(4) $$(\Psi \cdot \chi) = \int_D \Psi^* \chi \, d\tau = \sum_{k,l} c_k^* d_l \int \varphi_k^* \varphi_l \, d\tau = \sum_{k,l} c_k^* \, d_l \, \delta_{kl} = \sum_{k,l} c_k^* \, d_k \ ,$$

and one will have:

(5) $$(\chi \cdot \Psi) = (\Psi \cdot \chi)^* ;$$

one will indeed then have the generalization of the classical expression for the scalar product to complex vectors.

The scalar product of a vector Ψ with itself, which is analogous to the square of the length of an ordinary vector, is called the "norm" of that vector, and will have the value:

(6) $$N(\Psi) = (\Psi, \Psi) = \int_D | \Psi |^2 \, d\tau = \sum_k | c_k |^2 \ .$$

If the vector is normalized then one will have:

$$N(\Psi) = 1 \qquad \text{and} \qquad \sum_k | c_k |^2 = 1.$$

An operator on Hilbert space will correspond to an operator that makes one vector go to another one $\chi = A\Psi$, which will then define the operation that takes Ψ to χ, and one will have:

$$\sum_l d_l \varphi_l = A \sum_k c_k \varphi_k \ ,$$

so, upon multiplying by φ_j^* and integrating over D, one will get:

(8) $$d_j = \sum_k c_k \int_D \varphi_k^* A \varphi_k \, d\tau = \sum_k a_{jk} c_k \ .$$

The a_{jk}, which are elements of the matrix that is generated by A in the system of the φ_k, will then be the coefficients of the linear transformation that takes the components of Ψ to those of χ. The conservation of norm would impose the condition that the matrix a must be unitary.

If Ψ is once again the wave function of a "pure case" then imagine the operation on Hilbert space of "projecting onto the vector Ψ;" let P_Ψ be the corresponding operator. It is obvious that $P_\Psi^2 = P_\Psi$, and that, more generally, $P_\Psi^n = P_\Psi$. Since all of the powers of P are identical, one says that this operator is "idempotent."

Now, let there be a complete system of orthonormal basis functions $\varphi_1, \ldots, \varphi_n, \ldots$ We have a development for Ψ:

$$\Psi = \sum_k c_k \varphi_k \ , \qquad \text{with} \quad c_k = \int_D \varphi_k^* \Psi \, d\tau \qquad \text{and} \qquad \sum_k | c_k |^2 = 1.$$

One can obviously find an infinitude of orthonormal basis systems for which Ψ is one of the basis vectors. In one of these systems, the function φ_k will have a development of the form:

(9) $$\varphi_k = d\Psi + \ldots, \qquad \text{with} \quad d = \int_D \Psi^* \varphi_k \, d\tau = c_k^* .$$

The operator P_Ψ, which is the "projector" onto Ψ, is defined by:

(10) $$P_\Psi \, \varphi_k = d\Psi = c_k^* \Psi$$

for any φ_k. The matrix that is generated by the operator P_Ψ in the basis system of φ_k has an element with indices m, n:

(11) $$(P_\Psi)_{mn} = \int_D \varphi_m^* P_\Psi \varphi_n \, d\tau = c_n^* \int_D \varphi_m^* \Psi \, d\tau = c_m \; c_n^* .$$

Thus, the matrix P_Ψ that is attached to the pure case being considered is expressed with the aid of the coefficient of the development of Ψ in the basis system that is being utilized. One has thus defined what von Neumann called the "statistical matrix" that is attached to the pure case Ψ; formula (11) makes it obvious that this matrix is Hermitian.

The statistical matrix possesses two fundamental properties:

1. Its trace is equal to 1. Indeed:

(12) $$\text{Tr } P_\Psi = \sum_n (P_\Psi)_{nn} = \sum_n c_n^* c_n = 1.$$

2. It is idempotent. Indeed, one will have:

(13) $$(P_\Psi^2)_{mn} = \sum_p c_m^* c_p \cdot c_p c_n^* = c_m \; c_n^* = (P_\Psi)_{mn} ,$$

and thus, in terms of matrices, $P_\Psi^2 = P_\Psi$, and by recurrence, $P_\Psi^n = P_\Psi$.

Now, let A be a quantity in the system being considered. If the φ_k are functions of an arbitrary orthonormal basis (which are no longer proper functions of A, here) then we have seen that the mean value of A will be:

(14) $$\overline{A} = \sum_{k,l} c_k^* \, c_l \, A_{kl}^{(\varphi)} ,$$

where the $A_{kl}^{(\varphi)}$ are the elements of the matrix that is generated by the operator A in the system of the φ_k, and c_k is the component of Ψ along φ_k. One can also write:

(15) $$\bar{A} = \sum_{k,l} (P_\Psi)_{lk} A_{kl}^{(\varphi)} = \mathrm{Tr}(P_\Psi A) = \mathrm{Tr}(A P_\Psi).$$

Therefore, the knowledge of the statistical matrix will provide us with a simple means of calculating $\bar{A}$.

The statistical matrix of a pure case is frequently called an "elementary statistical matrix" (*Einzelmatrix*), in contrast to the more general statistical matrices that we shall encounter later on while studying mixtures of pure cases.

An elementary statistical matrix can be easily put into diagonal form. In order to this, it will suffice to take the basis system to be a system where the Ψ considered is one of the basis functions; for example, $\varphi_1 = \Psi$. The elementary statistical matrix will then take the form:

(16) $$\begin{vmatrix} 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdots \end{vmatrix}.$$

All of the terms will be zero, except for the first diagonal term, which is equal to 1; this results from (11) easily. The trace of the statistical matrix will be an invariant under changes of basis functions, and in turn, a known property of unitary transformations; it must then be equal to 1, as the table (16) shows. This table will also permit one to verify immediately that the statistical matrix is idempotent.

3. The statistical matrix for a mixture of pure cases. – We shall now consider a mixture of pure cases. We have already defined such a mixture by considering $\mathcal{N}$ systems, of which, $\mathcal{N}p_1$ are in the state $\Psi^{(1)}$, $\mathcal{N}p_2$ are in the state $\Psi^{(2)}$, …, with $\sum_k p_k = 1$. However, we can also introduce the idea of mixture for just one system. Indeed, it can happen that we are ignorant of the exact form of the wave function of a system, and that we know only that it has a probability p_1 of being in the state $\Psi^{(1)}$, a probability p_2 of being in the state $\Psi^{(2)}$, etc., a probability p_n of being in the state $\Psi^{(n)}$, with $\sum_k p_k = 1$. The state of our knowledge about the system is then represented by a mixture of pure cases with the statistical weights p_k.

Each of the pure cases in the mixture has an elementary statistical matrix $\Psi_{p^{(k)}}$. We attribute a Hermitian statistical matrix:

(17) $$P = \sum_{k=1}^{n} p_k P_{\Psi^{(k)}},$$

with

(18) $$P_{(m)} = \sum_{k=1}^{n} p_k c_l^{(k)} c_m^{(k)}$$

to it, where the statistical weights p_k are positive numbers between 0 and 1 whose sum is equal to 1. The $c_m^{(k)}$ are the components of the various $\Psi^{(k)}$ in the system with basis φ_1, …, φ_n . The statistical matrix (17) thus appears to be a superposition of elementary statistical matrices.

As an example, suppose that one has taken the basis functions to be the proper functions that relate to position $\delta(q - q')$, where δ is the singular Dirac function. The formula:

$$\Psi^{(k)}(q, t) = \int \Psi^{(k)}(q', t)\, \delta(q - q')\, dq' \tag{19}$$

will then show that the $c_i^{(k)}$ are equal $\Psi^{(k)}(q', t)$, and that one will find that the components of the statistical matrix are:

$$P(q', q'') = \sum_{k=1}^{n} p_k \Psi^{(k)}(q') \Psi^{(k)*}(q'') . \tag{20}$$

This is Dirac's statistical matrix.

The mean value of a measurable quantity A of the system will be:

$$\bar{A} = \sum_{k=1}^{n} p_k \bar{A}_{\Psi^{(k)}} , \tag{21}$$

where $\bar{A}_{\Psi^{(k)}}$ is the mean value that A will have when the system is in the pure state $\Psi^{(k)}$. From (15), we will then get:

$$\bar{A} = \sum_{k=1}^{n} p_k \sum_j (P_{\Psi^{(k)}} A)_{jj} = \sum_{k=1}^{n} \left(\sum_j p_k P_{\Psi^{(k)}} A \right)_{jj} . \tag{22}$$

The formula will therefore be the same as it is for the pure case.

The statistical matrix of a mixture, like that of a pure case, will always have a trace that is equal to 1, because:

$$\text{Tr } P = \sum_m P_{mm} = \sum_m \sum_{k=1}^{n} p_k c_m^{(k)} c_m^{(k)*} = \sum_{k=1}^{n} p_k \sum_m | c_m^{(k)} |^2 = 1. \tag{23}$$

By contrast, while the matrix of a pure case is always idempotent, the same thing is not true for the statistical matrix of a mixture. Indeed, one can prove that any idempotent statistical matrix is elementary. In order to do that, one assumes that $P^2 = P$, and one writes P in diagonal form, which is always possible. If p_i is the i^{th} diagonal element of P then the relation $P^2 = P$ will demand that one must have $p_i^2 = p_i$, and the p_i will then be equal to 0 or 1. The equation Tr $P = 1$ that is satisfied by all statistical matrices will then show that one of the p_i is different from 0, and therefore equal to 1. The system will then have a unique Ψ that agrees with one of the basis functions that reduces P to its diagonal

form. Therefore, the necessary and sufficient condition for a statistical matrix to be an idempotent is that it be elementary.

Now, consider the non-elementary statistical matrix of a mixture. If the $\Psi^{(1)}$, $\Psi^{(2)}$, …, $\Psi^{(n)}$ that define the pure cases that appear in the mixture are orthogonal (which can happen only in exceptional cases) then one can take them to be the first n basis function of an orthonormal system. One will then have $c_m^{(k)} = \delta_{km}$, since $\Psi^{(k)}$ will reduce to φ_k and $P_{(m)}$ will be zero for $l \neq m$, while the P_{kk} will be equal to p_k for $k \leq n$ and zero for $k > n$. The statistical matrix will then take the following diagonal form:

$$\begin{vmatrix} p_1 & 0 & 0 & 0 & \cdots \\ 0 & p_2 & 0 & 0 & \cdots \\ . & . & . & . & \cdots \\ 0 & 0 & 0 & p_n & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ . & . & . & . & \cdots \end{vmatrix}. \tag{24}$$

However, this is the one exceptional case. In general, the functions $\Psi^{(1)}$, …, $\Psi^{(n)}$ will not be orthogonal. One can nonetheless reduce the matrix P to diagonal form, even in this case, but the diagonal elements p'_k will no longer be equal to p_1, …, p_n, 0, 0, … Since the matrix P is Hermitian, the p'_k will be real numbers. Moreover, since Tr P = 1, one will have $\sum_k p'_k = 1$. We shall show that the p'_k cannot be negative. In order to do that, if ξ_k are the components of a vector Ξ in Hilbert space then consider the scalar product of Ξ with $P\Xi$. One will then have

$$(\Xi, P\Xi) = \sum_{m,n} \xi_m^* \sum_{k=1}^{n} p_k c_m^{(k)} c_n^{(k)*} \xi_n = \sum_{k=1}^{n} p_k \left|(\Xi, \Psi^{(k)})\right|^2 \tag{25}$$

for its value.

Since the square of a modulus is, *a fortiori*, positive or zero, we will see that the scalar product (25) is necessarily positive or zero. Now, if we put P into its diagonal form then that scalar product will have the following expression:

$$(\Xi, P\Xi) = \sum_m p'_m \left| \xi_m \right|^2, \tag{26}$$

which must be ≥ 0, and for any Ξ. Therefore, the p'_m must all be positive or zero. Since their sum is equal to 1, one will have $0 \leq p'_m \leq 1$. One infers from that that $p'_m - p'^2_m \geq 0$, so for any arbitrary vector Ξ in Hilbert space:

$$(\Xi \cdot (P - P^2)\, \Xi) = \sum_m (p'_m - p'^2_m) \left| \xi_m \right|^2 \geq 0. \tag{27}$$

4. Irreducibility of the pure case. – We now come to a theorem that plays a major role in the proof by which von Neumann wanted to establish the impossibility of explaining the present probabilistic character of wave mechanics with the aid of hidden variables.

The important theorem in question is stated as follows:

It is impossible to represent a pure case in the form of a mixture, or also: *A pure case is never reducible to a superposition of pure cases.*

He thus established the intrinsic special character of the pure cases.

Indeed, if this theorem were not true then it would have to be possible – at least, in some cases – to obtain a relation of the form:

(28) $$P = \sum_i \alpha_i Q_i \,,$$

in which, P and Q_i are elementary statistical matrices – i.e., idempotent Hermitian matrices with trace 1 – and the α_i are positive numbers such that $\sum_i \alpha_i = 1$. Now, one will then have:

(29) $$\begin{aligned} P^2 &= \sum_i \alpha_i^2 Q_i^2 + \sum_{i\neq j} \tfrac{1}{2}\alpha_i\alpha_j (Q_iQ_j + Q_jQ_i)\,, \\ &= \sum_i \alpha_i^2 Q_i^2 + \sum_{i\neq j} \tfrac{1}{2}\alpha_i\alpha_j [Q_i^2 + Q_j^2 - (Q_i - Q_j)^2]\,, \\ &= \sum_i \left[\alpha_i^2 + \alpha_i \sum_{j\neq i}\alpha_j\right] Q_i^2 - \sum_{i>j}\alpha_i\alpha_j (Q_i - Q_j)^2\,, \\ &= \sum_i \alpha_i Q_i^2 - \sum_{i>j}\alpha_i\alpha_j (Q_i - Q_j)^2\,, \end{aligned}$$

because $\sum_{j\neq i}\alpha_i = 1 - \alpha_i$. One will therefore have:

(30) $$P^2 - P = \sum_i \alpha_i (Q_i^2 - Q_i) - \sum_{i>j}\alpha_i\alpha_j (Q_i - Q_j)^2 \,.$$

However, $P^2 = P$ and $Q_i^2 = Q_i$, so:

(31) $$\sum_{i>j}\alpha_i\alpha_j (Q_i - Q_j)^2 = 0,$$

and since all of the α_i are positive:

(32) $$(Q_i - Q_j)^2 = 0.$$

Now, the square of a Hermitian matrix can be zero only if the matrix itself is zero. Indeed, if A is a Hermitian matrix then the elements of A^2 will be:

$$(a^2)_{ik} = \sum_l a_{il}\, a_{lk} = \sum_l a_{il}\, a^*_{kl}\ ,$$

and if the $(a^2)_{ii}$ are zero then one must also have $\sum_i \left| a_{il} \right|^2 = 0$, which will demand that $a_{il} = 0$, and in turn, that $A = 0$.

Since $Q_i - Q_j$ is a Hermitian matrix then the condition (32) will imply that $Q_i = Q_j$. All of the Q_i will be the same, and one will have:

$$P = \sum_i \alpha_i\, Q_i = Q_i\,, \qquad \text{since} \quad \sum_i \alpha_i = 1.$$

P will not be truly a sum of elementary statistical matrices then, which is contrary to hypothesis.

It is therefore indeed proved that the pure cases are irreducible and can never be reduced to a mixture of pure cases. The pure cases of wave mechanics will thus possess the following two properties:

1. They will be represented by elementary (i.e., idempotent) statistical matrices, while any mixture will have a matrix that is not elementary (i.e., not idempotent).
2. There will be no way of reducing a mixture to a pure case.

5. The statistical laws of quantum mechanics will be impossible to interpret by the introduction of hidden variables. – In classical physics, any time one must introduce probabilities in place of rigorous laws one will always assume that there exists determinism in the phenomena, but that this determinism is too complicated or too subtle for us to be able to follow it in detail, since the observable manifestations are of a statistical character and, for that reason, they will be expressed by probabilities. The laws of probability and the element of chance that they seem to introduce will not be the proof of a true contingency, but the result of our incapacity to follow a determinism that is too fine-grained or too complicated. That is the definition of chance that one finds in the writings of all thinkers who predated the development of wave mechanics, and in particular, in the works of Henri Poincaré.

The best-known example of such a pseudo-statistical theory in physics is the kinetic theory of gases. There, one will assume that the motions of the gas molecules, as well as their mutual collisions, are governed by the rigorous laws of classical mechanics, in such a way that there will be a subordinate determinism. However, the molecules are sufficiently numerous that their motions will be so complicated that we cannot actually follow this elementary determinism in all of its details. Moreover, the molecular motions complete elude our senses, and we can only predict the macroscopic effects of these motions, such as pressure, temperature, local fluctuations of density or energy, the Brownian agitation of a visible granule due to its irregular collisions with molecules, etc. Since these macroscopic phenomena will result from an enormous number of complicated, elementary phenomena, we seem to be constructing a statistical theory that will involve only probabilities, but that introduction of chance is only apparent, and, for example, the disorganized motions of a granule in its Brownian agitation will seem to us

to be ruled by a rigorous determinism if we know how to calculate all of the molecular motions of the ambient gas and their collisions with the granule.

Since that elimination of chance, to the profit of a subordinate determinism, has succeeded in classical physics, one might attempt to introduce it into quantum physics. We have found the laws of probability in wave mechanics. Can we not assume that this results from our ignorance of some hidden determinism? If one succeeds in this enterprise then one will have once more eliminated indeterminism and maintained the classical concept of chance. On the contrary, if one fails then one must abandon determinism and assume an absolute contingency of microphysical phenomena. To employ the language of von Neumann: In the latter case, wave mechanics will be a "truly statistical" theory. Now, von Neumann thought that he could resolve the question by proving a theorem, from which, he thought that he could deduce the impossibility of deducing the laws of probability in wave mechanics from some hidden determinism.

In order to establish its proof, von Neumann started with the following remarks:

To assume a subordinate determinism is to assume the existence of variables whose exact values we ignore (viz., hidden parameters), such as, for example, the positions and velocities of the molecules of a gas and the probabilities of then introducing them as a result of our ignorance of these hidden parameters. In a deterministic theory with hidden parameters, the real state of a gas, for example, is entirely determined at each instant. All of the molecules of the gas will have well-defined positions and velocities, and if we know all of these parameters then we can represent the state of the gas by a point in the extension-in-phase. However, we ignore the exact value of the hidden parameters, and in order to represent the global statistical picture, which is all that is accessible to our sense, we envision a "mixture" of elementary states with conveniently-chosen statistical weights. The elementary states will define a mixture that will correspond to well-determined values of all the quantities. They will therefore be indecomposable and also "dispersionless," because any quantity *A* that has a well-defined value will be equal to its mean value, and the dispersion $\sigma = \sqrt{\overline{A^2} - (\overline{A})^2}$ will be zero, as well as all of the differences $\overline{A^n} - (\overline{A})^n$, moreover.

In other words, any deterministic statistical theory with hidden parameters will introduce a collective system of individuals, for which all quantities that characterize them will have well-determined values, and will be, in turn, exempt from dispersion. The dispersions will appear for the collective system only as an ensemble. Under these conditions, the probability distributions that are valid for the collective system must satisfy the usual schema of statisticians that we have studied previously. Now, we know that this is not the case for the probability distributions of conventional wave mechanics, and that one already believes that one is authorized to deduce that wave mechanics cannot be interpreted by a deterministic theory with hidden parameters.

Von Neumann recovered this result by the following path: He started with the remark that a statistical theory can reduce to a deterministic schema with hidden parameters only if the probability distributions that appear in that theory can all be reduced to indecomposable mixtures of elementary states with no dispersion. He then proved that this is not the case for the distributions that are envisioned in wave mechanics by appealing to the following theorem:

The states that one encounters in wave mechanics are never dispersionless.

In other words, one cannot have $\overline{A^2} = (\bar{A})^2$ for any measurable quantity in any realizable state.

In reality, von Neumann's proof – which is beautiful, but somewhat cumbersome – tells us nothing new. Since one knows the uncertainty relations in the form $\sigma(x)\ \sigma(p_x) \geq h\ 4\pi$, one already knows that no quantity can be dispersionless for the probability distributions of conventional wave mechanics.

However, follow von Neumann's argument: We have seen that any state (where it is a pure state or a mixture) is characterized by a Hermitian statistical matrix with a trace that is equal to 1, such that the mean value of any quantity in this state will be given by the formula (15). Therefore, in order to a state to be dispersionless, it will be necessary that one have:

(33) $$\mathrm{Tr}\,(PA)^2 = [\mathrm{Tr}(PA)]^2$$

for any quantity A.

Now, let φ_1, φ_2, …, φ_i, … be a complete system of orthonormal basis functions. Consider the operator on Hilbert space that projects any vector in that space onto the vector φ_i. That projector is a Hermitian linear operator P_{φ_i}, and we can take $A = P_{\varphi_i}$. If the state were dispersionless then one would have, in particular:

(34) $$\mathrm{Tr}(PP_{\varphi_i}^2) = [\mathrm{Tr}(PP_{\varphi_i})]^2 .$$

However, since one has $P_{\varphi_i}^2 = P_{\varphi_i}$, one will get:

(35) $$\mathrm{Tr}(PP_{\varphi_i}) = [\mathrm{Tr}(PP_{\varphi_i})]^2 .$$

Now:

(36) $$\mathrm{Tr}(PP_{\varphi_i}) = \sum_k (PP_{\varphi_i})_{kk} = \sum_k \int_D \varphi_k^* PP_{\varphi_i}\varphi_k \, d\tau ,$$

and since $P_{\varphi_i}\varphi_k = \delta_{ik}\,\varphi_i$, one will finally get:

(37) $$\mathrm{Tr}(PP_{\varphi_i}) = \sum_k \int_D \varphi_k^* P\varphi_i \, d\tau\, \delta_{ik} = \int_D \varphi_i^* P\varphi_i \, d\tau = P_{ii} .$$

This trace must be equal to its square, so either $P_{ii} = 1$ or $P_{ii} = 0$, and this must be true for any index i, because we can reason the same way with all of the P_{φ_i}. However, one can assume that certain P_{ii} will be equal to 1 and others, to 0, and in order to satisfy the relation $\sum_i P_{ii} = 1$, one must then have all of the P_{ii} are zero, except for one of them. However, the latter hypothesis can be rejected because we can vary the orthonormal basis system in Hilbert space in a continuous fashion by an operation that corresponds to a rotation of the axes in that function space. We can then make each of the original axes coincide successively with the other axes, and by a continuous operation. Each of the P_{ii}

must vary in a continuous fashion in the course of that operation, and since they can take only the values 0 or 1, they must keep their initial values. Therefore, either the P_{ii} will all be equal to 1 or they will all be equal to 0. Now, neither of these two hypotheses is compatible with $\sum_i P_{ii} = 1$,, because $\sum_i P_{ii}$ will be infinite in one case and zero in the other.

Finally, there can exist no acceptable statistical matrix P that corresponds to an absence of dispersion for all of the quantities, and von Neumann concluded from this that it will be impossible to account for the probability distributions of wave mechanics by a hidden determinism.

6. Critique of the preceding conclusion. – After having reflected upon this for some time, I now think that von Neumann's proof does not carry the weight that is attributed to it. It indeed shows that the probability distributions of conventional wave mechanics are never completely dispersionless, so they cannot, in turn, correspond to any collection of individuals with well-determined properties. However, we have already confirmed that, and that result is, moreover, contained in the uncertainty relations.

As for deducing the impossibility of interpreting wave mechanics by a deterministic schema of hidden variables, that is another matter. We have already said that the probability distributions in $| c_k |^2$ are generally realized only *after* performing the corresponding measurement. Since the measuring devices for the various magnitudes are generally incompatible, one will have no reason, *a priori*, to expect that all of the probability distributions in $| c_k |^2$ will correspond to the same collective system, and in fact, they do not. However, nothing prevents one from imagining that by introducing hidden variables, one can define probability distributions, *which are also hidden*, that will correspond to a unique collective system in the initial state, before performing any measurement, and will permit one to obtain a deterministic schema. These probability distributions that exist in the initial state before any measurement will remain hidden, because, in general, performing the measurement of the quantity A, while acting on the system under study, will make that initial probability distribution disappear and will make the one that one habitually considers appear. We will confirm that this perfectly-admissible hypothesis is the one that corresponds to the causal interpretation of wave mechanics with the aid of the ideas of the double solution – or pilot-wave – and we will study them in detail later on. In the final analysis, von Neumann's celebrated theorem seems to me to not have the significance that has been attributed to it, and does not seem to me to constitute an insurmountable obstacle to the edification of a deterministic interpretation of wave mechanics that introduces hidden variables.

———

CHAPTER III

THE THEORY OF MEASUREMENT, ACCORDING TO VON NEUMANN (*CONT.*).

---

1. Generalities on measurement. – As we have seen, measurement plays an essential role in quantum physics. Indeed, its role in microphysics is completely different from its role in classical macroscopic physics. In classical physics, a measurement, at least, when it is performed with suitable precautions, is a simple "verification" that specifies our knowledge of objective reality without disturbing it appreciably. Since real elementary states are assumed to be determined perfectly, any ignorance on our part will translate into probabilities that pertain to a mixture with convenient statistical weights for the various elementary states, and measurements will be considered to be capable of diminishing our ignorance, or even suppressing it, by making a perfectly-determined elementary state known to us (for example, a corpuscle with exact values for the quantities of position and quantity of motion that characterize it).

The probability distributions that present themselves in classical physics will therefore always have the character of a mixture of elementary states in which all of the quantities have well-defined values (i.e., collective systems that consist of individual entities whose characteristics all have well-defined values). A measurement is then assumed to make us aware of the real value of a quantity that exists objectively before the measurement, and if that measurement is indeed performed, without modifying it appreciably.

Everything is completely different in quantum theory. There, the maximum knowledge that we can have about a system will be realized when we can consider it to be a pure case – i.e., to attribute a well-defined function Ψ to it. In that state of maximum knowledge, it will be impossible for us to specify the value of any of the quantities in the system, since no experimental device can give us all of them at once. The pure case Ψ corresponds to probability distributions (that will be, in principle, realized only after performing the corresponding measurement) that will involve non-zero dispersions for certain quantities. If a new measurement is performed then that can give us only give us, at best, a new pure case that also involves non-zero dispersions for certain quantities. It will augment our knowledge of some quantities, but only in such a fashion that our optimum knowledge of the state of the system will always remain represented by a pure case that has dispersions. Moreover, the measurement augments nothing concerning our knowledge of the state of the system prior to the measurement, because *it will create an entirely new state by its action on the system*.

I think that these are the results of quantum physics that a definitive character, but, contrary to what one assumes, they do not at all imply the impossibility of maintaining the classical idea of individual entities – i.e., corpuscles – for which all quantities have well-defined values. As we will see, one can imagine that there exists a unique collective system in any pure case that will give probability distributions for all of the quantities and will satisfy all of the usual rules of statistics [with the existence of $\rho(x, y, z)$], except that these probability distributions will not, in general, be the ones that one ordinarily envisions in wave mechanics, because each of them will be realized only after performing

the corresponding measurement. The distributions that we will introduce cannot be exhibited, since in order to create them, one must perform measurements, and any measurement, by the action that it exerts on the objects that it affects, will generally change the probability distributions. Finally, it is precisely the new role that is played by measurement in microphysics that will permit us to imagine probability distributions that remain hidden with no contradictions. We will return to that question.

We again make an important remark, to which we will also have to return. A measuring device that involves individual microphysical entities will necessarily involve the appearance of an observable macroscopic phenomenon that is triggered by a microphysical individual entity. This must be true, since the measurement can only result from an observation that is made by the physicist. Therefore, in a Wilson chamber, the observation of a corpuscular trajectory, which can permit one to either localize or evaluate its energy or quantity of motion, will result from a phenomenon of macroscopically-observable condensation droplets that is triggered by the ionizing action of the corpuscle in motion. The same thing will be true for a photographic record, where the elementary action of a corpuscle (e.g., photon or charged particle) will trigger a macroscopically-observable chemical phenomenon. That very important remark will ultimately serve to reconstruct the theory of measurement on a new basis.

2. The statistics of two interacting systems, according to von Neumann. – Recall von Neumann's analysis, and consider two corpuscles or two systems of corpuscles that are involved in a measurement. Von Neumann said that the former is the "system under study" and the latter is the "measuring apparatus." We will have to critique these terms, but for now, we shall let that pass.

Let $u_k(x)$ be an orthonormal set of proper functions for the former system and let $v_\rho(y)$ be the analogous set for the latter one. When the systems are isolated from each other (in the initial state), their wave functions $\Psi_{\rm I}$ and $\Psi_{\rm II}$ will evolve separately according to the corresponding wave equation, and one can set:

$$(1) \qquad \Psi_{\rm I} = \sum_k c_k(t)\,u_k(x), \quad \Psi_{\rm II} = \sum_k d_\rho(t)\,v_\rho(x).$$

Since system I is obviously in a pure case, it will remain in a pure case. The total system, whose Hamiltonian H is then the sum of Hamiltonians $H_1 + H_2$ of the two systems, will have the wave function:

$$(2) \qquad \Psi(x, y, t) = \Psi_{\rm I}(x, t)\,\Psi_{\rm II}(y, y) = \sum_{k,\rho} c_k(t)\,d_\rho(t)\,u_k(x)\,v_\rho(y).$$

It will represent a pure case of the system that will persist as long as the interaction has not yet commenced.

When the interaction does commence, one must add an interaction term H_i to the terms $H_1 + H_2$ in the global Hamiltonian that will depend upon the coordinates, x and y, of the two systems in a form that is not simply additive. The wave function of the global

system will then cease to be the product of a $u_k(x)$ with a $v_\rho(y)$, but since the products $u_k(x)\, v_\rho(y)$ will continue to define a complete and orthonormal basis system for the set of variables x and y, one can write:

$$\Psi(x, y, t) = \sum_{k,\rho} C_{k\rho}(t)\, u_k(x)\, v_\rho(y), \tag{3}$$

but the $C_{k\rho}$ are no longer of the form $c_k\, d_\rho$. Since we will always have a wave function Ψ for the total system that evolves according to the wave equation, the total state of the system will always remain a pure case. The corresponding statistical matrix will then be given by:

$$P_{k\rho,l\sigma} = C_{k\rho} C^*_{l\sigma}. \tag{4}$$

Here, we remark that it takes two indices to represent a state of the global system. We now direct our attention to system I, and envision a certain quantity A in that system such that the corresponding matrix is defined by:

$$A_{kl} = \int_D u^*_k(x)\, A u_l(x)\, dx. \tag{5}$$

The mean value of A during the interaction is:

$$\bar{A} = \int_D \Psi^* A \Psi\, d\tau = \sum_{k,\rho,l,\sigma} C^*_{k\rho} C_{l\sigma} \int_D u^*_k A u_l\, dx \int_D v^*_\rho\, v_\sigma\, dy = \sum_{k,\rho,l} C^*_{k\rho} C_{l\rho} A_{kl}\ . \tag{6}$$

Now, the statistical matrix of system I during the interaction must be such that:

$$\bar{A} = \mathrm{Tr}(P_\mathrm{I}\, A), \tag{7}$$

which leads one to write:

$$(P_\mathrm{I})_{lk} = \sum_\rho C_{l\rho} C^*_{k\rho}. \tag{8}$$

One will likewise find that the statistical matrix of system II is:

$$(P_\mathrm{II})_{\sigma\rho} = \sum_k C_{k\sigma} C^*_{k\rho}. \tag{9}$$

The statistical matrix P of the total system is Hermitian, has a trace that is equal to 1, and is idempotent, as one easily verifies from (4) by taking into account the orthonormal character of the products $u_k(x)\, v_\rho(y)$; it is therefore an elementary statistical matrix. The same thing is not true for the matrices P_I and P_II, which permit one to calculate the mean value of the quantities for one or the other system. They are indeed Hermitian matrices with trace 1, but they are not idempotent. Therefore, the statistics of systems I and II, when considered separately, are no longer those of pure cases, but mixtures.

In order to specify the composition of these mixtures, recall formula (3). For a given value of the index ρ, we will have a sum of terms of the form:

$$v_\rho(y)\ \sum_k C_{k\rho}(t)\ u_k(x), \qquad \text{with} \quad \sum_{k,\rho} \left| C_{k\rho} \right|^2 = 1$$

in the development (3). One can then say that for a given value of ρ (i.e., for a certain state of system II), system I will have a probability that is proportional to $| C_{k\rho} |^2$ of being found in the state *k*. The absolute value of that probability will be equal to $\left| C_k^{(\rho)} \right|^2$ if one sets:

(10) $$C_k^{(\rho)} = \frac{C_{k\rho}}{\sqrt{\sum_l | C_{l\rho} |^2}},$$

in such a fashion that one has $\sum_k \left| C_k^{(\rho)} \right|^2 = 1$.

One can then write:

(11) $$(P_{\mathrm{I}})_{kl} = \sum_\rho p_\rho C_k^{(\rho)} C_l^{(\rho)*},$$

with:

(12) $$p_\rho = \sum_l \left| C_{l\rho} \right|^2,$$

and one will likewise find that:

(13) $$(P_{\mathrm{II}})_{\rho\sigma} = \sum_k p_k C_\rho^{(k)} C_\sigma^{(k)*},$$

with:

(14) $$C_k^{(\rho)} = \frac{C_{k\rho}}{\sqrt{\sum_\sigma | C_{k\sigma} |^2}}, \qquad p_k = \sum_\sigma \left| C_{k\sigma} \right|^2.$$

The matrices P_{I} and P_{II} thus indeed appear to be defining mixtures with statistical weights p_ρ and p_k, respectively.

Therefore, whereas the total system will remain in a pure case, despite the interaction, each of the two partial systems, which are considered to be isolated, will be transformed into a mixture by the interaction, and von Neumann added: “Whereas our knowledge of the global system will remain a maximum, that of the two component systems will cease to be a maximum. Since each partial system can be considered to be found in a pure case that we are ignoring, the mixture will represent that ignorance. A simple verification can then suffice to eliminate that ignorance by making the effectively-realized pure case known.”

By studying the forms of the statistical matrices P_{I} and P_{II}, one confirms that for each system the mixture is determined by the states of the other system. This is what translates – for example, in formula (11) – into the fact that the sum in the right-hand side involves an index ρ that relates to the second system. It is by verifying the state of the second system (i.e., the value of ρ, when effectively realized) that we can say what the pure case is that we can attribute to the first one. However – and this is a point of paramount importance that is not sufficiently emphasized in von Neumann’s theory – in

order to verify the state of the second system, it is necessary that it must trigger a macroscopic phenomenon that we can observe directly. This is a point that will seem clearer when we return to the question in a more physical manner.

3. The measurement of a quantity in the von Neumann formalism. – We just studied the interaction of two systems, but in order for that interaction to be able to provide the measurement of a quantity in the first system, the result of the interaction must be of a special type. In other words, not just any kind of interaction can serve to measure a quantity in the first system. Indeed, we have seen that by macroscopically verifying the state of the second system after the measurement, one can deduce only that the first one will be found in a certain pure case. However, since a physical quantity does not generally have a precise value in a pure case, we will therefore not generally obtain a measurement of the quantity that we are interested in.

Let A be the physical quantity in the first system that we desire to measure. Take the proper functions of A to be the basis functions of the first system. In order for the interaction with the second system to serve as a measurement of A, it is necessary that there exist a magnitude B of the second system such that if $v_\rho(y)$ are the proper functions of B then the Ψ of the total system after the interaction will have the form:

$$\Psi = \sum_{k,\rho} C_{k\rho} u_k(x) v_\rho(y), \tag{15}$$

with $C_{k\rho} = C_k \, \delta_{k\rho}$; i.e., one will have:

$$\Psi = \sum_{k} C_k u_k(x) v_k(y). \tag{16}$$

One can then establish a bijective correspondence between the v and the u, or, if one prefers to say this in another way, between the observable phenomena that are triggered by the second system and the value of A for the first one. We shall ultimately return to this point in detail in a manner that will make things much clearer.

Calculate P_{I} when (16) is realized. We will have:

$$C_k^{(\rho)} = \frac{C_{k\rho}}{\sqrt{\sum_l \left| C_{l\rho} \right|^2}} = \frac{C_k \delta_{k\rho}}{C_\rho} \tag{17}$$

and

$$p_\rho = \sum_l \left| C_{l\rho} \right|^2 = \sum_l \left| \delta_{l\rho} C_l \right|^2 = | C_\rho |^2, \tag{18}$$

so:

$$(P_{\mathrm{I}})_{kl} = \sum_\rho p_\rho C_k^{(\rho)} C_l^{(\rho)*} = \sum_\rho p_\rho \delta_{k\rho} \frac{C_k}{C_\rho} \delta_{l\rho} \frac{C_k^*}{C_\rho} = \delta_{kl} \, p_k = \delta_{kl} \, | C_k |^2, \tag{19}$$

so P_{I} is a diagonal matrix whose diagonal terms are the $| C_k |^2$. One easily sees that the same thing will be true for P_{II}, which is identical to P_{I} .

One will then have a mixture of states that each correspond to one value of α_k and one value of β_k, *which is a one-to-one correspondence*, and in which the probability of the pair of values α_k, β_k will be $|C_k|^2$. The verification of the value β_k of B by an observable phenomenon that is triggered by the second system will then permit us to attribute the value α_k to A; it is therefore truly a "measurement." The verification that we just supposed resulted from a macroscopic phenomenon that one could observe or record that makes our knowledge of A more precise by showing us that it will be the value of A that is effectively realized in the mixture that is produced by the interaction.

We now examine the conditions under which the hypothesis that we made on the form (16) of Ψ can be found to be satisfied. Suppose that before the measurement system II is in the state $v_0(y)$ and system I is in the state $u_k(x)$. The wave function of the global system in the initial state will then be:

$$\Psi(x, y) = v_0(y)\, u_k(x). \tag{20}$$

The hypothesis that was made on the final form of Ψ will be realized if, for any proper function $u_k(x)$ that is realized at the origin, one has:

$$\Psi(x, y) = u_k(x)\, v_k(y) \tag{21}$$

at the end of the process of interaction, where $v_k(y)$ is a proper function of the quantity B that corresponds bijectively to $u_k(x)$. Indeed, due to the linear character of the wave equation, if the initial state, instead of being represented by (20), is represented by the superposition:

$$\Psi(x, y) = \sum_k C_k v_0(y) u_k(x) \tag{22}$$

so at the end of the interaction the wave function will indeed have the form (16) and the measurement of A will be impossible.

In the presentation of the theory of measurement according to von Neumann that we just made, we avoided saying, as one generally does in the usual presentations, that system II is a measuring apparatus and that quantity B is, for example, the position of a needle. Indeed, system II must be a microscopic system at the atomic level, like system I, and its role is to *trigger* an observable macroscopic phenomenon in a measuring device. No measurement can be made by making a corpuscle act in a microscopic body directly.

Moreover, in our opinion, von Neumann's theory presents an exaggeratedly abstract character, like many of the present theories of quantum physics, moreover. It does not sufficiently specify the physical conditions of the process of measurement and the necessity of triggering an observable macroscopic phenomenon in this process. It contains a pure formalism, so it gives us no precise physical image of the manner by which a measurement process can take us from a pure case to a mixture.

4. Less-admissible consequences of the theory of measurement in the present interpretation of wave mechanics. – As von Neumann said, the evolution of the wave Ψ continually unwinds during the measurement, such that the global system will remain

in a pure case, while the state of each of the partial systems will be a well-defined mixture. There is a break in the continuity of that evolution, with the creation of a new situation, when the observer who is verifying the state of system II can attribute a wave function to system I that corresponds to a well-defined value of the quantity *A*. In this manner of looking at things, it is therefore the "perception of the observer" that, upon verifying the state of system II, permits one to reduce the mixture that relates to the state of the system studied that would result from the interaction to one of these terms.

It is indeed obvious that one's knowledge of a quantity that is being measured will result from the observer becoming aware of the result of the measurement. However, this seems to imply that one can give the wave Ψ only a subjective significance with the present interpretation of wave mechanics, so von Neumann and his commentators have arrived at some concepts that are truly difficult to accept. I would like to give a summary of them following the presentation that was recently given by London and Bauer ([1]).

Consider three partial systems: the object under study (x), the measuring apparatus (y), and the observer (z), which collectively define a unique global system. Describe the global system with the aid of the wave function:

$$\Psi(x, y, z) = \sum_k c_k \, u_k(x) \, v_k(y) \, w_k(z) . \tag{23}$$

For the global system, we have a pure case that persists during the measurement, and for the partial systems, we have a mixture. The wave function (23) will give a maximum knowledge of the global system without one knowing the state of the object being studied (x) precisely.

However, the observer has another viewpoint because for him it is only the object (x) and the measuring apparatus (y) that belong to the external objective world. As far as that is concerned, he is in a very different situation because he possesses the awareness or faculty of introspection that permits him to know his state directly. It is by virtue of that immediate knowledge that he claims the right (?) to create his proper objectivity by cutting the chain of statistical coordinations that are expressed by Ψ, and by verifying that: "I am in the state w_k, so the measuring apparatus is in the state v_k, and in turn, the object is in the state u_k ," a verification that implies the attribution of a well-defined value to the quantity A; i.e., a measurement of that quantity.

Such is the presentation of London and Bauer, who add that: "It is therefore not a mysterious interaction between the measuring apparatus and the object that produces the appearance of a new Ψ for the system. It is only the awareness of a 'Me' that separates it from the old function $\Psi(x, y, z)$ and constitutes a new objectivity by virtue of his conscious observation by henceforth attributing a new wave function $u_k(x)$."

I cited the phrase, "this 'Me' that separates the wave function," even though I do not understand it very well, which seems to me to be much more mysterious than an interaction between the object and the measuring apparatus would be. One understands why Schrödinger said, with an ironic twist: "The theory of the wave Ψ is becoming a psychological theory." It serves no great purpose to add that these considerations only support the opinion of Bohr, who said that in quantum physics one cannot draw an exact

([1]) *See* bibliography [**2**].

boundary between the objective and the subjective, because that statement is itself hardly comprehensible and explains nothing. The more that you think about it, the more you get the impression that all of that interpretation should be recast upon a different basis.

We continue our study of the London-Bauer pamphlet. The authors remark that their ideas will raise one difficulty: If objective reality is created only by an act of perception on the part of the observer then does that reality not vary from one observer to another? Now, it is certain that this is not true, because otherwise no collective science – i.e., a science that is common to everyone – would be possible. However, we are told, one must remark that the verification that permits the measurement is a macroscopic verification that does not modify the state of thing that is being observed. Nothing, in turn, will prevent another observer from making the same statement, and it is a fact of experience that all observers will make the same verification, up to errors in observation. We add that it is this fact that permits one to abstract from the personality of the observer and to create a science that has an objective character. In summation, in the mixture that results from the interaction of the measurement, there is one *and only one* possibility that it will prove to be realized for all *all* observers.

That explanation seems insufficient to us, because it amounts to confirming the fact that one would like to explain. The existence of a science that is common to everyone seems to us quite difficult to comprehend in a theory that would like to describe everything with the aid of a function Ψ that has a subjective character, since it is a function that will depend upon what happens in the perception of the observer. The undeniable agreement between the observations that are made by different observers seems to us to be comprehensible only if one does not assume the existence of an objective reality, and if one does assume its existence then that objective reality must be capable of being described by something other than the subjective function Ψ that the present orthodox interpretation of wave mechanics condemns us to envision uniquely.

The present interpretation of wave mechanics thus seems to be lost in the contradictions, because one does now know what exact meaning one is to attribute to the wave Ψ. Logically, one is led to attribute the meaning to it of a simple representation of a purely subjective probability that capable of reducing the probability packet when the user receives new information. However, it can then no longer account for the existence of a science that is common to everyone and an objective reality that is independent of the observer. Moreover, there exist some arguments for attributing an objective reality to the wave Ψ. Therefore, suppose that an observer has knowledge of the state Ψ of a corpuscle (or a system) and that he calculates the probabilities of the result of a measurement that he is to perform with that wave function. If, before he performs the projected measurement, another observer makes a measurement of the corpuscle *without the knowledge of the first one* then generally the statistical predictions of the first observer will be found to be false. Therefore, it is the action of the measuring device, and not the perception of the observer, that modifies Ψ, and that would seem to impose a certain character of objective reality on the wave function. Bohr always seemed to have recognized this character, but his very subtle thinking is often quite obscure. In reality, almost all authors that have presented the current interpretation of wave mechanics have alternatively passed from the idea of a function Ψ that is a simple subjective representation of probability to the idea of a wave that preserves a certain character of

reality, and it is only with the aid of this unconscious subterfuge that they can avoid the all-too-flagrant contradictions.

As for the corpuscle, one further knows even less of its exact nature, and one confines oneself to saying: “In quantum microphysics, a corpuscle is endowed with not just quantities of determinate values, but also with a set of potential distributions that refer to each measurable quantity, which are distributions that can each come into play only when the corresponding measurement has been made,” and this does not give one a very clear idea of what a corpuscle can be.

We shall now examine whether one cannot arrive at a better comprehension of what happens in the process of measurement by adopting a causal and objective interpretation of wave mechanics that seems clearer.

———

CHAPTER IV

CAUSAL INTERPRETATION OF WAVE MECHANICS (THEORY OF THE DOUBLE SOLUTION)

---

1. Ideas at the basis for the theory of the double solution. – I would like to rapidly summarize the bases for the interpretation of wave mechanics by the theory of the "double solution" that I sketched out in 1927, and which I have returned to developing for some years now, as a result of a paper by David Bohm, in collaboration with J. P. Vigier. I made a summary of that theory in a book that appeared recently, to which one can refer ([1]).

At the beginning of my work on wave mechanics, my initial idea was to preserve the idea of a physical reality that was independent of the observer, and to seek, as one always does in classical physics, a clear representation of physical processes in the context of space and time. I was thus led to seek a synthetic viewpoint of the duality between waves and particles that would be compatible with the ideas that I introduced (*Mécanique ondulatoire*, 1923-1924), and which was confirmed in a remarkable fashion (e.g., the work of Schrödinger in 1926, the discovery of the diffraction of electrons in 1927). Following a current of ideas that was manifested in the work of Mie and Einstein, I sought to represent the corpuscle as a sort of local accident – i.e., a singularity – within an extended wave phenomenon. That led me to represent physical reality, not by the continuous solutions Ψ of the wave equation that were considered exclusively by Schrödinger and his school, but by other solutions of that equation that I will denote by u, in order to distinguish them from the regular solutions Ψ, and which involve a singularity. Upon reflection, I immediately saw a great advantage to that concept of a corpuscle being "incorporated" into an extended wave field, and consequently consolidated with the global motion of that field. It seems to me to permit one to comprehend that the corpuscle is localized and that its motion can nonetheless be influenced by presence of obstacles that are at distant from its trajectory, which must necessarily be interpreted, by preserving the idea of a localized corpuscle, as the existence of interference phenomena and diffraction.

Nevertheless, it seems to me that the probabilistic interpretation of the regular wave Ψ, which originated in the work of Born, and was confirmed by its success, must be preserved. Whereas the wave u will be the true description of the structure of physical unity, the wave Ψ will be a fictitious wave with a subjective character that is capable of providing us with exact statistical information about the position and motion of the corpuscle. However, in order for it to be able to fulfill that role, it is further necessary that it be related to the wave u in some fashion.

My first researches into wave mechanics led me to attribute a particular importance to the "phase" of the wave that one associates with a corpuscle. It is essentially the

---

([1]) Bibliography [**3**]. *See* Vigier's thesis [**4**], as well.

agreement between phases of the corpuscle, which is considered to be a sort of clock, and the surrounding wave that led me to write the fundamental formulas of wave mechanics ($w = hv$; $\lambda = h / p$). It is therefore the frequency and wavelength – which are elements that are contained in the phase – that therefore establish a bridge between the propagation of the wave and the motion of the corpuscle. That led me to write the wave function that is usually envisioned in the form:

$$\Psi = a\, e^{\frac{2\pi i}{h}\varphi}, \tag{1}$$

in which a and φ are real, and to attribute a profound physical significance to the phase φ (which will coincide with the Jacobi function S in the geometrical optics approximation). On the contrary, the amplitude a, which is continuous, does not seem to me to have any objective significance, but only a statistical significance.

Among the probabilities that are envisioned by the probabilistic interpretation of wave mechanics that is already assumed in this epoch, the probability of presence $| \Psi |^2 = a^2$ seems to me to have a sort of priority, because, in my way of looking at things, it will correspond to the possibility that the corpuscle is at a a given point, independently of any measurement process. The other probabilities, such as $| c(\mathbf{p}) |^2$ for the value $\mathbf{p}$ of the quantity of motion [where $| c(\mathbf{p}) |^2$ is the Fourier coefficient that corresponds to $\mathbf{p}$ in the development of Ψ in monochromatic plane waves], must have a less immediate sense, from my viewpoint. They will be valid only after the action of a measuring device for the quantity envisioned on the real wave u, into which a corpuscle has been incorporated, when one does not know the result of that measurement, moreover.

Endowed with these general ideas, I have assumed the principle to which I gave the name of "the principle of the double solution":

Any regular solution of the type (1) *of the wave equation of wave mechanics must correspond to a singular solution of the type:*

$$u = f\, e^{\frac{2\pi i}{h}\varphi} \tag{2}$$

that has the same phase φ *as the solution* (1), *but with an amplitude f that presents a point-like singularity that is generally mobile.*

In the period of time when I wrote my book on the double solution, which was in the Spring of 1927, one knew the Schrödinger wave equation:

$$\Delta\Psi - \frac{8\pi^2 m}{h^2} V\Psi = \frac{4\pi i m}{h}\frac{\partial \Psi}{\partial t} \tag{3}$$

that corresponds to the motion of a corpuscle of mass m in a field that is derived from a potential function $V(x, y, z, t)$, which is supposed to be known. Today, one must consider equation (3) as valid only in the Newtonian approximation for the corpuscles of spin 0. Some time after Schrödinger's first papers, it became apparent that there must be an equation that generalized equation (3) when one accounted for the corrections of

relativity. This new equation, which one habitually calls the "Klein-Gordon equation," and which constitutes the relativistic wave equation for particles of spin 0, is written:

$$(4) \qquad \Box\Psi - \frac{4\pi i\varepsilon}{c^2} V \frac{\partial \Psi}{\partial t} - \sum_{x,y,z} \frac{4\pi i}{h} \frac{\varepsilon}{c} A_x \frac{\partial \Psi}{\partial x} + \frac{4\pi^2}{h^2}[m_0^2 c^2 - \varepsilon^2 (V^2 - A^2)]\Psi = 0,$$

where m_0 is the proper mass of the particle, ε is its electric charge, c is the velocity of light in vacuo, and **A** is the scalar and vector potential whose derivative is the electromagnetic field to which the corpuscle is subjected.

Since equation (4) is the most general one, and it contains equation (3) as a degenerate form in the non-relativistic approximation, it will be equation (4) that I will use as the basis for my reasoning. In the case of the absence of a field, one will then have:

$$(5) \qquad \Box\Psi + \frac{4\pi^2}{h^2} m_0^2 c^2 \Psi = 0,$$

and the simplest solution to this equation when one confines oneself to continuous waves will be the monochromatic plane wave:

$$(6) \qquad \Psi = a\, e^{\frac{2\pi i}{h}(Wt - pz)},$$

with a constant and $W^2 / c^2 = m_0^2 c^2 + p^2$. W is the energy of the corpuscle in motion with the velocity $v = \beta c$, so $W = \frac{m_0 c^2}{\sqrt{1-\beta^2}}$, and **p** is its quantity of motion $\mathbf{p} = \frac{m_0 \mathbf{v}}{\sqrt{1-\beta^2}}$, where the direction of motion is taken to be the z-axis.

Now, I easily found that the Klein-Gordon equation also admits the moving singular solution:

$$(7) \qquad u(x, y, z, t) = \frac{\text{const.}}{\sqrt{x^2 + y^2 + \frac{(z - ct)^2}{1 - \beta^2}}} e^{\frac{2\pi i}{h}(Wt - pz)},$$

which will take the form:

$$(8) \qquad u(x_0, y_0, z_0, t_0) = \frac{\text{const.}}{r_0} e^{\frac{2\pi i}{h} m_0 c^2 t_0} \qquad (r_0 = \sqrt{x_0^2 + y_0^2 + z_0^2})$$

when $v = 0$.

The solution (7) has the same phase as the solution (6), but its amplitude will present a point-like singularity at the point $x = y = 0$, $z = vt$ that displaces with the velocity v in the direction of wave propagation, which provides a clear image of the motion of the corpuscle. In this particular case, one will thus obtain what one seeks exactly, and the constant value of the wave amplitude Ψ will appear to simply have the following significance: If one ignores the position of the corpuscle-singularity then one must

consider all of the parallel trajectories and all of the possible positions of the corpuscle at any instant *t* to be equally probable.

Encouraged by this first success, I then considered the general case of equation (4), with potentials *V* and **A** that are given continuous functions of *x*, *y*, *z*, *t*, and in 1927 I proved the following remarkable result:

1. If there exist two solutions Ψ and *u* of equation (4), one of which has a continuous amplitude, while the other one has an amplitude that involves a point-like moving singularity *that has the same phase* φ [which must say that they can be written in the forms (1) and (2)] then the singularity of *u* will displace in space with an instantaneous velocity **v** that is defined by the formula:

$$\mathbf{v}(x, y, z, t) = -c^2 \frac{\operatorname{grad}\varphi + \frac{\varepsilon}{c}\mathbf{A}}{\frac{\partial\varphi}{\partial t} - \varepsilon V}. \tag{9}$$

This is the "guidance formula," which will give simply:

$$\mathbf{v} = -\frac{1}{m}\operatorname{grad}\varphi, \tag{10}$$

when one can neglect the relativistic corrections and suppose that the magnetic field is zero (i.e., one sets $\partial\varphi / \partial t - eV \sim m_0c^2$ and $\mathbf{A} = 0$), which is a form that will correspond to the Schrödinger equation (3). If propagation takes place in the geometrical optics approximation, moreover, then one can set $\varphi \sim S$, where *S* is the Jacobi function, and (10) will then be nothing but the classical formula $m\mathbf{v} = -\operatorname{grad} S$ of Hamilton-Jacobi theory.

2. The motion of the corpuscle is the same as if it were subjected, moreover, to a classical force that is derived from the potentials *V* and **A** and a "quantum" force that is equal to $-\operatorname{grad} Q$, where *Q* is a "quantum potential" that is ignored in classical theories, and which is written simply as:

$$Q = -\frac{h^2}{8\pi^2 m}\left(\frac{\Delta f}{f}\right) = -\frac{h^2}{8\pi^2 m}\left(\frac{\Delta a}{a}\right) \tag{11}$$

in the non-relativistic approximation of equation (3), where the quantities in the parentheses are calculated at the point where one finds the corpuscle at the instant *t* and the equality of the two expressions (11) for *Q* automatically follows from the hypothesis that the two waves Ψ and *u* have the same phase φ.

The guidance formula and the definition of the quantum potential will permit one to give a Lagrangian form to the dynamics of the corpuscle that is incorporated into its wave as a singularity.

2. Another manner of expressing the guidance formula, and some generalizations. – The guidance formula gives a mathematical form to the fact that

because the corpuscle is integrated into the wave it will be analogous to a clock that displaces while remaining in phase with the wave. From this viewpoint, it is the crowning achievement of my initial considerations on waves and corpuscles in wave mechanics. However, one can give it another form that will permit a greater generalization.

All of the forms of wave mechanics that are currently known will permit one to construct a hydrodynamical image that is associated with the propagation of a wave; i.e., to define a fictitious fluid whose density ρ and flux density $\rho\mathbf{v}$ are given at each point and each instant by functions that are bilinear in the wave function and its complex-conjugate function.

Therefore, in the case of the Schrödinger equation (3), the fictitious fluid and its motion will be given by formulas that were originally used by Madelung:

$$\rho = \Psi\Psi^* = |\Psi|^2, \qquad \rho\mathbf{v} = -(\Psi^* \operatorname{grad}\Psi - \Psi \operatorname{grad}\Psi^*), \tag{12}$$

where Ψ^* is the complex-conjugate quantity to Ψ. Thanks to (1), one can also write:

$$\rho = a^2, \qquad \mathbf{v} = -\frac{1}{m}\operatorname{grad}\varphi. \tag{13}$$

One sees from the expression for **v** that the guidance formula can be expressed by saying that the corpuscle follows *one* of the streamlines.

In the case of the Klein-Gordon equation (4), the fictitious fluid will be defined by:

$$\begin{cases} \rho = \dfrac{h}{4\pi i}\dfrac{1}{m_0c^2}\left(\Psi^*\dfrac{\partial\Psi}{\partial t} - \dfrac{\partial\Psi^*}{\partial t}\Psi\right) - \dfrac{1}{m_0c^2}V\Psi^*\Psi, \\ \rho\mathbf{v} = -\dfrac{h}{4\pi i m_0}\left(\Psi^*\operatorname{grad}\Psi - \Psi\operatorname{grad}\Psi^*\right) - \dfrac{1}{m_0c}\mathbf{A}\Psi^*\Psi, \end{cases} \tag{14}$$

or, thanks to (1):

$$\rho = \frac{1}{m_0c^2}\frac{\partial\varphi}{\partial t}a^2 - \frac{\varepsilon}{m_0c^2}Va^2, \qquad \rho\mathbf{v} = -\frac{1}{m_0c^2}a^2\operatorname{grad}\varphi - \frac{\varepsilon}{m_0c^2}a^2\mathbf{A}, \tag{15}$$

so:

$$\mathbf{v} = -c^2\,\frac{\operatorname{grad}\varphi + \dfrac{\varepsilon}{c}\mathbf{A}}{\dfrac{\partial\varphi}{\partial t} - \varepsilon V}. \tag{16}$$

We thus recover the guidance formula for the Klein-Gordon equation; the corpuscle then follows a streamline. We also see, moreover, that the hypothesis that is expressed by (1) and (2), according to which Ψ and u have the same phase φ, amounts to supposing that the streamlines are the same for two waves, or if one prefers, that the velocity vector field $\mathbf{v}(x, y, z, t)$ is the same for both of them.

In that form, the relationship that is established between the two waves u and Ψ can be generalized to the equations of particles with non-zero spin. Therefore, in order for the particles with spin $h / 4\pi$ – such as electrons – the wave Ψ will have four components Ψ_k that obey four simultaneous partial differential equations (viz., the Dirac equations):

$$\left(\frac{h}{2\pi i} \frac{\partial}{\partial t} - \varepsilon V \right) \Psi_k = \left[\sum_{j=1}^{3} \left(\frac{h}{2\pi i} \frac{\partial}{\partial x_j} - \frac{\varepsilon}{c} A_j \right) \alpha_j + m_0 c \, \alpha_4 \right] \Psi_k , \tag{17}$$

with k = 1, 2, 3, 4. The matrices α_1, α_2, α_3, α_4 are matrices with four rows and four columns, which are such that:

$$\alpha_i \, \alpha_j + \alpha_j \, \alpha_i = 2 \, \delta_{ij} \, I, \tag{18}$$

where I is the unit matrix. The fictitious fluid is then defined by:

$$\rho = \sum_{k=1}^{4} | \Psi_k |^2 , \qquad \rho v_j = - c \sum_{k=1}^{4} \Psi_k^* \alpha_k \Psi_k \qquad (j = 1, 2, 3), \tag{19}$$

so one will have the following components for the fluid velocity:

$$v_j = - c \, \frac{\sum_{k=1}^{4} \Psi_k^* \alpha_j \Psi_k}{\sum_{k=1}^{4} \Psi_k^* \Psi_k} = - c \, \frac{\sum_{k=1}^{4} u_k^* \alpha_j u_k}{\sum_{k=1}^{4} u_k^* u_k} . \tag{20}$$

In the theory of the double solution, it is this velocity that one agrees to attribute to the corpuscle-singularity in such a way that (20) will constitute the guidance formula in the Dirac theory (where one can no longer introduce a unique phase φ for the four components of the wave, in general). Here, one must replace the postulate on the phases φ with the one that of the velocity field **v** must be common to Ψ and u, which will justify the equality of the two expressions (20) for v_j .

For particles with spin greater than $h / 4\pi$ (e.g., photons, α-particles, gravitons, etc.), one will have wave functions with more than four components that will always obey a system of simultaneous partial differential equations. However, one can always define the density ρ and the flux ρ**v** of a fictitious fluid by means of bilinear formulas that analogous to (14) and (19) and obtain the corresponding guidance formula by assuming that the corpuscle-singularity of the wave u will always follow one of the streamlines that are common to the waves Ψ and u.

Now, it is a fundamental fact that for all of the wave equations that one has to consider the fictitious fluid will be conservative and obey the equation of continuity:

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{v}) = 0, \tag{21}$$

which is a consequence of the wave equations. It is equation (21) that will permit one to take the density ρ to be the probability of presence and to "normalize" Ψ (which is a simple representation of probability) by the formula $\int_D \rho \, d\tau = 1$. In the case of the Schrödinger equation, one will thus obtain $\int_D |\Psi|^2 \, d\tau = 1$ as the normalization formula, and one must take $|\Psi|^2 = a^2$ to be the probability of presence. We shall see that it is by starting with equation (21) that one can obtain the proof of the guidance formula in all cases.

3. Proof of the guidance formula. – We commence by remarking that if one assumes that any regular solution Ψ of the wave equation will correspond to a solution u with a moving singularity *that has the same streamlines* then the two densities $\rho(u)$ and $\rho(\Psi)$ will obey the same continuity equation, since the vector field **v** will be the same in the two cases, but, whereas $\rho(\Psi)$ is everywhere regular, $\rho(u)$ must present a point-like singularity that is mobile, in general.

A first manner of obtaining the guidance formula, which is basically equivalent to the one that we gave in 1937 ([1]), consists of writing the equation of continuity for $\rho(u)$ in the form:

$$\frac{\partial}{\partial t}\rho(u) + \mathbf{v}\ \text{grad}\ \rho(u) + \rho(u)\ \text{div}\ \mathbf{v} = 0; \tag{22}$$

hence, after dividing by $\rho(u)$:

$$\frac{\partial}{\partial t}\log \rho(u) + \mathbf{v}\ \text{grad}\ \log \rho(u) = -\ \text{div}\ \mathbf{v}. \tag{23}$$

If u, and in turn, $\rho(u)$ takes on very high values in a small region (which is obviously around the singularity) then $\log \rho(u)$ and its derivatives will have very high values there. For a given **v**, the right-hand side of (23) will then be negligible compared to the first one, and upon letting $\mathrm{D} / \mathrm{D}t$ denote the total derivative with respect to time along the streamline ($\mathrm{D} / \mathrm{D}t = \partial / \partial t + \mathbf{v} \cdot \text{grad}$), one will have:

$$\frac{\mathrm{D}}{\mathrm{D}t}\log \rho(u) = 0. \tag{24}$$

Therefore, $\log \rho(u)$, and in turn, $\rho(u)$ will remain constant when one follows a streamline with the velocity **v**. Thus, while there will be convergence and divergence of the streamlines for moderate values of ρ, in general (i.e., for $\text{div}\ \mathbf{v} \neq 0$), and in turn, ρ will not preserve a constant value when one displaces along a streamline with the velocity **v**, the same thing will no longer be true when ρ takes on extremely high values. The locally-elevated values of the density displace along streamlines with the velocity **v** without dispersing or damping out. One then sees that the singularity of $\rho(u)$ will follow one of the streamlines that are, by hypothesis, common to u and Ψ with a corresponding velocity **v**, which will give us the general form of the guidance formula.

([1]) One will it in [**3**], pp. 101, *et seq.*

Another method of obtaining the guidance formula consists of integrating the continuity equation, when it is written in the form:

$$\frac{\partial \rho}{\partial t} + v_X \frac{\partial \rho}{\partial x} + v_Y \frac{\partial \rho}{\partial y} + v_Z \frac{\partial \rho}{\partial z} + \rho \operatorname{div} \mathbf{v} = 0, \tag{25}$$

by the well-known method.

One knows that the integration of such a linear, first-order partial differential equation can be reduced to the integration of a system of differential equations:

$$\frac{dx}{v_x} = \frac{dy}{v_y} = \frac{dz}{v_z} = dt = -\frac{d\rho}{\rho \operatorname{div} \mathbf{v}}, \tag{26}$$

where v_x, v_y, v_z are functions of x, y, z, t that are assumed to be known. The integration of the first three differential equations (26) will give integrals of the form:

$$f_1(x, y, z, t) = \lambda, \qquad f_2(x, y, z, t) = \mu, \qquad f_3(x, y, z, t) = \nu. \tag{27}$$

When λ, μ, ν have constant values, these formulas will define a world streamline in space-time; i.e., a world-line, at each point of which $\frac{dx}{dt}$, $\frac{dy}{dt}$, $\frac{dz}{dt}$ will be equal to their values v_x, v_y, v_z at that point, respectively. This world streamline will, at the same time, represent the trajectory and the motion of the molecules of the fictitious fluid.

Equations (27) permit one to express x, y, z, t as functions of λ, μ, ν, t, and in turn, to express div **v** in the form $F(\lambda, \mu, \nu, t)$. In order to obtain the integration of the partial differential equations, it will then suffice to write the fourth differential equation (26) in the form:

$$dt = -\frac{d\rho}{F(\lambda, \mu, \nu, t)\rho}, \tag{28}$$

and then integrate this with λ, μ, ν constant, which will give:

$$\rho = e^{-\int^t F(\lambda, \mu, \nu, t)\, dt}\, \Phi(\lambda, \mu, \nu), \tag{29}$$

where the integration in the exponent of the exponential must be performed over t with λ, μ, ν constant, and where Φ is an arbitrary function. Since the continuity equation is, by hypothesis, valid for both Ψ and u with the same values of v_x, v_y, v_z, one will have:

$$\begin{cases} \rho(\Psi) = e^{-\int^t F(\lambda, \mu, \nu, t)\, dt}\, \Phi_1(\lambda, \mu, \nu), \\ \rho(u) = e^{-\int^t F(\lambda, \mu, \nu, t)\, dt}\, \Phi_2(\lambda, \mu, \nu). \end{cases} \tag{30}$$

Since $\rho(\Psi)$ is regular, the two factors in its expression must be regular. In the expression for $\rho(u)$, the first factor is the same as the one in the expression for $\rho(\Psi)$, so it must be regular. The singularity of $\rho(u)$ must then be provided by Φ_2 . It will then result that Φ_2 must have a singularity for a certain value of λ, μ, ν, namely, $\lambda = \lambda_0$, $\mu = \mu_0$, $\nu = \nu_0$, which translates into the existence of a point-like singularity of u that occupies a position x_0, y_0, z_0 at the instant t. However, that singularity will then be found along the world streamline that is defined by the values of λ_0, μ_0, ν_0, of λ, μ, ν. In other words, $\rho(u)$, and in turn, u will present a point-like singularity in space at any instant t, and the motion of that singularity in the course of time will be represented in space-time by the world streamline that is defined by $\lambda = \lambda_0$, $\mu = \mu_0$, and $\nu = \nu_0$.

When the singularity occupies the position x, y, z at the instant t it will then be animated with a velocity $\mathbf{v}(x, y, z, t)$. This is, once more, the guidance theorem in its most general form, and we can state our result by saying: *If two solutions of the wave equations of wave mechanics are such that one of them is regular and the other one has a moving, point-like singularity and they admit the same streamlines then the singularity of the second solution will follow one of these streamlines.*

It is important to remark that our proof will be further valid if the solution u, instead of presenting a true mathematical singularity, involves only a very small region, which is generally moving, in which it attains very high values, while the regular solution Ψ that it couples to involves no such analogous accident. The expressions (30) then show that the existence of that "singular region" must translate into a particular form of Φ_2 that must present very high values when λ, μ, ν have values that are close to certain values of λ_0, μ_0, ν_0 . However, this will further signify that the motion of the very small region where u takes on very large values in the course of time will be represented in space-time by a very fine world-tube, whose axis is defined by $\lambda = \lambda_0$, $\mu = \mu_0$, $\nu = \nu_0$.

In the general form that we just gave it, the theory of guidance permits one to better perceive the agreement between the wave u and the wave Ψ. These waves must have the same streamlines, so the wave Ψ will just as well represent the wave u, *which is the set* of possible motions for the corpuscle, but it will lack an essential element, which is the corpuscle itself that describes one of the streamlines. This is why, according to this viewpoint, if the wave Ψ can give can give an *exact statistical image* of the motion of the corpuscle then it cannot constitute a *complete description* of physical reality. Here, we come back to an opinion that Einstein has always maintained.

4. Introduction of nonlinearity and the form of the wave function u. – When I reprised the study of the double solution some years ago, with the active collaboration of Vigier, we were both immediately struck with the deep analogy that it presented with Einstein's ideas on the coupling between corpuscles and fields, as well as those of Mie in his nonlinear theory of electromagnetism. In the theory of the double solution, as in the thinking of Einstein and Mie, the desired goal is to *incorporate* the corpuscle in the field in the form of a very small region where the field takes on very high values (which might or might not involve a true mathematical singularity). However, in the theory of the double solution, the field that one seeks to incorporate the corpuscle in will no longer be

the electromagnetic field or the gravitational field, but the quantum wave field u, which must give an objective representation of the structure of the corpuscle.

Having arrived at this viewpoint, we now perceive the necessity of introducing a new idea.

Indeed, when one assumes linear equations of evolution for a field u, even when one introduces source terms that are independent of the field, as in Lorentz's theory, one cannot comprehend how a corpuscle can have a motion that is determined by the evolution of the field; this is a point upon which Einstein often insisted. In order to escape this difficulty, it will be necessary to suppose that the field equations are *nonlinear*. The idea of nonlinearity, which is new in wave mechanics, has nonetheless been introduced in recent years by some authors – notably, by Heisenberg – but in the context of ideas that are very different from ours.

Meanwhile, what we said before leads us think that if the wave equation u is nonlinear then the nonlinear terms that appear in it must be important only in small singular regions where the values of u are very high, which are the very small regions that constitute a corpuscle. Outside of that very small region, the nonlinear terms must be very small, and the equation of propagation of u must become approximately linear and coincide with the usual equation of propagation that is assumed for the wave Ψ, which brings us back to the hypothesis that we assumed previously, namely, that u and Ψ must obey the same equation.

Upon digging further into that idea, Vigier and myself eventually perceived that in order to be able to account for the success of the usual calculations of the phenomena of interference and diffraction, and also the success of the usual calculations of the proper values of the energies that correspond to the stationary states of quantized systems, it will be necessary to specify the form of the wave u by the following hypothesis: In the singular region where the equation for u is very approximately linear, u must have the form:

$$u = u_0 + v, \tag{31}$$

where u_0 is a solution of the linear equation with a point-like singularity at the center of the singular region, and v is a regular solution of the same equation. We shall specify how the form of the two terms in the expression (31) will follow. The term u_0 must be extremely small with respect to v outside of the immediate neighborhood of the singular region; that hypothesis will have great importance. The function u_0 will increase extremely rapidly when one approaches the singular region and it will become infinite at the center of that region if one can prolong to it; i.e., if the linear equation remains valid in the interior of the singular region. As for v, it will be a regular solution of the linear equation that must – at least, in general – coincide with the usual form that is assumed for the wave Ψ in the problem considered, up to a constant factor. Later on, we shall show that the set of hypotheses that we just posed can be found to be realized.

Therefore, externally, the solution u_0 of the linear equation seems to be a sort of very fine pointer that is implanted in a wave v that has the same form as the wave Ψ. Now, from the theory of guidance, the pointer function u_0 must displace along one of the streamlines of the wave v. However, if the wave equation u is everywhere linear and coincides everywhere with the usual equation of wave mechanics then the solution u_0 and v will be totally independent; there will be no reason for them to admit the same

streamlines, and the displacement of the pointer u_0 will be determined by the streamlines of v. Things are completely different if the equation for u is not linear in the singular region or the nonlinear terms are important by reason of the large values of u_0 . In that region, the terms u_0 and v will be linked to each other by the nonlinearity. In other words, the nonlinear equation in u will admit *one* solution u and the decomposition (31) will be valid only approximately in the region that is external to the singular regions; this will appear very clearly in an example that I would like to give shortly.

Moreover, the very localized nonlinearity of the wave equation u appears to be essential if one is to understand the meaning of the guidance theorem. The proof that we gave of this theorem by starting with the linear equation was based upon the hypothesis that the regular solution v (or Ψ) and the regular solution $u = u_0 + v$ have the same streamlines. Now, that hypothesis is entirely arbitrary in the context of an everywhere-linear theory. It will cease to be true if there exists a local nonlinearity in the small, singular region, because then the nonlinearity, although quite localized, is, in a sense, the "cement" that unites the solutions u_0 and v ([1]).

We shall now give an example that illustrates all of these considerations in a very useful way.

5. Illustration of the hypotheses made on u by an example. – Consider the simple case of a corpuscle with spin 0 that is at rest in a Galilean system, with the center of its singular region taken to be the origin of the coordinates, and suppose arbitrarily that the wave equation for u is the nonlinear equation:

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} - \Delta u + k_0^2 u = -\frac{1}{C^2} e^{-2\frac{a}{r}} \frac{a^2}{r^2} u^2 u^*, \tag{32}$$

with $k_0 = \dfrac{2\pi}{h} m_0 c$. When equated to zero, the left-hand side of (32) will give us the Klein-Gordon equation. The nonlinear right-hand side has a form in $u^2 u^*$ that was already envisioned by the authors who recently sought to introduce nonlinearity into wave mechanics. It will contain a function of the distance r from the origin that we have chosen arbitrarily and two constants C and a, the first of which is a numerical constant, and the second of which is a very small length that defines a "radius," in some sense, of a corpuscle with spherical symmetry. Later on, we shall set $\varepsilon = Ca$, by definition.

As one must have $u = f\, e^{ik_0 ct}$, where is f a function of only r, one will find the following equation for f:

([1]) We remark that the guidance theorem can be expressed by saying that in space-time the very high values of the wave function u are contained in the interior of a very thin world-tube whose walls are defined by streamlines of the "external" wave v (viz., the regular part of the wave u). When one states the guidance formula in that form, its lineage with the manner by which Georges Darmois and André Lichnérowicz stated the geodesic principle in general relativity becomes obvious (*see*, for example, LICHNÉROWICZ, *Théories relativistes de la gravitation et de l'Électromagnétisme*, Masson, 1955, book I, chap. III).

(33) $$\Delta f = \frac{\partial^2 f}{\partial r^2} + \frac{2}{r}\frac{\partial f}{\partial r} = \frac{1}{C^2} e^{-2\frac{a}{r}} \frac{a^2}{r^4} f^3,$$

which is an equation that will admit the solution:

(34) $$f = C\, e^{a/r}.$$

That solution will take on very high values in the neighborhood of the origin (i.e., for $r \ll a$), and it will even have a singularity at $r = 0$.

Here, can assume that the region that is external to the singular region is defined by $r \gg a$. One then sees that f will take on the approximate form:

(55) $$f = C + \frac{\varepsilon}{r} \qquad (\varepsilon = C\, a)$$

in the external region thus defined.

This is explained by the easily-verified fact that $\frac{\partial^2 f}{\partial r^2}$ and $\frac{1}{r}\frac{\partial f}{\partial r}$ are of the same order as $\frac{C}{r^2}\frac{a}{r}$, whereas the right-hand side of (33) is of the same order as $\frac{C}{r^2}\left(\frac{a}{r}\right)^2$, and is thus negligible with respect to the left-hand side if $r \gg a$. In the external region, the nonlinear equation (33) will reduce reasonably to the linear equation $\Delta f = 0$, so it is natural to find that f will take on the approximate form of a spherically-symmetric solution to that equation.

In the singular region that surrounds the origin, where r is of order a or less than a, the two sides of the nonlinear equation will become of the same order of magnitude, and one must take f to be the rigorous expression (34).

Now, if the linear equation $\Delta f = 0$ is valid everywhere then its general spherically-symmetric solution will be:

(36) $$f = A + \frac{B}{r},$$

where A and B will have arbitrary constant values. In the external region where the wave equation for u reduces reasonably to the Klein-Gordon equation, we have found the approximate form (35) for f, which coincides quite well with the general form (36), but with special well-defined values for A and B – namely, $A = C$ and $B = e = Ca$ – and one indeed sees that these special values are imposed by the nonlinearity of the wave equation for u in the very small singular region that surrounds the origin.

Moreover, the approximate solution (35) will have the form $u_0 + v$, where u_0 will have a singularity and v will be a regular function. Furthermore, since one has $\frac{u_0}{v} = \frac{\varepsilon}{Cr} = \frac{a}{r}$ here, one also sees that u_0 becomes much smaller than v as one gets more distant from the singular region. We have thus recovered the decomposition $u = u_0 + v$ in the external region with all of the characteristics that we have wished for.

One can, moreover, note that u is found to be the sum of the singular solution $\frac{\varepsilon}{r} e^{\frac{2\pi t}{h} m_0 c^2 t}$ that I already envisioned in 1927 for a corpuscle at rest that obeys the Klein-Gordon equation in the absence of a field ([1]) and a regular solution $Ce^{\frac{2\pi t}{h} m_0 c^2 t}$ that will take on the form of the classical monochromatic plane wave that was envisioned in the early years of wave mechanics when it is referred to a reference system in which the corpuscle has a uniform, rectilinear motion. In that reference system, the form of u will thus be, from (7):

$$u = \left(C + \frac{1}{\sqrt{x^2 + y^2 + \frac{(z-ct)^2}{1-\beta^2}}} \right) e^{\frac{2\pi i}{h}(Wt - pz)} \tag{37}$$

outside the moving singular region. All of this overlaps quite well.

We add a further remark: The solution (14) presents a singularity at $r = 0$. If, in accord with an opinion that is frequently expressed by Einstein, one considers it to be desirable to avoid any mathematical singularity in the representation of the corpuscle that is incorporated in the field then one will need only to take the nonlinear equation of u to be, in place of equation (32), the one that one obtains by replacing the variable r with the variable $\rho = r + \alpha$, where a is a positive length that is very small with respect to a ($0 \leq \alpha \ll a$). Thanks to that artifice, one can transform the solution (34) into:

$$f = C\, e^{a/\rho} = C\, e^{\frac{a}{r+\alpha}}, \tag{38}$$

in such a way that f will possess a very high, but finite, value for $r = 0$. The modification that was introduced will obviously make sense only in the center of the singular region, where $r \ll a$ will become of order α. It results easily from this that the values that we deduced from equation (32) will still remain valid.

Naturally, we have no reason to think that equation (32), with its right-hand side chosen arbitrarily, will be the true nonlinear wave equation for a corpuscle u of spin 0. However, it has the advantage of offering us a simple example of the manner by which a very localized nonlinearity in the equation for the wave field u can fuse together the two terms in the expression $u = u_0 + v$ that is valid in the exterior of the singular region and which completely determines the value of the coefficients C and ε that figure in u_0 and v.

6. The relationship between u and Ψ. – We shall now seek to specify the relationship between the function v and the function Ψ that is utilized in wave mechanics. Since the function u has an objective physical reality that is independent of the knowledge of the observer in the theory of the double solution, the function v that is a part of u and which practically agrees with u when one gets distant from the singular

([1]) *See* equation (8), above.

region (because $u_0 \ll v$ then) will also have an objective reality. In particular, v must have a perfectly-determined amplitude that is not at the disposition of the user and cannot be normalized at his discretion. However, the observer can mentally construct a function Ψ that must be, in principle, everywhere proportional to v, but with a coefficient of proportionality C that the user, who is free to give it any value that he desires, can choose in such a fashion that the function Ψ will be normalized. That function will then be a mental construct with a subjective character that will uniquely play the role of permitting one to calculate certain probabilities, but they must be constructed by the user, to the extent that his information on the form of the function v is exact, with the aid of the relation:

(39) $$\Psi = C\, v.$$

It is because the function Ψ was constructed by starting with v that it will have an objective reality that will permit an exact statistical evaluation of the probabilities, despite its subjective character ([1]).

If the function v occupies several disjoint regions in physical space, and the corpuscle is found in one of them, then the user can, according to the state of his knowledge of the position of the corpuscle, choose the constant C in a different fashion for each of the regions in question, and one can easily see how that will permit one to interpret the reduction of the probability packet.

Along the same lines, it is interesting to reflect upon the idea of a "pilot-wave" that I introduced in 1927, and which was reprised in some recent papers, notably in those of David Bohm. I remarked in 1927 that since, according to the guidance formula, the corpuscle must follow one of the streamlines of the wave Ψ, one can adopt the following viewpoint: Consider only the wave Ψ of conventional wave mechanics and arbitrarily add the notion of a corpuscle that displaces along one of the streamlines of the wave that will be found to be guided by the wave Ψ, which will permit one to give it the name of pilot-wave. However, I then consider – and I consider it more than ever today – the theory of the double solution that incorporates the corpuscle in the wave as being much more profound. Moreover, the wave Ψ of conventional wave mechanics has, without a doubt, a subjective character, since it changes with our information, and one cannot assume that the "guidance" of the corpuscle by something subjective would be real.

The question will be clarified if one distinguishes v from Ψ. The wave u involves a very localized accident that is represented by u_0, so everything happens as if that accident (i.e., the corpuscle) were guided in its motion by the wave v by following one of its streamlines. In reality, with the concepts that were discussed above, this will be true, since u_0 and v define a unique ensemble, namely, the wave function u (which is equal to $u_0 + v$ outside of the singular region), in which u_0 and v are fused together by the nonlinearity in the singular region. However, one can, by abstracting from these

([1]) On the subject of the formula $\Psi(x, y, z, t) = c\, v(x, y, z, t)$, one can remark, with Jean-Louis Destouches, that despite the equality of the two sides of that equation, the significance of the letters x, y, z is not the same on the right and the left. In v, they denote the current *variables* of space, while in Ψ they represent the *coordinates* of the corpuscle. In the case of one corpuscle in a given field, which is the only one that we have studied here, that remark, which is exact, can seem to be a little subtle. It will take on all of its validity when one studies the interpretation of wave mechanics for systems of corpuscles in configuration space by the theory of the double solution.

profound reasons, consider the corpuscle to be piloted by the wave *v*. Here, there is no longer any paradox, because the wave *v* has a physical reality, and consequently, the corpuscle can be guided by it. However, since the wave Ψ must be, in principle, chosen to be proportional to *v* and have, in turn, the same streamlines, one will get the impression that the corpuscle is guided by the wave Ψ, which is paradoxical. We can then appeal to the image of a corpuscle that is guided by a regular wave such that it will follow one of its streamlines, but with the condition that we remember that the regular wave is the wave *v*, and that the corpuscle is not an object that is arbitrarily superimposed on that wave, but constitutes a unique reality with it, namely, the wave *u* in the singular region.

We also think that the distinction that was just made precise between the wave Ψ and the wave *v* will permit one to understand why for the last thirty years or so theoreticians seem to have constantly oscillated more or less consciously between the idea of a wave with an objective character and that of a simple wave function that is an abstract representation of probabilities.

CHAPTER V

SOME COMPLEMENTARY NOTIONS TO THE THEORY OF THE DOUBLE SOLUTION AND GUIDANCE.

---

1. Existence of singular solutions in the exterior problem. – We shall call the study of the solutions u to the wave equation in the region that is external to the singular region *the exterior problem*, when that equation agrees – at least, in the first approximation – with the linear equation of propagation that was envisioned for the corpuscle that is considered by conventional wave mechanics, and when it admits an approximate solution of the form $u_0 + v$.

In 1927, I employed the Klein-Gordon equation exclusively, along with its degenerate form, the Schrödinger equation, and I did not distinguish v from Ψ. I would now like to prove that each solution Ψ in conventional wave mechanics will already correspond to a solution u_0 with a moving singularity that will have the same phase as Ψ. In the case of the absence of a field, I found the solution that was pointed out previously ([1]); however, that is just one very special case, and I have not seen how one could establish the existence of the function u_0 in a general fashion.

Today, when the theory of the double solution has taken on a more precise and coherent form, the study of the existence of singular solutions to the exterior problem and the coupling with the regular solutions continues to be interesting in its own right. Some notable progress in that direction was recently made in the thesis of Francis Fer. In that paper, the author considered a type of partial differential equation that contained the Klein-Gordon equation as a special case. Utilizing the general methods of integration for partial differential equations, he proved the existence of solutions with singularities that are expressed by formulas of the same type as retarded potentials. By studying the agreement between these singular solutions and the regular solutions, Fer was led to recover the same motion for the singularity that was predicted by the guidance formula. The work of Fer thus seems to make an important contribution to the establishment of singular solutions to the exterior problem in the theory of the double solution.

Naturally, that exterior problem, which is analogous to the one that is posed in general relativity when one studies the field outside of a very thin world-tube that is filled with matter, will correspond to an incomplete viewpoint if one assumes that the true equation that is satisfied by u is nonlinear and that the decomposition $u = u_0 + v$ is only an approximate expression that is valid only in the external region.

Without being able to give a general proof of the existence of u_0 in the exterior problem, I would like to insist on a method that seems to permit one to effectively construct the function u of the exterior problem in the case of stationary states.

([1]) Chap. IV, form. (7) and (8).

2. The Rayleigh-Sommerfeld formula. – There is something curious about the method that we shall present, namely, the fact that it has point of departure in the assertion of a fact that seems to constitute a very important difficulty for the theory of the double solution.

That difficulty originates in the theory of Green functions for wave equations, which is discussed in many books ([1]), and which is intimately linked to the general theory of linear integral equations.

Consider a wave function u that obeys a wave equation such that if u is a monochromatic solution – i.e., it depends upon time only through a factor e^{ikct} – then it will take the form:

(1) $$\Delta u + [k^2 - F(x, y, z)]\, u = 0.$$

We know that this is the case for the Schrödinger equation, for example, and that k^2 will then be proportional to the energy E of the corpuscle.

Envision a domain D in physical space that is connected and *finite-dimensional*. The stationary waves that can be based in the domain D are defined to be monochromatic wave solutions of (1) that are annulled on the boundary of the domain D. If, as in conventional wave mechanics, one confines oneself to the consideration of regular solutions Ψ of equation (1) then one will prove that the stationary wave exists only if the constant k has one of the values in a sequence $k_1, \ldots, k_n, \ldots$, whose set forms the spectrum of "proper values" of the problem considered. In wave mechanics, as one knows, the proper values define the quantized energies of the corpuscle in the domain D. The wave functions Ψ_n that are regular and zero on the boundary, and that correspond to them will be the "proper functions."

However, one can also envision solutions to equation (1) that will be zero on the boundary of the domain D, but which will present a point-like singularity at a point Q in that domain. These solutions are the "Green functions" of equation (1) for the domain D and the "source" point Q. These Green functions $G(M, Q)$ thus depend upon the current point M and the source point Q. One restricts them by the following two conditions:

1. They are zero on the boundary of the domain D.
2. The function $G(M, Q)$ has a point-like singularity at the point Q such that when M tends to Q, it will increase like $\dfrac{1}{MQ} = \dfrac{1}{r}$.

Now, the general theory of linear integral equations will give us the following theorem on the subject of the existence of Green functions:

The Green function $G(M, Q)$ will always exist when the constant k has a value that does not coincide with any of the proper values $k_1, k_2, \ldots$ If k does coincide with one of the proper values k_n then the Green function $G(M, Q)$ will exist only in the very special case in which the corresponding proper function $\Psi_n(M)$ is zero at the point Q.

([1]) In particular, *see* the bibliography [**5**] and [**6**].

One can prove this result by effectively constructing the Green function with the aid of a formula that was once given by Lord Rayleigh and then used frequently by Sommerfeld.

In order to prove it, we start with the remark that by reason of the point-like singularity in $1 / r$ that the Green function must present at the point Q, one must write the equation that is satisfied by $u = G(M, Q)$ in the form:

$$\Delta u + [k^2 - F(x, y, z)]\, u = \varepsilon\, \delta(M - Q), \tag{2}$$

where one has introduced a source term into the right-hand side that involves a numerical coefficient ε that is analogous to an electric charge whose value is arbitrary and the singular Dirac function $\delta(M - Q)$. Equation (1) will then be satisfied by u everywhere except for the point Q, where it will have a singularity in $1 / r$.

Now, one can develop $\delta(M - Q)$ in proper functions $\Psi_i(M)$ in the form:

$$\delta(M - Q) = \sum_i c_i \Psi_i(M), \tag{3}$$

with

$$c_i = \int \delta(M - Q) \Psi_i^*(M)\, d\tau = \Psi_i^*(Q), \tag{4}$$

hence:

$$\delta(M - Q) = \sum_i \Psi_i^*(Q) \Psi_i(M). \tag{5}$$

If we likewise develop $u = G(M, Q)$ in the form:

$$u = \sum_i d_i \Psi_i(M) \tag{6}$$

then we must have:

$$[\Delta + k^2 - F(x, y, z)] \sum_i d_i \Psi_i(M) = \varepsilon \sum_i \Psi_i^*(Q) \Psi_i(M), \tag{7}$$

and since $\Psi_i(M)$ is the solution of equation (1) with $k = k_i$, this will become:

$$\sum_i (k^2 - k_i^2)\, d_i \Psi_i(M) = \varepsilon \sum_i \Psi_i^*(Q) \Psi_i(M), \tag{8}$$

from which, one will infer:

$$d_i = \frac{\varepsilon \Psi_i^*(Q)}{k^2 - k_i^2}, \tag{9}$$

since the Ψ_i form a complete system.

By substituting this in (6), one will obtain the *Rayleigh-Sommerfeld* formula:

(10) $$u(M) = G(M, Q) = \sum_i \frac{\varepsilon \Psi_i^*(Q)\Psi_i(M)}{k^2 - k_i^2}.$$

It is appropriate to remark that the series that appears in the right-hand side of (10) is not absolutely convergent. Despite that defect, it can be used safely in general, as Sommerfeld has shown.

One easily recovers the results that are provided by the theory of linear integral equations from formula (10). Indeed, if the constant k does not coincide with any of the k_i then formula (10) will give the Green function whose existence is thus found to be proved. On the contrary, if k coincides with one of the k_i then the term in the sum (10) that has the index n will infinite, and the formula will not give an acceptable Green function, except for the very special case where $\Psi_n(Q)$ is zero. We will thus indeed recover the results that were stated above.

Now, upon reflection, these results first appear to be disastrous for the theory of the double solution. Indeed, in that theory, it obviously seems that one must make a function $u = G(M, Q)$ that is zero on the boundary of D correspond to the stationary state that is usually represented by the function $\Psi_n(Q)$ and to the value k_n, and like Ψ_n, that function will also present a point-like singularity at the point Q where the corpuscle is found and correspond to the value k_n of the constant k. However, it is precisely that Green function that will not exist, or at least, it will exist only if the corpuscle is found at a point Q such that $\Psi_n(Q) = 0$. Unfortunately, by virtue of the statistical significance of $|\Psi|^2$, which is certainly exact, the corpuscle will then have a zero probability of being found at Q. The contradiction is flagrant, and seems to constitute a redoubtable objection again the theory of the double solution.

Nevertheless, we shall see that when we look at things more closely the Rayleigh-Sommerfeld formula, far from constituting an objection against the existence of the wave u in the stationary case, will, on the contrary, provide the means to construct it.

3. Construction of the function u with the aid of the Rayleigh-Sommerfeld formula in the case of stationary states. – We shall start with the following remark: Since $u = u_0 + v$ is assumed to be zero on the boundary of the domain D, the function v must not be rigorously zero on that boundary, but equal to $-u_0$. Upon discarding the extremely improbable case where Q is situated so close to the boundary of D that the very small singular region that surrounds Q will touch that boundary, the values $-u_0$ that v must present on the boundary will be extremely small everywhere, but they will nonetheless not be rigorously zero. As a result, v cannot be considered to be exactly proportional to the proper function Ψ_n that is usually calculated. Therefore, v must be a solution to the linear wave equation that corresponds to a value of k that is extremely close to k_n, but not exactly equal to k_n.

We are thus led to think that the function u that corresponds to the stationary state with index n must be equal to the Green function $G(M, Q)$ that corresponds to a value of k that is slightly different from k_n. That Green function will then exist, and it must be given by the Rayleigh-Sommerfeld formula!

In order to examine the form of that function *u*, we write formula (10) by isolating the term with the index *n* and denoting the difference $k - k_n$ by δk_n , which, to abbreviate, I will call the "frequency shift." One can then write, very approximately:

$$n(M, Q) = \sum_{i \neq n} \frac{\varepsilon \Psi_i^*(Q)\Psi_i(M)}{k^2 - k_i^2} + \frac{\varepsilon \Psi_n^*(Q)\Psi_n(M)}{2k_n \delta k_n}. \tag{11}$$

Let $\Psi'_n(M)$ be a function then that is a solution of the wave equation for $k = k_n + \delta k_n$, where δk_n corresponds to a very small shift in frequency whose value we will determine later on. Since the function Ψ'_n is very close to the proper function Ψ_n, we will set:

$$\Psi'_n(M) = \Psi_n(M) + \delta\Psi_n(M),$$

in which $\delta\Psi_n$ is the very small variation of Ψ_n that will result when k_n varies by δk_n ([1]). One will then have:

$$u(M, Q) = \sum_{i \neq n} \frac{\varepsilon \Psi_i^*(Q)\Psi_i(M)}{k^2 - k_i^2} - \frac{\varepsilon \Psi_n^*(Q)\delta\Psi_n(M)}{2k_n \delta k_n} + \frac{\varepsilon \Psi_n^*(Q)\Psi'_n(M)}{2k_n \delta k_n}. \tag{13}$$

We are then assured of having obtained a solution of the exterior problem that is zero on the boundary of *D* and presents a point-like singularity in r^{-1} at the point *Q*, since we have only to apply the Rayleigh-Sommerfeld formula for a value of *k* that is different from all of the k_i .

Since the function $\Psi'_n(M)$ is regular, the singularity can affect only the first two terms in the right-hand side of (12). Therefore, if we set:

$$\begin{cases} u_0(M,Q) = \sum_{i \neq n} \frac{\varepsilon \Psi_i^*(Q)\Psi_i(M)}{k^2 - k_i^2} - \frac{\varepsilon \Psi_i^*(Q)\delta\Psi_i(M)}{2k_n \delta k_n} \\ v(M) = \frac{\varepsilon \Psi_i^*(Q)}{2k_n \delta k_n}\Psi'_n(M) = C\Psi'_n(M), \end{cases} \tag{13}$$

with

$$C = \frac{\varepsilon \Psi_i^*(Q)}{2k_n \delta k_n}, \tag{14}$$

then we will have finally converted the function *u* into the form $u = u_0 + v$, where u_0 and *v* are solutions of the linear equation, u_0 has a point-like singularity at *Q*, and *v* is regular.

([1]) Ψ'_n is equal to $\Psi_n + \frac{\partial \Psi_n}{\partial k_n}\delta k_n$. Upon neglecting the terms in δk_n^2, one will see that Ψ'_n is asolution (which is non-zero on the boundary of *D*) of the equation $[\Delta - F + (k + \delta k_n)^2]\ \Psi'_n = 0$ by taking into account the equation $[\Delta - F + k_n^2]\ \Psi_n = 0$ and its derivative with respect to k_n .

Moreover, we see that v is of the form $C\Psi'_n \approx C\Psi_n$; i.e., it is very reasonably proportional to the function Ψ_n , which is the proper function that is considered in conventional wave mechanics.

If the ratio ε / C is considered to be well-defined then we will obtain an expression for the value of the frequency shift δk_n that is determined entirely by the position of the singularity:

$$\delta k_n = \frac{\varepsilon}{C}\frac{\Psi_n^*(Q)}{2k_n}. \tag{15}$$

However, in the exterior problem, where one considers only the linear wave equation, the value of ε / C will be arbitrary. Indeed, ε is introduced into the right-hand side of (2) artificially with a well-defined value. As for C, that constant has no value imposed upon it *a priori*. Therefore, formula (15) will not provide us with a well-defined value for δk_n .

We previously saw that the viewpoint that is adopted in the exterior problem, where one implicitly considers the linear equation (1), with a zero right-hand side, to be valid everywhere except for a point Q, is insufficient. We must assume that in the immediate neighborhood of Q there will exist a very small singular region where equation (1) is no longer valid, and where one must take into account a nonlinear right-hand side. Now, as I showed in an example in the preceding chapter (§ 5), the localized nonlinearity in the singular region can suffice to impose perfectly-determined values on ε, C, and ε / C. It results from this that this localized nonlinearity must permit one to obtain a perfectly-determined and extremely small value of the frequency shift δk_n from formula (15).

Here, it is appropriate to make a remark that can be important. If the preceding theory is exact, since the frequency k of the "true" wave u will differ very slightly from k_n, then the usual method of calculating the quantized energies in wave mechanics that deduces these energies from the proper values of the Schrödinger equation will be tainted with a very slight inexactitude. However, in the present state of the theory of the double solution, we can always suppose that the ratio ε / C is very small so there will be no observable effect, even in the most precise spectroscopic measurements.

I have, moreover, given ([1]) the complete calculation of the function u in the case of a corpuscle at rest at the center of a spherical enclosure and showed that it can be represented by the Rayleigh-Sommerfeld formula. The calculation is rendered very easy by the fact that the proper functions and the Green function will then have very simple forms $\left(\text{viz.,}\ \frac{\sin k_n r}{r}\ \text{and}\ \frac{\cos kr}{r},\ \text{resp.}\right)$. Andre Rot just made ([2]) an analogous calculation in the more general case of a corpuscle that occupies an arbitrary position in a spherical enclosure, and then extended it to the case of any finite domain D when there is separation of the variable, and even in certain cases of infinite domains.

Be that as it may, it seems that, at least in the case of finite domains, and while ignoring the examination of certain questions of convergence, the Rayleigh-Sommerfeld formula, which seems to constitute a grave difficulty for the theory of the double

([1]) *See* [**3**], pp. 226-230.
([2]) *C. R. Acad. Sc.* **243** (1956), pp. 483 and 1281.

solution, provides, on the contrary, a method for constructing a wave function *u* in the exterior domain that possesses all of the requisite properties.

4. Interpretation of the statistical significance of $|\Psi|^2$ in the stationary states. – We shall now study a problem that was often considered as providing a strong objection against the guidance formula.

In all of the attempts at a causal interpretation of wave mechanics, one must demand to know how one can justify the fact that for quite some time it has been well-established that the square of the modulus of the wave function Ψ gives that probability of presence of the corpuscle at each point and each instant. Bohm and Vigier ([1]) made an important contribution to the solution of that problem by showing that if the motion of the corpuscle that is defined by the "guidance formula" is constantly subject to small random perturbations then the probability of presence in $|\Psi|^2$ must be established very rapidly. These small random perturbations play the same role as "molecular chaos" does in Boltzmann's statistical mechanics. To what can these incessant small random perturbations be due? To interactions with other systems that come close (i.e., collisions), to feeble fluctuations of the boundary conditions that are imposed upon the wave, perhaps even, as Vigier suggested, to interactions with a turbulent and uncoordinated wave field that fills up what we call "the vacuum."

From a general viewpoint, one can remark that in any theory that imposes a well-defined law of motion upon a corpuscle, it will be necessary to introduce a random element in order to obtain statistical mechanics (e.g., Boltzmann's molecular chaos in classical mechanics, Bohm and Vigier's hypothesis of perturbation in the casual interpretation of wave mechanics). However, the statistical result that the introduction of that random element permits one to justify will be, in some way, already contained in the equations of motion that one starts with, which will allow one to predict that result *a priori*. Therefore, in the context of the old mechanics of Newton and Einstein, one can prove Liouville's theorem, which asserts the conservation in the course of time of the domain in the extension-in-phase that is occupied by the representative points in that abstract space of a cloud of corpuscles that displaces in space according to the laws of dynamics. That theorem makes it probable *a priori* that the fundamental statistical principle in classical or relativistic statistical mechanics must be the equal probability of elements that are equal in the extension-in-phase. However, the rigorous proof of that proposition, which is the objective of ergodic theories, always seems to demand the more or less explicit introduction of a random element that is analogous to Boltzmann's molecular chaos.

Similarly, in the theories of the double solutions or the pilot wave (the distinction between the two is unimportant here), the role that is played in the old mechanics by Liouville's theorem belongs to the continuity equation $\left[\frac{\partial \rho}{\partial t} + \operatorname{div} \rho \mathbf{v} = 0\right]$ that is valid for the fictitious fluid that is associated with the propagation of the regular wave. That equation makes it probable *a priori* that in the new dynamics that arises from the

([1]) D. BOHM, Phys. Rev. **85** (1952), pp. 166 and 180; D. BOHM and J. P. VIGIER, Phys. Rev. **96** (1954), 208.

guidance formula, the quantity $\rho\, dv$ (where $\rho = |\Psi|^2$ for the Schrödinger equation) will be the probability for the corpuscle to be present in the volume element dv of physical space at the instant t. However, here again, that assertion cannot be truly justified by an argument that is analogous to that of Bohm and Vigier, namely, that one can introduce a random element in the form of incessant, small perturbations that we spoke of above.

No matter what the physical origin of these perturbations, we can represent them in the following manner: Suppose that, abstracting from these perturbations, the regular wave that is associated with a corpuscle (either the wave Ψ or the wave v, if one assumes that they are proportional) is of the form $a\, e^{\frac{2\pi i}{h}\varphi}$ with a and φ real. The motion of the corpuscle that is incorporated into that "unperturbed" wave will be defined by the guidance formula, which will be written:

(15 *bis*) $$\mathbf{v} = -\frac{1}{m}\ \text{grad}\ \varphi,$$

upon confining oneself to the simple case of the Schrödinger equation.

Introduce small perturbations: Although they are very numerous during each unit of our macroscopic time (for example, per second), we will assume that they are very long with respect to their duration. During one of these perturbations, the wave will take the form $(a + \varepsilon)\, e^{\frac{2\pi i}{h}(\varphi+\eta)}$, where ε and η are small perturbations of the amplitude and phase, respectively. By reason of the random character of the perturbations, it is natural to assume that the mean values in time $\overline{\varepsilon}$ and $\overline{\eta}$ will be zero. During the duration of the perturbation, the velocity of the corpuscle will become the sum of the unperturbed velocity that is given by (15 *bis*) and the additional velocity $\mathbf{v} = -\frac{1}{m}\ \text{grad}\ \eta$. Although the mean value of $\mathbf{v}$ is zero, these additional velocities will make the corpuscle move from its initial, unperturbed trajectory to another unperturbed trajectory, and then to a third, etc. Finally, although the duration of each of these perturbations is, by hypothesis, much shorter than that of the interval during which the corpuscle describes an unperturbed trajectory, the enormous number of perturbations that the latter is subject to will have the effect that after a length of time that is very short on our scale, the probability of presence $|\Psi|^2 = a^2$ will be found to be realized; this seems to prove the argument of Bohm and Vigier. Moreover, if one only goes to first order then since $\overline{\varepsilon} = 0$, that probability will also be found to be equal to the mean value of the square of the perturbed amplitude $\overline{(a+\varepsilon)^2}$.

We now arrive at the application of the guidance formula to the stationary states, which is an application that seems to lead to a great difficulty at first. Consider a stationary state of a quantized system; for example, an electron in the hydrogen atom. In general, the corresponding wave function will be of the form $a(x, y, z)\ e^{\frac{2\pi i}{h}E_n t}$, in which E_n is the quantized value of the energy, and a is a real function of the variables x, y, z. Formula (15) then tells us that the electron must remain at rest at an arbitrary, but well-defined, point in the atom. That would correspond to the fact that the quantum force – grad Q, which is derived from the quantum potential Q, can then equilibrate the

electrostatic force. In other cases, one can find that the electron is animated with a simple periodic motion. Therefore, for the electron in the hydrogen atom, when the wave function is of the form $\Psi = F(r, \theta)\, e^{im\alpha}\, e^{\frac{2\pi i}{h}E_n t}$, where r, θ are the polar coordinates around the kernel, since the phase is then a linear function of the longitude angle α, the electron must, from formula (15), describe a "parallel" around the polar axis with a uniform velocity. In all cases where the electron is either at rest or animated with a very simple periodic motion, one cannot at all see how the probability of presence $|\Psi|^2$ can be realized. This objection seems to be grave.

However, one now introduces small random perturbations that are brief and spaced-out from each other and begins to imagine the case where the electron of the hydrogen atom has a uniform, circular motion for its unperturbed motion. One can easily see that the length of the circular trajectory must be of order 10^{-8} to 10^{-9} cm, and the velocity of the electron, of order 10^9 cm/s. One then assumes, by way of example, that it produces a million brief perturbations per second, in the mean. Nevertheless, the corpuscle will have enough time to describe a million orbits around its unperturbed trajectory in each time interval between two consecutive perturbations, in the mean. That example shows that the corpuscle can be considered to be *almost constantly* animated with the unperturbed motion that is defined by formula (15), although it will change its circular trajectory a million times per second. This permits us to understand how, despite the circular form of the unperturbed trajectories, one can expect to find the electron at any point of the atom with the probability $|\Psi|^2$.

In the case where the electron remains at rest at a point of the atom in its unperturbed state, one can say that the unperturbed motion will reduce to a state of rest. However, if we always assume that it produces a million perturbations per second, in the mean, then the electron will be propelled from one position to another a million times per second, in the mean, and after a second, it will have occupied a million different positions in the atom, and that will be true despite the fact that it remains at rest, in the mean, in each of these positions during a length of time that is very long with respect to the period of its wave (which, being always close to $\frac{h}{m_0 c^2}$, will be of order 10^{-20} s). Here again, we arrive at an understanding of how one can realize the probability of presence with $|\Psi|^2$, even though the corpuscle remains *almost constantly* at rest, thanks to the continual jittering of the corpuscle that is due to perturbations.

5. Two theorems in the theory of the double solution-pilot wave. – We shall now prove two interesting theorems in the causal interpretation of wave mechanics that one can state in the language of the theory of the pilot wave, which will be equivalent to the theory of the double solution, here. These theorems have been known to us for quite some time. Moreover, they have been given by other authors, notably by Herbert Franke.

a. Theorem on the expression for the kinetic energy. – In conventional wave mechanics, one considers the wave function Ψ to be an indecomposable, complex quantity in which one does not make the modulus and the argument enter *separately*. One takes the Hamiltonian operator be:

(16) $$H = \frac{1}{2m} P^2 + V \qquad \left(P^2 = -\frac{h^2}{4\pi^2}\Delta \right),$$

where the operator $P^2 / 2m$ corresponds to the kinetic energy T of the classical theory. The mean value of the total energy E in the state Ψ is then:

(17) $$\bar{E} = \int_D \Psi^* \left(-\frac{h^2}{8\pi^2 m}\Delta + V \right) \Psi \, d\tau ,$$

in the usual formalism.

In the theory of the double solution-pilot wave, one writes $\Psi = a\, e^{\frac{2\pi i}{h}\varphi}$, and makes the amplitude a and the phase φ play distinct roles. By substituting this into the wave equations, one obtains the generalized Jacobi equation:

(18) $$\frac{\partial \varphi}{\partial t} \equiv E = \frac{1}{2m} (\text{grad } \varphi)^2 + V + Q,$$

with

(19) $$Q = -\frac{h^2}{8\pi^2 m} \frac{\Delta a}{a}.$$

Moreover, one also obtains the continuity equation:

(20) $$\frac{\partial}{\partial t}(a^2) + \text{div} \left(-\frac{a^2}{m} \text{grad } \varphi \right) = 0.$$

Since, from the guidance formula (15), the kinetic energy of the corpuscle has the well-defined value $T = \frac{1}{2m} (\text{grad } \varphi)^2$, one sees that the total energy E is the sum of the kinetic energy, the classical potential energy V, and the quantum potential energy Q. Since the probability of presence density is $|\Psi|^2$, one is led to write:

(21) $$\bar{E} = \int_D \left(\frac{1}{2m} \text{grad}^2 \varphi + V + Q \right) a^2 \, d\tau .$$

Now, by taking equation (20) into account, one easily finds that:

(22) $$-\frac{h^2}{8\pi^2 m}\Delta\Psi = \frac{1}{2m} (\text{grad } \varphi)^2 - \frac{h^2}{8\pi^2 m} \frac{\Delta a}{a}.$$

Upon substituting (22) into (17), and upon comparing that with (21), one will see that:

1. The usual expression (17) for $\bar{E}$ will coincide with the expression (21) that is given by the theory of the double solution-pilot wave.

2. In the expression (21) for $\bar{E}$, the term $-\frac{h^2}{8\pi^2 m}\Delta$ in the usual expression (17) does not correspond to the kinetic energy *T* that is defined by the guidance formula, but to *the sum of that kinetic energy and the quantum potential.* If that potential does not figure explicitly in formula (17) then that will be because it is contained in the term $\frac{1}{2m}P^2$ that the usual theory considers as corresponding to the kinetic energy, but which we interpret differently here.

This theorem is important for the exact comparison of the usual theory with the causal interpretation and the guidance formula.

b. Virial theorem. – In classical statistical mechanics, one proves a theorem that is known by the name of the "virial theorem," which is a theorem that notably plays a role in the kinetic theory of gases. I shall first recall the classical proof of that theorem. The motion of a corpuscle with a quantity of motion that equals **p** in a force field that is derived from a potential *V* is:

$$\frac{d\mathbf{p}}{dt} = -\operatorname{grad} V. \tag{23}$$

From this, one deduces that:

$$\frac{d}{dt}(\mathbf{r}\cdot\mathbf{p}) = \mathbf{p}\cdot\mathbf{v} + \mathbf{r}\cdot\frac{d\mathbf{p}}{dt} = 2T - \mathbf{r}\cdot\operatorname{grad} V, \tag{24}$$

where **r** is the radius vector that defines the position of the corpuscle. One then sees that *for a periodic motion* the left-hand side of equation (24) must be zero in the temporal mean, so one will get:

$$2\bar{T} - \overline{\mathbf{r}\cdot\operatorname{grad} V} = 0 \tag{25}$$

for such a motion. The quantity $\overline{\mathbf{r}\cdot\operatorname{grad} V}$ is called the "virial of the forces," and formula (25) expresses the classical virial theorem.

This theorem can be transposed into conventional wave mechanics. In order to do that, we introduce the following definitions:

$$\begin{cases} N = \dfrac{h}{2\pi i}\displaystyle\int_D \Psi^*(\mathbf{r}\cdot\operatorname{grad} V)\Psi\, d\tau \\ R = -\displaystyle\int_D (\mathbf{r}\cdot\operatorname{grad} V)\Psi^*\Psi\, d\tau, \\ T = \dfrac{1}{2m}\left(\dfrac{h}{2\pi i}\right)^2 \displaystyle\int_D \Psi^*\Delta\Psi\, d\tau. \end{cases} \tag{26}$$

It is obvious that R represents the mean value of the virial, here. Starting with the Schrödinger equation and performing some integrations by parts, one then proves that one will have:

(27) $$\frac{dN}{dt} = 2\bar{T} + R.$$

If the wave is stationary $\left(\Psi \sim e^{\frac{2\pi i}{h}Et}\right)$ then the left-hand side of (27) will be zero, and what remains is:

(28) $$2\bar{T} + R = 0,$$

which is obviously the transposition of the classical virial theorem into conventional wave mechanics.

We shall interpret formula (28) by taking the viewpoint of the causal interpretation.

Upon taking theorem a into account and introducing the quantum potential $Q = -\frac{h^2}{8\pi^2 m}\frac{\Delta a}{a}$ and its mean value $\int_D Q a^2 d\tau$, we rewrite formula (28) in the form:

(29) $$2\bar{T}' + 2\bar{Q} + R = 0,$$

where $\bar{T}'$ is the "true" kinetic energy $\frac{1}{2m}(\text{grad } \varphi)^2$, here, which corresponds to the guidance formula. Now, in the theory of the double solution-pilot wave, the virial theorem must obviously be written in the form:

(30) $$2\bar{T}' + R' + R = 0,$$

where

(31) $$R' = -\int_D Q a^2 d\tau = 2\bar{Q}$$

is the mean value of the virial of the quantum force, which one obviously must add to the mean value R of the virial of the classical force here.

In order to prove (31), it will suffice to show that:

(32) $$R' = 2\int_D (\mathbf{r} \cdot \text{grad } Q)\, a^2 d\tau,$$

namely:

(33) $$2\int_D a \Delta a \, d\tau = -\int_D \left(\mathbf{r} \cdot \text{grad}\frac{\Delta a}{a}\right) a^2 d\tau.$$

Now, one easily verifies that:

(34) $$-\int_D \left(\mathbf{r} \cdot \text{grad}\frac{\Delta a}{a}\right) a^2 d\tau = \int_D \frac{\Delta a}{a}(3a^2 + \mathbf{r} \cdot \text{grad } a^2)\, d\tau$$

$$= \int_D (3a\,\Delta a + 2\Delta a\,\mathbf{r}\cdot\operatorname{grad} a)\,d\tau$$
$$= 2\int_D a\,\Delta a\,d\tau,$$

because one easily sees, by a sequence of integrations by parts, that:

(36) $$2\int_D \Delta a\,\mathbf{r}\cdot\operatorname{grad} a\,d\tau = -\int_D a\,\Delta a\,d\tau.$$

The virial theorem in the form (30) is thus proved.

6. Some words about the wave mechanics of systems in configuration space. – One knows that in those beautiful papers in 1926, in order to construct the wave mechanics of systems of corpuscles in such a fashion that it admits the classical Hamilton-Jacobi theory as its geometrical optics approximation, Schrödinger was led to associate the motion of a system with the propagation of a wave in configuration space that was defined by the set of the 3*N* coordinates of the *N* corpuscles that constituted the system. He wrote the equation of propagation in configuration space in the form:

(36) $$\sum_{k=1}^{N}\frac{1}{m_k}\Delta_k\Psi - \frac{8\pi^2}{h^2}V\Psi = \frac{4\pi i}{h}\frac{\partial\Psi}{\partial t},$$

where m_k is the mass of the k^{th} corpuscle, whose coordinates are x_k, y_k, z_k, and $\Delta_k = \frac{\partial^2}{\partial x_k^2}+\frac{\partial^2}{\partial y_k^2}+\frac{\partial^2}{\partial z_k^2}$. The potential *V* corresponds to both the interactions that can be exerted on the system from the exterior, if it is not isolated, and to the interactions between the corpuscles of the system. In that case where $N = 1$, one comes back to the equation that is valid for just one corpuscle in a given external field.

By thus putting the propagation of the wave Ψ of a system in configuration space and that of the wave Ψ of a corpuscle in physical space on the same basis, one eliminates any character of physical reality from the wave Ψ, because the propagation of a wave in abstract configuration space can only be purely fictitious. Even in the case of one corpuscle, if one considers the wave equation of that corpuscle to be a special case of equation (36) for $N = 1$ then one will obtain an equation of propagation in the configuration space of the corpuscle that is defined by its coordinates *x*, *y*, *z*, and not an equation of propagation in the physical space that is defined by the spatial variables *x*, *y*, *z*. The wave must then be a purely abstract quantity.

Naturally, in the era when I sought to preserve the character of objective reality for the wave of wave mechanics, I could not assume that viewpoint. For me, any real phenomenon can be described in framework of space and time. It seems inadmissible to me that one can treat the problem of *N* corpuscles in interaction only by considering a type of wave propagation that is obviously fictitious in an entirely abstract configuration space. To my eyes, it must be possible to pose, and even to solve, that problem by considering the propagation in the physical space of *N* waves *u* with singularities that

mutually influence each other. One must then be able to prove that the statistical result of these interactions is provided by the consideration of Schrödinger's wave Ψ in configuration space, where since the wave Ψ is only a representation of a probability, it can have only an abstract character. In short, it will be only a statistical representation of the correlations that are established between the positions of the singularities of the wave *u* under the influence of their interactions.

Conforming to this program, one must seek to represent a system of *N* corpuscles as being composed of *N* trains of waves *u* that each carry a singular region and evolve in physical space in the course of time, such that the propagation of each of the wave trains is influenced by the actions that are exerted on them by the singular regions of the other wave trains. In my paper in 1927, I already made a first attempt to justify this by directing my attention to the role of the wave Ψ in configuration space. In these latter years, I have reprised my work along those lines, and I have presented it in my recent book ([1]). Certainly, one cannot say that a clear and rigorous proof has been obtained up to now, but I have reason to think that one will be obtained. In the present presentation, I also assume that when the traces of the waves *u* that correspond to various corpuscles interact – for example, in a measuring device – the statistical correlations that are established between the positions of the corpuscles by the theory of the wave Ψ in configuration space will be exact. This hypothesis will permit us to obtain the results that we desire without having to treat the problem of the motion of each wave train *u* during the period of interaction.

We shall now return to the problem of measurement, while nonetheless assuming the viewpoint of the theory of the double solution and subjecting it to a more detailed analysis than one has habitually done up to now.

([1]) *See* [**3**], chap. XII, and also *C. R. Acad. Sc.*, **244** (1957), 529.

CHAPTER VI.

POSITION OF THE CAUSAL INTERPRETATION IN REGARD TO MEASUREMENTS IN MICROPHYSICS.

---

1. The special role played by the position of the corpuscle. – The theory of the double solution re-establishes a description of phenomena in the context of space and time. It is therefore inclined to give a special role to the measurement of position. This seems natural, moreover, if one remarks that all observations are necessarily carried out in the context of physical space.

If one contemplates the manner in which one can perform the determination of the position of a corpuscle then one will be led to the following assertions: First, since the corpuscle is not directly observable, its presence can be detected only by way of some local macroscopic effect that it will provoke, and the same thing will be true for any microphysical system. That is why a photon that arrives at the sensitive coating of a photographic plate will produce a photoelectric effect, and the emitted photo-electron will trigger a cascade of chemical phenomena by ionization effects that translate into a local reduction of the silver bromide, and a local darkening of the sensitized plate that will be visible after photographic developing. Likewise, an electric corpuscle that penetrates a Wilson chamber will trigger condensation vapor droplets by an ionization effect that will leave a trace in the chamber, and a series of consecutive analogous actions will then provoke the appearance of a filament of vapor droplets that roughly sketch out the trajectory of the corpuscle.

Upon pondering that, it seems that any observable phenomenon that is provoked by corpuscles at the atomic level will be detectable in only that way. There will always be the local action of a corpuscle that finally triggers an observable macroscopic phenomenon at the origin of the observation. This is an essential point that was not sufficiently brought to light in von Neumann's analysis and the comments that he made. By contrast, in that analysis, one makes the measuring device, and similarly the indicator of the measuring device, play a role that seems exaggerated to us. In reality, the measuring device and its indicator can play only a secondary role in the measurement of *macroscopic* phenomena that are triggered by the local action of the corpuscle: For example, a galvanometer can serve to measure the current discharge that is provoked by the arrival of a corpuscle in a scintillation counter, but it is the triggering of the discharge that is the essential thing and not its measurement by the galvanometer. The role of the measuring instrument appears to us to be much less important than one frequently hears that it is, and there are even cases in which one can ignore it completely (for example, in the direct visual observation of a photographic plate).

In the usual interpretation, one very often considers that one measures the position of a corpuscle by making it pass through a hole that is pierced in a screen that is open for a very short period of time. One will then have a determination of the position of the corpuscle that will be as exact as one desires. However, independently of the fact that

one never performs such a measurement of the position of a corpuscle in practice, one must remark that there would have to be, in turn, some macroscopic phenomenon that was triggered by the corpuscle when it traversed the screen; otherwise, one would observe nothing. One can perform the experiment in the following fashion: Receive the wave train that contains the corpuscle on a pierced screen that has an infinitude of very close holes (e.g., a sieve) and place a photographic plate behind that screen.

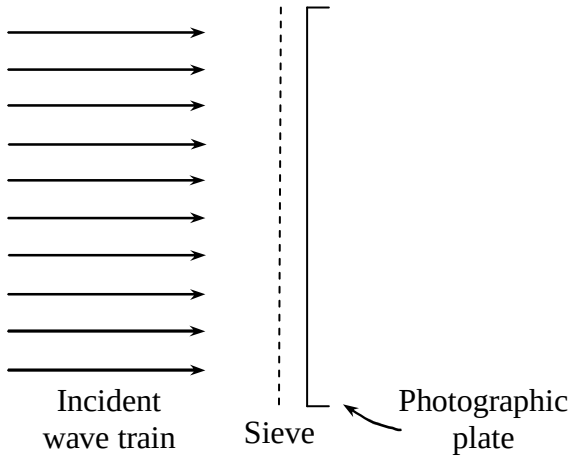

Figure 1.

The observation of a local darkening of the photographic plate will permit one to say that the corpuscle has passed through the hole that is situated in front of that darkening. The determination of the position, thus performed, will always be imprecise, since the dimensions of the black spot on the photographic plate will have a macroscopic order of magnitude (if that spot is observable), and in turn, will be much larger than the dimensions of the corpuscle. Nevertheless, one will arrive at a considerable improvement in the precision of the position of the corpuscle, since the dimensions of a hole in the sieve will be much smaller than the transverse dimensions of the incident wave train.

We arrive at the general idea, which is too often unknown, that if one cannot determine the position of a corpuscle in a very precise fashion then, meanwhile, any observation or measurement that relates to a microscopic corpuscle will always amount to observing a very localized macroscopic phenomenon that is triggered by the action of the corpuscle.

Contrary to what is said in the very abstract theory of representations in the usual wave mechanics, the position of a corpuscle will thus play a role that is completely different from that of the other measurable magnitudes. We repeat that this is quite natural since any observation is performed in the context of physical space. By disregarding that fact, the theory of representation tends to put the space of momenta (p_x, p_y, p_z) and physical space (x, y, z) on exactly the same footing, but that is too much of an abstraction: The physicist, his laboratory, and his instruments are in physical space, and the space of mometa exists only in the mind of the theoretician.

Since the theory of the double solution deals with more concrete ideas and establishes the privileged role that is incontestably played by experiments in physical space, it is not

surprising that it must lead us to attribute a special role to the probability distribution that relates to the position ($|\Psi|^2$, in the case of the Schrödinger equation). Indeed, it will teach us that this probability distribution corresponds to a collective property that one must associate with the initial state before the measurement. On the contrary, the probability distribution that the usual formalism attributes to a quantity that is not simultaneously measurable with the position is not generally realized in the initial state: It will correspond to a collective property that is associated with the state that will exist after the action of a device on the corpuscle that permits one to measure the quantity that is imagined.

2. Any measuring device will involve a separation of wave trains in space. – We shall now insist upon another circumstance that has also been ignored in the theory of measurement up to now.

We first place ourselves in the case where one must measure a quantity that relates to a corpuscle without making another corpuscle intervene. One must then employ a macroscopic device whose action on the propagation of the wave, into which the corpuscle is incorporated, will finally have the effect of *spatially* separating the wave trains that each correspond to a given value of the measurement. That conclusion is the immediate consequence of the fact that was brought to light in the last paragraph that any observation of a corpuscle will always consist of its localization. In order for the localization of a corpuscle after the action of a measuring device to permit us to say what the value of that quantity is at that moment, it is necessary that one have a one-to-one correspondence between the localization of the corpuscle and the value of the quantity that is being measured, and that is what will demand the separation of the wave trains in space after the measurement.

The device that we imagined above (Fig. 1) for the measurement of position satisfies that condition, since it has the effect of isolating the wave trains with very small lateral dimensions, which permits one to measure (somewhat imprecisely) the coordinates of the corpuscle in the plane of the screen, thanks to the very localized darkening that is triggered in the photographic plate that is placed behind the sieve.

Now, imagine the measurement of the quantity of motion, the knowledge of which will provide us with the energy. In order to measure the quantity of motion of a photon (and therefore, its energy, frequency, and "color"), one passes the incident sheaf through a device like a prism or grating that spatially separates the various wave trains that correspond to different frequencies by bending them into different directions. When the incident wave train is not monochromatic, the device will realize a true spectral decomposition by spatially isolating the various Fourier components of the incident wave. However, the same separation would obtained if the device successively received monochromatic wave trains that had different frequencies, because each of them will be bent into the direction that corresponds to its frequency. Since the wave trains can be superposed at the exact point when they leave the device, the localization of the photon in that region by the observation of a macroscopic phenomenon that it triggers will not permit one to attribute a well-defined frequency to it. One also ordinarily places a lens behind the prism or grating that will separate the various monochromatic sheaves and make them converge onto some small separated regions in its focal plane, where they will

each give a colored image of the source. If an observable macroscopic phenomenon (for example, the local darkening of a photographic plate) is triggered by the arrival of a photon in one of these regions then one can attribute a well-defined frequency to it. The initial sheaf is therefore divided by the action of the device (e.g., grating + lens) into a series of wave portions that will strike the photographic plate in spatially-disjoint regions, and it is that spatial separation that will permit one to measure the frequency, and in turn, the quantity of motion, of the photon. The fact that we just reasoned with a photon has no particular importance, because we know today that any corpuscle can give us phenomena of optical type, and we can construct devices for electrons, for example, that are analogous to a prism or a lens. There is therefore no essential difference between the photon and the other corpuscles in the problem that we are examining.

More generally, we can analyze this kind of measurement in the following fashion: Suppose that we would like to measure a quantity A that relates to a corpuscle. If the initial wave train R_0 is represented by the wave function:

$$\Psi = \sum_k c_k \varphi_k \,, \tag{1}$$

where φ_k is the proper function of A that corresponds to the proper value α_k, then we will send that wave train into a device D (grating + lens, in the case that was studied above) that separates the components $c_k\ \varphi_k$ in such a way that upon leaving the device D, each of them will occupy a region R_k that is spatially separated from the regions that are occupied by the other ones.

If we then observe (with the aid of a photographic record, or some other one) a macroscopic phenomenon that is triggered by the corpuscle in the region R_j then we can say that the quantity A of that corpuscle will have the value α_j after the action of the device, and we will have thus performed a measurement of A. The formalism of wave mechanics tells us that the value α_j will have the probability $|\ c_j\ |^2$; i.e., if we perform the same measurement experiment a great number of times with the wave trains represented by the same wave function (1) then the proportion of cases in which we will obtain the value α_j will be given by $|\ c_j\ |^2$.

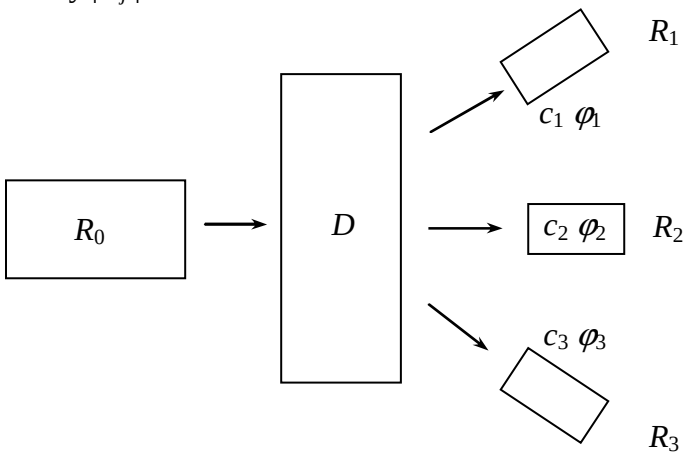

Figure 2.

In the usual interpretation, where one adds nothing to the concept of the wave Ψ, there will be no localization of the incident corpuscle in R_0, nor will there be any in the regions R_1, R_2, … after the action of the device D. It will be only at the moment when an observable phenomenon is produced in R_j that the corpuscle will be *briefly* localized in that region. In the von Neumann-London-Bauer theory, one must even say that it is the awareness of the macroscopic phenomenon by the observer that localizes the corpuscle in R_j . However, that seems truly unacceptable! It seems obvious that the macroscopic phenomenon will be produced, even if the observer has his eyes closed, in such a way that the awareness of the observer should have nothing to do with that.

What is equally incomprehensible in the present explanation is how it happens that the triggering of an observable macroscopic phenomenon in R_j will *instantly* prevent the corpuscle from manifesting itself in any other region R_k . Something even more surprising is that the various R_k can be found to be very distant from R_j at the moment of localization of the corpuscle.

It was, in short, that very difficulty that Einstein pointed out, in a somewhat different form, to the Solvay Council in 1927, and which has never been neatly eliminated. Einstein said: "Consider a planar screen that is pierced by a hole upon which a train of waves Ψ falls normally."

"Behind the screen, if the hole is very small then the wave will take the form of a spherical wave whose center will be the hole. Then place a hemispherical film F behind the screen. If the corpuscle is manifested by a point A of F then that will be interpreted very easily if the corpuscle has followed a well-defined trajectory (such as the one that is represented by the dashed line in Fig. 3) that takes it to A. However, if the corpuscle is not localized then it will be spread out into the potential state in all of the hemispherical wave that is behind the screen. How can the fact that it is manifested at A instantaneously prevent it from manifesting itself at any other point B of the film, which is a point that can be at a great distance from A?"

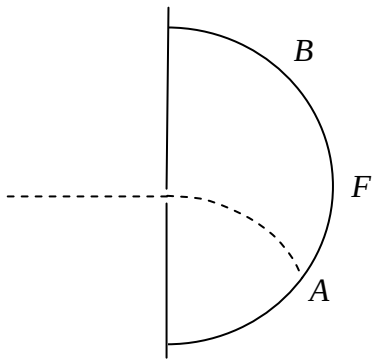

Figure 3.

One see that this objection is indeed the same as the one that we presented above, because Einstein's device is a measuring device for the position of the corpuscle. Moreover, one will be led from this device to the one that we envisioned previously (Fig. 1) by supposing that the hemispherical film is placed immediately behind a likewise hemispherical screen that is pierced by an infinitude of holes.

One can object to Einstein's argument by saying that the corpuscle does not manifest its presence at the point *A*, but in a very small region around *A*. However, since the surface of that region is very small with respect to that of the hemisphere *F*, Einstein's objection remains valid.

Return to Figure 2. One will encounter difficulties that will seem insurmountable if one does not want to assume that the corpuscle is localized, but everything will be clear if one establishes the localization of the corpuscle as the theory of the double solution does. Indeed, the corpuscle must then have a position in the initial wave train. Moreover, we cannot know that position, because in order to measure it, we will be obliged to employ a device that completely perturbs the initial wave train. However, we assume that this position exists, and that the probability for the corpuscle to be found at the point M_0 of the wave train at the initial instant t_0 will be given by $| \Psi(M_0, t_0) |^2$. (One will refer to paragraph 4 of the last chapter for the justification of the latter hypothesis.) From the guidance theorem, the corpuscle that starts at the initial position M_0 at the instant t_0 must follow the streamline that passes through M_0 . The motion that results for it is generally very complicated: It is rectilinear and uniform only before the action of the device when the wave train is reasonably monochromatic. However, we know that after the passage through the device that separates the wave trains R_1, R_2, …, the motion of the corpuscle will finally lead to one of the wave trains R_k , and the probability for it to be in R_j is obviously:

$$\int_R | \Psi |^2 \, d\tau = \int_{R_j} | c_j |^2 \, | \varphi_j |^2 \, d\tau = | c_j |^2,$$

since φ_j is normalized ([1]) and zero outside of R_j . If the corpuscle triggers an observable macroscopic phenomenon in R_j then that will be because it has arrived in R_j and the magnitude *A* will have the value α_j .

Obviously, it is necessary that the observer must confirm the triggering of the macroscopic phenomenon in order for him to become aware of the fact that *A* has the value α_j when the corpuscle is in R_j . However, this fact is independent of the awareness of the observer, and everything will become clear again.

3. Recovering the usual schema of statisticians. – We shall now show that the ideas of the theory of the double solution (which one can apply here in the simple form of a pilot wave) will immediately lead to an organization of the issues that are concerned with the question of the probability distribution.

We shall envision the case in which the quantity that one must measure is the quantity of motion **p**. We suppose that the Schrödinger equation is valid, and we write:

$$\rho(\mathbf{r}) = | \Psi(\mathbf{r}) |^2 \tag{2}$$

for the probability density of the position **r**. We then set:

([1]) In the development (1), the φ_k are, in reality, proper differentials that represent the wave group.

(3) $$\Psi = a\, e^{2\pi i \varphi / h} = \int c(\mathbf{p})\, \Psi(\mathbf{p}, \mathbf{r})\, d\mathbf{p} = \sum_k b(\mathbf{p}_k) \Psi_{\mathbf{p}}$$

in the initial state ([1]).

In the initial state, if the corpuscle is at a point **r** then, from the guidance formula, it will have a quantity of motion that is:

(4) $$\mathbf{p} = m\mathbf{v} = -\operatorname{grad} \varphi(\mathbf{r}).$$

Before the measurement of **P**, one will have:

(5) $$\rho_{\mathbf{R}}(\mathbf{r}) = |\, \Psi(\mathbf{r})\, |^2, \qquad \rho_{\mathbf{P}}^{\mathbf{R}}(\mathbf{p}, \mathbf{r}) = \delta(\mathbf{p} + \operatorname{grad} \varphi)$$

for the two random variables **R** and **P** in the initial state, where the second formula signifies that if one knows that value of **r**, and therefore $\varphi(\mathbf{r})$, then **p** will have the value that is given by the guidance formula. One will further have:

(6) $$\rho(\mathbf{r}, \mathbf{p}) = |\, \Psi(\mathbf{r})\, |^2\, \delta(\mathbf{p} + \operatorname{grad} \varphi),$$

and one will verify that:

(7) $$\int \rho(\mathbf{p}, \mathbf{r})\, d\mathbf{p} = |\, \Psi(\mathbf{r})\, |^2 = \rho_{\mathbf{R}}(\mathbf{r}), \qquad (d\mathbf{p} = dp_x\, dp_y\, dp_z).$$

One will likewise have:

(8) $$\rho_{\mathbf{P}}(\mathbf{p}) = \sum_i |\, \Psi(\mathbf{r}_i)\, |^2 ,$$

where the $\mathbf{r}_i$ correspond to the positions of the corpuscle for which grad φ has the value **p** in question, and one will verify that ($d\mathbf{r} = dx\, dy\, dz$):

(9) $$\int \rho(\mathbf{r}, \mathbf{p})\, d\mathbf{r} = |\, \Psi(\mathbf{r})\, |^2\, \delta(\mathbf{p} + \operatorname{grad} \varphi)\, d\mathbf{r} = \sum_i |\, \Psi(\mathbf{r}_i)\, |^2 = \rho_{\mathbf{P}}(\mathbf{p}).$$

In order to complete the statistical schema of the classical type that relates to the initial state, one must further define $\rho_{\mathbf{R}}^{\mathbf{P}}(\mathbf{p}, \mathbf{r})$, which one does by setting:

(10) $$\rho_{\mathbf{R}}^{\mathbf{P}}(\mathbf{p}, \mathbf{r}) = \frac{\rho(\mathbf{r}, \mathbf{p})}{\rho_{\mathbf{P}}(\mathbf{p})} = \frac{|\, \Psi(\mathbf{r})\, |^2\, \delta(\mathbf{p} + \operatorname{grad} \varphi)}{\int |\, \Psi(\mathbf{r})\, |^2\, \delta(\mathbf{p} + \operatorname{grad} \varphi)\, d\mathbf{r}}.$$

Finally, if one considers all of the possible positions **r** of the corpuscle in the initial wave train and the corresponding values of **p** that are given by the guidance formula then one will have defined a collection of individual entities that have well-defined positions

([1]) In order for the measurement of **p** that separates wave trains in space to be performed, it is necessary that the $\mathbf{p}_k$ must define a discontinuous sequence. Nevertheless, one can employ the integral by considering it to act on proper differentials.

and momenta, for which one will have a statistical schema of the classical type that will correspond to the following table:

$$\text{(I)} \qquad \left\{ \begin{array}{c} \rho_{\mathbf{R}}(\mathbf{r}) = |\,\Psi(\mathbf{r})\,|^2, \qquad \rho_{\mathbf{P}}(\mathbf{p}) = \sum_i |\,\Psi(r_i)\,|^2, \\ \rho(\mathbf{r},\mathbf{p}) = |\,\Psi(\mathbf{r})\,|^2\ \delta(\mathbf{p} + \operatorname{grad}\varphi), \\ \rho_{\mathbf{P}}^{(\mathbf{R})}(\mathbf{r},\mathbf{p}) = \delta(\mathbf{p} + \operatorname{grad}\varphi), \qquad \rho_{\mathbf{R}}^{(\mathbf{P})}(\mathbf{r},\mathbf{p}) = \dfrac{|\,\Psi(\mathbf{r})\,|^2\ \delta(\mathbf{p} + \operatorname{grad}\varphi)}{\int |\,\Psi(\mathbf{r})\,|^2\ \delta(\mathbf{p} + \operatorname{grad}\varphi)\, d\mathbf{r}}, \end{array} \right.$$

with the likewise classical relations:

$$\text{(I′)} \qquad \left\{ \begin{array}{l} \int \rho(\mathbf{r},\mathbf{p})\, d\mathbf{r} = \rho_{\mathbf{P}}(\mathbf{p}), \qquad \int \rho(\mathbf{r},\mathbf{p})\, d\mathbf{p} = \rho_{\mathbf{R}}(\mathbf{r}), \\ \rho_{\mathbf{P}}^{(\mathbf{R})}(\mathbf{r},\mathbf{p}) = \dfrac{\rho(\mathbf{r},\mathbf{p})}{\rho_{\mathbf{R}}(\mathbf{r})}, \qquad \rho_{\mathbf{R}}^{(\mathbf{P})}(\mathbf{r},\mathbf{p}) = \dfrac{\rho(\mathbf{r},\mathbf{p})}{\rho_{\mathbf{R}}(\mathbf{p})}. \end{array} \right.$$

However, we can insist upon an important point: Except for $\rho_{\mathbf{R}}(\mathbf{r})$, the probability distributions that we just defined will be "hidden," in the sense that we cannot determine them experimentally. Indeed, except for $\rho_{\mathbf{R}}(\mathbf{r}) = |\,\Psi(r)\,|^2$, which we can determine directly with the aid of the screen-sieve device in Fig.1, we cannot determine the values of **p** without making a measurement of that quantity, a measurement that will change the wave state completely by destroying the superposition of $\Psi_{\mathbf{p}_k}$ and the original collection that it is associated with, since that measurement will involve a spatial separation of the components $\Psi_{\mathbf{p}_k}$ in the development (3). We have thus indeed defined a collection with the aid of table (I), but it will be a *hidden* collection.

Now, study the situation after a measurement of **p**. The measuring device has partitioned the initial wave into trains of partial waves that each correspond to one of the components $\Psi_{\mathbf{p}_k}$. We have seen above that we will then have $\rho_{\mathbf{P}}(\mathbf{p}_k) = |\,c_k\,|^2$, which will be written:

$$\text{(11)} \qquad \rho_{\mathbf{P}}(\mathbf{p}) = |\,c(\mathbf{p})\,|^2,$$

in continuous notation, and we will also have:

$$\text{(12)} \qquad \rho_{\mathbf{R}}(\mathbf{r}) = \int |\,c(\mathbf{p})\,|^2\ |\,\Psi(\mathbf{p},\mathbf{r})\,|^2\, d\mathbf{p}$$

and

$$\text{(13)} \qquad \rho(\mathbf{r},\mathbf{p}) = |\,c(\mathbf{p})\,|^2\ |\,\Psi(\mathbf{p},\mathbf{r})\,|^2.$$

Since the proper function $\Psi(\mathbf{p},\mathbf{r})$ is normalized ([1]), one will verify immediately that:

$$\text{(14)} \qquad \int \rho(\mathbf{r},\mathbf{p})\, d\mathbf{p} = \rho_{\mathbf{R}}(\mathbf{r}), \qquad \int \rho(\mathbf{r},\mathbf{p})\, d\mathbf{r} = \rho_{\mathbf{P}}(\mathbf{p}).$$

([1]) *See* note pp. 89.

Since the probabilities are linked, we will find that:

$$(15) \qquad \rho_{\mathbf{P}}^{(\mathbf{R})}(\mathbf{r},\mathbf{p}) = |\,\Psi(\mathbf{p},\,\mathbf{r})\,|^2, \qquad \rho_{\mathbf{R}}^{(\mathbf{P})}(\mathbf{p},\mathbf{r}) = \int_{\mathbf{R_p}} \delta(\mathbf{r}'-\mathbf{r})\,d\mathbf{r}',$$

in which, $\mathbf{R_p}$ is the region where the wave will reduce to its component $\mathbf{p}$.

The last formula (15) expresses the idea that for a known $\mathbf{p}$, if the point $\mathbf{r}$ is in $\mathbf{R_p}$ then one will have $\rho_{\mathbf{R}}^{(\mathbf{P})} = 1$, and that if $\mathbf{r}$ is not in $\mathbf{R_p}$ then one will have $\rho_{\mathbf{R}}^{(\mathbf{P})} = 0$, and this is nothing but the mathematical expression of the fact that if the corpuscle manifests its presence in $\mathbf{R_p}$ (e.g., by triggering an observable phenomenon in it) then one must attribute the value $\mathbf{p}$ to the quantity of motion. That is precisely why the separation of wave trains will permit the measurement of the quantity of motion.

Briefly: After the measurement, upon considering an infinitude of corpuscles that are divided between the trains $\mathbf{R_p}$ according to the proportions $|\,c(\mathbf{p})\,|^2$ and then divided in the interior of each wave train $\mathbf{R_p}$ according to the density $|\,\Psi(\mathbf{p},\,\mathbf{r})\,|^2$, we will obtain a collection that is composed of individual entities that have perfectly-determined positions and quantities of motion, which is a collection that will correspond to the following statistical schema of classical type:

$$(\mathrm{II}) \qquad \begin{cases} \rho_{\mathbf{R}}(\mathbf{r}) = \int |\,c(\mathbf{p})\,|^2\,|\,\Psi(\mathbf{r})\,|^2\,d\mathbf{p}, \qquad \rho_{\mathbf{P}}(\mathbf{p}) = |\,c(\mathbf{p})\,|^2, \\ \rho(\mathbf{r},\mathbf{p}) = |\,c(\mathbf{p})\,|^2\,|\,\Psi(\mathbf{p},\mathbf{r})\,|^2, \\ \rho_{\mathbf{P}}^{(\mathbf{R})}(\mathbf{r},\mathbf{p}) = |\,\Psi(\mathbf{p},\mathbf{r})\,|^2, \qquad \rho_{\mathbf{R}}^{(\mathbf{P})}(\mathbf{r},\mathbf{p}) = \int_{\mathbf{R_p}} \delta(\mathbf{r}'-\mathbf{r})\,d\mathbf{r}', \end{cases}$$

and the following classical relations will exist between these quantities:

$$(\mathrm{II}') \qquad \begin{cases} \int \rho(\mathbf{r},\mathbf{p})\,d\mathbf{r} = \rho_{\mathbf{P}}(\mathbf{p}), \qquad \int \rho(\mathbf{r},\mathbf{p})\,d\mathbf{p} = \rho_{\mathbf{R}}(\mathbf{r}), \\ \rho_{\mathbf{P}}^{(\mathbf{R})}(\mathbf{r},\mathbf{p}) = \dfrac{\rho(\mathbf{r},\mathbf{p})}{\rho_{\mathbf{R}}(\mathbf{r})}, \qquad \rho_{\mathbf{R}}^{(\mathbf{P})}(\mathbf{r},\mathbf{p}) = \dfrac{\rho(\mathbf{r},\mathbf{p})}{\rho_{\mathbf{R}}(\mathbf{p})}. \end{cases}$$

The relations (II′) are verified immediately, except for the penultimate one, which is written:

$$(16) \qquad |\,c(\mathbf{p})\,|^2\,|\,\Psi(\mathbf{p},\,\mathbf{r})\,|^2 = \int |\,c(\mathbf{p}')\,|^2\,|\,\Psi(\mathbf{p}',\mathbf{r})\,|^2\,d\mathbf{p}' \int_{\mathbf{R_p}} \delta(\mathbf{r}'-\mathbf{r})\,d\mathbf{r}'.$$

Here is how one proves this: The last integral in (16) will be equal to 1 if the point $\mathbf{r}$ is in $\mathbf{R_p}$, and then the quantity of motion will be equal to $\mathbf{p}$; it will be zero if $\mathbf{r}$ is not in $\mathbf{R_p}$. It will then result that the right-hand side of (16) reduces to $|\,c(\mathbf{p})\,|^2\,|\,\Psi(\mathbf{p},\,\mathbf{r})\,|^2$, where $\mathbf{p}$ will have the value that corresponds to the known position $\mathbf{r}$, in such a way that equation (16) will indeed be verified.

We remark that, neither of the two probability distributions $\rho_{\mathbf{R}}(\mathbf{r})$ and $\rho_{\mathbf{P}}(\mathbf{p})$ will be "hidden" for the collection after the measurement. This will result from the fact that, on the one hand, the distribution $\rho_{\mathbf{R}}(\mathbf{r})$ that relates to the position is, as we know, always

verifiable by a statistical experiment (for example, with the aid of the screen-sieve device), and that on the other hand, the distribution $\rho_{\mathbf{P}}(\mathbf{p})$ will result from the same action of the device that permits one to measure **p**. By contrast, the probability distribution that corresponds to a quantity in the final collection that is not simultaneously measurable with either the position or the quantity of motion will have a perfectly-determined value, but it will remain "hidden," since the measurement of that quantity will destroy the collection. One then sees that in the theory of the double solution-pilot wave, any state will correspond to a well-defined collection, but there will always exist quantities whose probability distribution is hidden, because their measurement would destroy the collection ([1]).

It would now be very interesting to compare the probability distributions (I) and (II) that correspond to the collection that is realized before the measurement of **p** and the collection that is realized after the measurement of **p**, respectively.

First of all, the comparison of $\rho_{\mathbf{R}}(\mathbf{r})$ in (I) and in (II) will show us that we recover all of the usual formulas for the interference of probabilities.

On the other hand, the distribution $\rho_{\mathbf{R}}(\mathbf{r}) = |\Psi|^2$ in (I) and $\rho_{\mathbf{P}}(\mathbf{p}) = |c_p|^2$ in (II) are both the ones that are considered in the usual probabilistic formalism. However, here one will neatly see that they refer to the various collections that are realized, one of them, *before* the measurement, and the other one, *after* the measurement. That is the reason why these probability distributions cannot verify the usual schema of statisticians, which assumes the existence of a unique collection.

One now sees very neatly the flaw that vitiates the proof of the celebrated theorem of von Neumann. His argument shows that is it indeed impossible, even by introducing hidden variables, to construct a collection that simultaneously corresponds to the probability distributions $|\Psi|^2$ and $|c_p|^2$ that are habitually envisioned for the conjugate canonical quantities of "position" and "quantity of motion." However, it does not prove that one cannot construct the collections (by partially-hidden probability distributions) that are of the usual type and which correspond to the initial state before the measurement and the final state after the measurement, respectively, by introducing hidden variables. The usual probabilities that are considered will figure in these collections, but not in the *same* collection. We just constructed the collection in question in detail in the case of the measurement of **p**, and we now indeed perceive that von Neumann's theorem does not have the weight that one usually attributes to it.

4. Interpretation of the uncertainty relations. – In the usual interpretation, the fact that the probability distributions for a coordinate x and the conjugate momentum p_x will correspond to "dispersions" $\sigma(x)$ and $\sigma(p_x)$, such that:

$$\sigma(x)\,\sigma(p_x) \geq \frac{h}{4\pi}, \tag{17}$$

([1]) One must remark that each region $\mathbf{R_p}$ will be occupied by a "wave group" that is representable by a proper differential and will correspond, not to a completely exact value of **p**, but to extremely close values of **p**, in such a way that the uncertainty relations will remain valid for each wave group $\mathbf{R_p}$.

which will permit one to write the qualitative relation:

$$\delta x \, \delta p_x \geq h, \tag{18}$$

where δx and δp_x are the uncertainties in the values of x and p_x, must be interpreted by saying that the uncertainties in the values of x can never be both zero at the same time. However, from that prudent assertion, one generally proceeds to a much more audacious assertion. One assumes that quantities such as δx and δp_x do not represent simple experimental *uncertainties* in the values of these quantities that result from the conditions themselves that relate to the measurements of microphysics, but that they correspond to true *indeterminacies* that will always affect a part of the quantities that relate to a corpuscle. In that manner of looking at things, which seems to have been assumed implicitly by the authors, one is then obliged to consider the corpuscle as being statistically divided between the various states, which will produce an image of the corpuscle that is much less intelligible.

On the contrary, with the theory of the double solution, in which the probability distributions $\rho_{\mathbf{x}}(x)$ and $\rho_{P_x}(p_x)$ refer to different states, the interpretation of the uncertainty relations will no longer be the same. In each state, the corpuscle will have a position in the wave that is well-defined and a quantity of motion that is likewise well-defined as a function of position by the guidance. All of the quantities that relate to the corpuscle that are all expressed with the aid of the coordinates of the corpuscle and therefore the corresponding Lagrange momenta will also have well-defined values at each instant. However, these values cannot all be known simultaneously. Indeed, with the exception of position and the quantities that are measurable at the same time as position, the action of the device that will be necessary to measure a quantity A will have the effect of sending that corpuscle into one or the other of the wave trains that finally correspond to a given value of A in a perfectly-well-defined manner, in such a way that there will be an *a priori* uncertainty in the result of the measurement of A that provides the uncertainty in the (hidden) position of the corpuscle in the initial wave train, which is an uncertainty that corresponds to the probability distribution $|\Psi|^2$. The dispersion $\sigma(p_x)$ of the possible *final* values of p_x after the measurement will be related to the dispersion of the possible *initial* values of the position by the relation (17), and one can say, in that sense, that the position and quantity of motion of a corpuscle will always be affected with uncertainties δx and δp_x such that the relation (18) is verified.

However, from our standpoint, these are only uncertainties in the possible results of two incompatible measurements (viz., the measurement of position and the measurement of the quantity of motion) that demand different measuring devices. They are not both real indeterminacies in the position and quantity of motion of a corpuscle at each instant. These indeterminacies (at least, for quantities other than position) are produced by the action of the measuring device on the wave phenomena to which the corpuscle in incorporated (i.e., the wave v, not the wave Ψ, although the two are proportional). Since the coupling of the corpuscle to its wave involves Planck's constant essentially (whose true physical significance will remain mysterious), one can remain in agreement with Bohr, who said that the Heisenberg uncertainties are an inevitable consequence of the quantum of action, by the intermediary of the reaction of the wave propagation on the

motion of the corpuscle, but he reiterated that there does not result a true indeterminacy in the position and quantity of motion of the corpuscle in a necessary fashion.

The Heisenberg uncertainties will keep all of their value in the theory of the double solution, but they must be interpreted with more prudence than one ordinarily invests.

Moreover, some analogous considerations are applicable to the notion of complementarity. In the theory of the double solution, one can, if one so desires, preserve it, but on the condition that one indeed limits its scope. It will signify only that the same measurement operation cannot simultaneously provide the values of quantities that describe the wavelike aspect (such as the components of the quantity of motion) and quantities that describe the corpuscular aspect (such as the coordinates of the corpuscle). However, one cannot conclude from this that these non-simultaneously measurable quantities do not have a perfectly-well-defined value at each instant. Thus limited, the notion of complementarity will not raise any essential difficulties. It will no longer have the much less intelligible significance that one generally attributes to it, and as a result of which, what we call a “corpuscle” will be a protean entity that is capable of alternately taking on a wave-like aspect and a granular aspect, as a result of processes that will defy any rational representation.

CHAPTER VII.

MEASUREMENT OF QUANTITIES BY THE INTERACTION OF TWO CORPUSCLES.

---

1. The inconvenience of the measurement that was envisioned previously for just one corpuscle. – Much earlier, we envisioned the measurement of a quantity A that is effected by sending the wave train that carries the corpuscle to a measuring device that is capable of chopping the initial wave train into spatially-separated wave trains that each correspond to a well-defined value of the quantity A; i.e., of materially producing the spectral decomposition that relates to A.

However, this measuring process has an inconvenient aspect to it. When the corpuscle has triggered an observable macroscopic phenomenon in the region R_j (which is indispensible if one is to measure anything), the observation can attribute the value α_j to the quantity A. However, it is probable that when the observer has obtained the value for A, it will not be exact. Indeed, the triggering of the observable macroscopic phenomena will, in general, cause a reaction in the motion of the corpuscle, and the quantity A will, in turn, no longer have the same value α_j that it had before the triggering.

It is therefore preferable to proceed in a different manner, and to perform the measurement by appealing to an interaction between the corpuscle "under study" and another corpuscle that we call the "indicator" corpuscle. Whereas the measuring processes that were studied in the preceding chapter can be called "measurements of the first kind," the ones that we shall study can be called "measurements of the second kind." Let 1 denote the corpuscle under study and let 2 denote the indicator corpuscle. To begin with, the two corpuscles – which are, moreover, interacting – are attached to wave trains that occupy regions $R_0^{(1)}$ and $R_0^{(2)}$ that are separated in space and are represented by wave functions $\Psi_0^{(1)}(\mathbf{r}_1, t)$ and $\Psi_0^{(2)}(\mathbf{r}_1, t)$, resp. The wave function of the system in configuration space will then be:

$$\Psi_0(\mathbf{r}_1, \mathbf{r}_2, t) = \Psi_0^{(1)}(\mathbf{r}_1, t)\Psi_0^{(2)}(\mathbf{r}_2, t). \tag{1}$$

Following von Neumann, we assume that in order for a process to serve as a measurement of a quantity A that relates to the corpuscle under study, it will be necessary for the final wave Ψ to be of the form:

$$\Psi = \sum_k c_k \varphi_k(\mathbf{r}_1)\chi_k(\mathbf{r}_2), \tag{2}$$

where the φ_k are proper functions of the quantity A of the corpuscle under study that correspond to the proper values α_k, and the χ_k are proper functions of the indicator corpuscle that relate to a quantity B for that corpuscle that has proper values β_k .

Therefore, after the action of the measuring device, the values of A and B will be "correlated," so the value α_k of A will correspond to the value β_k of B.

However, that will not suffice. One must further have that after the end of the measuring interaction, the functions $\chi_k(\mathbf{r}_2)$ must correspond to wave portions that are spatially-separated, and consequently occupy disjoint regions $R_k^{(2)}$ in space. If the indicator corpuscle triggers an observable macroscopic phenomenon in the region $R_j^{(2)}$ then we can assert that its quantity B will have the value β_j, *and as a result*, the quantity A of the corpuscle under study will have the value α_j that is correlated with β_j. The probability for us to find that $A = \alpha_j$ will be, moreover, equal to $|c_j|^2$.

One sees the advantage of that measurement of the second kind over the measurement of the first kind. The triggering of observable macroscopic phenomena by the indicator corpuscle can perturb the motion of that corpuscle and, in turn, make its quantity B no longer have the value β_j. However, since the two corpuscles are entirely independent and separated in space at the end of the interaction, the triggering of the observable phenomena in $R_j^{(2)}$ can have no influence on the corpuscle under study, and one can state with certainty that the quantity A of the corpuscle under study has the value α_j after one has confirmed the observable phenomenon.

Moreover, here, as with the measurement of the first kind, it is not necessary to introduce the coordinates of the "measuring apparatus" into our analysis, whose role can only be that of permitting us to observe the observable macroscopic phenomenon with precision.

We remark that in order to make a measurement of A, it is not necessary that the functions $\varphi_k(\mathbf{r}_1)$ must correspond to spatially-separated regions $R_k^{(2)}$ in the final state. What will be indispensible is that the regions $R_k^{(2)}$ must be disjoint. However, for more clarity, we shall first suppose that the regions $R_k^{(2)}$ are separated, while we shall return later on to the case in which that hypothesis is not realized. We will then have the following schema:

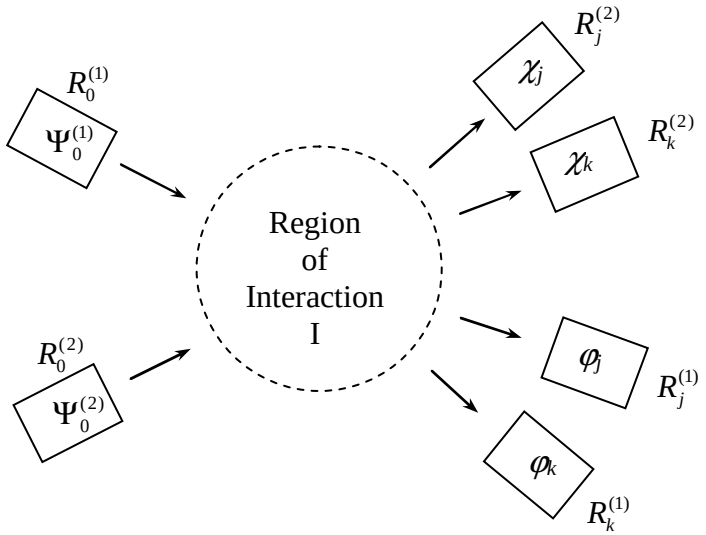

Figure 4.

Before the interaction, the two wave trains $R_0^{(1)}$ and $R_0^{(2)}$ are separated and independent. They then approach each other and enter into interaction in region I. After the interaction, the wave function in configuration space will, by hypothesis, have the form (2), so the various functions φ_k will likewise correspond to spatially-separated wave trains $R_k^{(1)}$ in physical space, and the various functions χ_k will likewise correspond to spatially-separated wave trains $R_k^{(2)}$ in physical space. The statistical correlations that are established by the interaction between the presences of the two corpuscles in the different regions *R* in physical space and translate into the form (2) of the wave function will then tell us that if an observable, macroscopic phenomenon is triggered in the region $R_j^{(2)}$ by the indicator corpuscle then the corpuscle under study will be necessarily found in $R_j^{(1)}$ with $A = \alpha_j$.

We shall specify an example of this measuring process in order to show that it indeed corresponds to experimental conditions that are often realized. Suppose that we initially have two corpuscles whose energies and quantities of motion have known values. The two wave trains $R_0^{(1)}$ and $R_0^{(2)}$ will be in interaction (e.g., collision) in a neighborhood of a point *O*, which we will take to be the origin of the coordinates. After the interaction, the correlated wave trains $R_j^{(1)}$ and $R_j^{(2)}$ will be distant from the point *O* in directions that are defined by the angles φ_1 and φ_2 with respect to an axis *Ox* that is contained in the symmetry plane of the phenomenon.

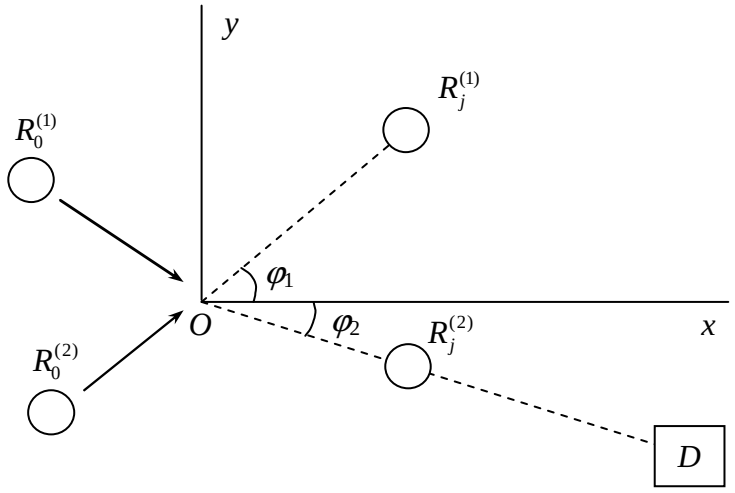

Figure 5.

All of the phenomena are produced in the symmetry plane *xOy*, so one knows that the correlation between $R_j^{(1)}$ and $R_j^{(2)}$ will be expressed by three conservation relations for the energy and the two *x* and *y* components of the quantity of motion. We thus have three relations in the final state between the angles φ_1 and φ_2 , and the magnitudes p_1 and p_2 of the quantities of motion. If the corpuscle 2 then arrives at a device *D* (for example, a scintillation counter) where it produces an observable, macroscopic phenomenon then one can determine the angle φ_2 , and one will obtain p_1 and φ_1 by eliminating p_2 from the

three conservation equations, ; i.e., one will have measured the magnitude and direction of the final quantity of motion of the first corpuscle, and that measurement will obviously exert no effect on the corpuscle 1 under study, since that corpuscle can be found to be very distant from the device *D* at the moment when it produces the observable macroscopic phenomenon in it.

2. Interpretation of the measurement of the second kind by the usual theory. – How are we to interpret the measurement of the second kind in the current wave mechanics, where the corpuscle is not localized in the wave Ψ?

The interpretation that conforms to the von Neumann-London-Bauer theory of measurement consists in saying: It is the observer's perception of the macroscopic phenomenon that is triggered in $R_j^{(2)}$ that will localize the corpuscle 1 in $R_j^{(1)}$ that was hitherto divided statistically between all of the wave trains $R_k^{(1)}$. Such an interpretation seems inadmissible: Something that happens in the perception of an observer cannot provoke a physical effect at a distance.

Suppose that there are two observers in the region $R_j^{(2)}$, one of whom has his eyes open, and is thus aware of the observable macroscopic phenomenon, while the other one has his eyes closed, so he is not aware. Does the awareness of the first observer provoke the localization of the corpuscle 1 in $R_j^{(1)}$, or does the unawareness of the second observer prevent that localization? The question remains unanswered, because it is absurd. However, one thing seems certain: The observer who has his eyes open, after having confirmed that a macroscopic phenomenon has been triggered by the corpuscle 2, will replace the wave function $\Psi = \sum_k c_k \varphi_k(\mathbf{r}_1)\chi_k(\mathbf{r}_2)$ with the new wave function $\Psi = \varphi_k(\mathbf{r}_1)\chi_k(\mathbf{r}_2)$, and that "reduction of the probability packet" will permit him to then make exact statistical predictions. Therefore, what really matters is not the knowledge of the observer, but the physical fact that the triggering of the observable phenomenon constitutes.

An interpretation that might appear to be more reasonable will then consist of saying: It is the observable, macroscopic phenomenon that is provoked by the corpuscle 2 in $R_j^{(2)}$ that briefly localizes the corpuscle 1 under study in the correlated region $R_j^{(1)}$. In reality, that interpretation is no more admissible than the previous one. The observable phenomenon that is manifested in $R_j^{(2)}$ can in no way localize the corpuscle 1 in $R_j^{(1)}$, and that will be all the more the case when the regions $R_j^{(1)}$ and $R_j^{(2)}$ are extremely distant from each other at the moment when that phenomenon is produced. For such a localization to be produced briefly at no particular distance by the phenomenon that is observed in $R_j^{(2)}$ would be inconceivable. While discussing that objection, Schrödinger wrote: "That would be magic!" and indeed that would be magic.

In the final analysis, it thus seems clear that in a theory where corpuscles are not localized in their waves, no reasonable interpretation can be given for correlations that are represented by the wave Ψ in configuration space and which permit measurements.

We shall see that the theory of the double solution will provide a clear and intelligent interpretation for measurements of the second kind by establishing the position of the corpuscle in the wave.

3. Interpretation by the theory of the double solution. – Return to Fig. 4. In the theory of the double solution, we must suppose that the regions $R_0^{(1)}$ and $R_0^{(2)}$ are initially occupied by the singular regions of the waves *u* of each of the two corpuscles, respectively. Upon arriving in I, the two wave trains will begin to integrate; i.e., the propagation of each of them will depend upon the action that they exert upon each other. As I pointed out at the end of chapter V, I will assume (while hoping that this hypothesis can be justified rigorously in the theory of the double solution) that the Schrödinger wave Ψ in configuration space of two corpuscles will permit one to represent precisely the correlations between the possible positions of the corpuscle in the course of time and at the end of the interaction. That will be a reason for why the wave Ψ in configuration space, although obviously fictitious and representing no real evolution of the phenomenon in physical space, still gives a statistically exact view of the *possible* results of the interaction.

We are thus led to think that after the end of the interaction the wave *u* of corpuscle 1 will have been chopped into a series of wave trains $R_0^{(1)}$, …, $R_k^{(1)}$, … that are spatially-separated, and that, similarly, the wave *u* of the corpuscle 2 will have been chopped into a series of wave trains $R_1^{(2)}$, …, $R_k^{(2)}$, … that are likewise disjoint. Moreover, since the statistical correlations that are represented by the final form (2) of the wave Ψ are supposed to be exact, if the singular region that constitutes corpuscle 2 finally arrives in $R_j^{(2)}$ then the one that constitutes corpuscle 1 will arrive in $R_j^{(1)}$, and the probability of that eventuality will be given by $|c_j|^2$.

In other words, in the theory of the double solution the two corpuscles-singular regions will have trajectories that are determined entirely by their initial positions in the wave trains $R_0^{(1)}$ and $R_0^{(2)}$, which are trajectories that will necessarily lead them to occupy two well-defined positions in the two "correlated" wave trains $R_0^{(1)}$ and $R_0^{(2)}$ in the final state. Therefore, according to a wish that was often expressed by Einstein, the description of the interaction by the wave Ψ in configuration space will remain an exact theory, but it will not be a complete description, which would be the theory that establishes the localization of corpuscles and the determinism in their motion.

Here, the interpretation of the measure of the second kind becomes quite clear and can be expressed in several ways. If an observable, macroscopic phenomenon is triggered by corpuscle 2 in the region $R_j^{(2)}$ then that will be, quite simply, because that corpuscle is effectively found to be in that region, and then the corpuscle 1 under study will necessarily be found in the correlated region $R_j^{(1)}$, which will permit us to attribute the value α_j to the quantity *A* with some confidence. Here, there is no inadmissible instantaneous action-at-a-distance in the observed phenomena. We are relieved to see that: "There is no longer any magic!"

As for the perception of the observable phenomenon by the observer that permits one to attribute the value α_j to A, it is only the perception of an external objective reality, and it will thus recover the completely reasonable meaning that it has in classical physics.

We would like to insist a little bit on the manner by which the question of the relationship between the wave u and Ψ should present itself here. In the regions $R_0^{(1)}$ and $R_0^{(2)}$, the wave u of each corpuscle will have the form $u = u_0 + v$ outside of a very small singular region, where $v \gg u_0$ is a regular wave. The wave u will have a perfectly-defined amplitude, since it has objective reality, but the physicist will have the right to define a *fictitious* wave function Ψ by setting $\Psi = Cv$ and choosing the constant C in such a manner that Ψ will be normalized. One will then define the individual functions Ψ in $R_0^{(1)}$ and $R_0^{(2)}$ for the two corpuscles in the initial state, and then one will construct the wave function in configuration space by taking the product of the two individual wave functions.

In the final state, after the end of the interaction, the wave u of the corpuscle 1 will have been chopped in physical space into wave trains $u_1^{(1)}$, …, $u_k^{(1)}$, … that occupy spatially-disjoint regions $R_1^{(1)}$, …, $R_k^{(1)}$, …, while the wave u of corpuscle 2 will likewise have been chopped into wave trains $u_1^{(2)}$, …, $u_k^{(2)}$, … that occupy spatially-disjoint regions $R_1^{(2)}$, …, $R_k^{(2)}$, … in physical space. However, corpuscle 1 will only be found in one of the regions $R_k^{(1)}$. The part u_0 of u will exist in only that region, and u will reduce to v in the other regions $R^{(1)}$. An analogous situation will be realized for the second corpuscle. However, the wave function (2) must represent the statistical correlations in the final state exactly, so the two corpuscles will certainly be found in the two correlated wave trains. When the observer confirms the triggering of an observable, macroscopic phenomenon in $R_j^{(2)}$ by the indicator corpuscle, he will say that the corpuscle under study is in $R_j^{(1)}$, and he must then construct a new individual wave function for corpuscle 1 by setting $\Psi = Cv_j^{(1)}$, where $v_j^{(1)}$ is the function v for corpuscle 1 in $R_j^{(1)}$ and C is chosen in such a fashion that Ψ will be normalized in $R_j^{(1)}$. This is the rupture of the statistical correlations that von Neumann spoke of, and there is also reduction of the probability packet, since the wave Ψ of corpuscle 1, instead of being divided between all of the regions $R_j^{(1)}$, is briefly reduced to no longer occupying that $R_j^{(1)}$.

Here, the reduction of the probability packet takes on a very clear significance. It signifies simply that the observer, having acquired information about physical reality, consequently modifies the function Ψ that serves to represent the probability of the observable phenomenon. However, the information that is acquired by the observer cannot modify the physical reality itself that is described by the wave u. The distinction between objective and subjective is thus established in a completely clear and satisfying fashion.

We remark that if the "subjective" function Ψ can provide exact statistical predictions then that will be because the user must construct it to be proportional to the wave u (or rather, to its part that is "exterior" to the singular region, by abstracting from that singular region) *in the region where one knows that the corpuscle can be found.* An open-eyed

observer who has confirmed the triggering of the observable, macroscopic phenomenon in the region $R_j^{(2)}$ will employ the function $\Psi = Cv_j^{(1)}$ for the corpuscle under study, and can make exact statistical predictions with it. The closed-eyed observer, who will observe nothing, will utilize a function Ψ for corpuscle 1 that will remain proportional to $v_k^{(1)}$ in each region $R_k^{(1)}$, and he will make inexact statistical predictions with it, since he assumes that corpuscle 1 can be found in regions $R_k^{(1)}$ other than $R_j^{(1)}$, which is not true.

One then sees that if the subjective wave Ψ can be of service then that will be due to the fact that it is constructed by the user as a function of his knowledge of a certain exterior, physical reality. It is, moreover, obvious that a subjective wave function can be constructed arbitrarily, and that it will then be incomprehensible that it will lead to exact predictions. In other words, a purely subjective interpretation of the wave Ψ is impossible; it is only necessary that there be some objective reality behind it. Now, the wave Ψ, which undergoes the reduction of the probability packet, cannot itself have objective reality, but it can be a reflection of the user's knowledge of objective reality.

We add that in the case of the measurement by interaction of two corpuscles, as we did in the case of the measurement of the first kind with just one corpuscle that traverses a device, one can construct collective systems that correspond to the initial and final states and represent the concepts of localization and motion of the corpuscles in the theory of the double solution. As in the preceding chapter, one will recover the probability distributions that are usually envisioned, but they will belong to different collective systems, with the consequences that this implies. One will also recover the interference of probabilities and the interpretation of the Heisenberg relations as representing, not an indeterminacy in the position and motion of corpuscles, but the uncertainties that are introduced by the intervention of the wave nature of corpuscles in any measurement process.

We further point out a difficult, but important, question that is posed when one applies the concepts of the double solution to the problem that we just studied. When the observer has observed the phenomenon that is produced in $R_j^{(2)}$, he will know that the corpuscle or singular region of the wave $u^{(1)}$ is found in $R_j^{(2)}$, and he will reconstruct his wave $\Psi^{(1)}$ as a consequence, in order to make ulterior statistical predictions. However, the fragments of the wave $u^{(1)}$ that go into regions $R_k^{(1)}$ other than $R_j^{(1)}$ must persist, since, having objective reality, they cannot depend upon the information of the observer. They will constitute pieces of the wave $u^{(1)}$ of the corpuscle under study *that do not contain the singular region.* What do these fragments of the wave u with no singular region become then, and how, on the other hand, does the wave fragment $u^{(1)}$ that arrives in $R_j^{(1)}$ and carries the singular region evolve, which is a wave fragment whose exterior part $v^{(1)}$ was weakened with respect to what it was in the initial state by chopping up the wave $u^{(1)}$? These questions pertain to a type of difficult question that the theory of the double solution is forced to address. However, it is probable that if it succeeds in answering them, then that will be by making nonlinear phenomena intervene essentially, and in

particular, of transitory states that have representation in the present linear theory ([1]). However, I would not like to insist here upon a problem that is, moreover, difficult to undertake today.

Finally, note that the analysis that we just made of the measuring process by the interaction of two corpuscles can be, it seems to me, transposed into the study of two systems of corpuscles with no difficulty. The complications that one might encounter will be only complications of notation, while the general ideas will remain the same.

4. Case of a measurement process involving the interaction of two corpuscles, whose singular regions $R_k^{(1)}$ are not spatially disjoint. – We have already remarked that since the function Ψ in configuration space is assumed to have the form (2) after the end of the interaction of the two corpuscles, it will suffice that the wave trains $\chi_k(\mathbf{r})$ of the indicator corpuscle are spatially-separated in order for the triggering of the observable, macroscopic phenomena in $R_j^{(2)}$ to permit us to attribute the value α_j to the quantity A without the $R_k^{(1)}$ being spatially separated.

It is easy to give a concrete example: Let there be a hydrogen atom whose peripheral electron will play the role of the corpuscle under study. The wave of that electron will occupy a region of the atom that we will represent by a spherical region $R^{(1)}$. Another corpuscle that plays the role of indicator then passes in the neighborhood of the atom, and during that passage, it will interact with the atomic electron. The wave train of that indicator corpuscle will be initially contained in a region $R_0^{(1)}$ in space. At the end of that interaction, the wave function Ψ in configuration space of the two corpuscles will take the form (2), by hypothesis, and the χ_k now correspond to spatially-disjoint regions $R_k^{(2)}$, while the φ_k will always correspond to the same region $R^{(1)}$.

In its initial state, the atomic electron has a wave function of the form $\Psi_0^{(1)} = \sum_k c_k^0 \varphi_k(\mathbf{r})$, where the φ_k are the proper functions of a certain quantity A, which might be energy, for example. When one has noted an observable, macroscopic phenomenon that is triggered by the indicator corpuscle in $R_j^{(2)}$, one can attribute the value α_j to A and take $\Psi^{(1)}$ to be equal to φ_j in that final state.

Here, one recovers the same considerations as in the preceding case that was studied. There can be no question of saying, as one does in the present interpretation, that it is the perception of the observable phenomenon by the observer or the triggering of observable phenomena that makes the atom pass briefly from its initial state to its final state; this will always be magical. The perception of the observer has nothing to do with this case, and $R_j^{(2)}$ can be very distant from $R^{(1)}$, so an instantaneous influence of the phenomenon that is triggered on anything that happens in $R^{(1)}$ would be inconceivable.

One can illustrate this by a striking example, which is not described in the same fashion as the preceding one in the present state of our formalism, but which is

([1]) I have touched upon these questions in the last chapter of my recent book [**3**]. *See* the very interesting recent book by Jean-Louis Destouches [**7**], as well.

completely analogous to it physically, namely, the example of the emission of a photon by an atom (a hydrogen atom, for example). If the quantity A is the energy of the atomic electron then the $\alpha_k = E_k$ and the φ_k will be the corresponding proper values and proper functions, respectively. We assume that the atom is initially in a quantized state of energy E_k and that, in turn, its wave function is $\Psi_0^{(1)} = \varphi_k$. If it passes through a quantum transition from that initial state to the quantized state of energy E_l then there will be emission of a photon of frequency $\nu = (E_k - E_l) / h$. Collect the emitted photon at a distance in a mesh + lens device that will, as we have seen, permit one to attribute a well-defined frequency to it by making it produce an observable effect – for example, the darkening of a photographic plate at a point. From the observation of that phenomenon, the observer can deduce that the atom has passed from the initial state $\Psi_0^{(1)} = \varphi_k$ to the final state $\Psi^{(1)} = \varphi_j$.

Now, nothing prevents one from supposing that the atom is on the star Sirius and that the observer is in Paris. Is it admissible that the perception by a Parisian observer of a local darkening on photographic plate will change the state of an atom on Sirius? Is it admissible that the physical triggering of this darkening produced that effect? Obviously no, in both cases, and the argument is, just the same, very striking here.

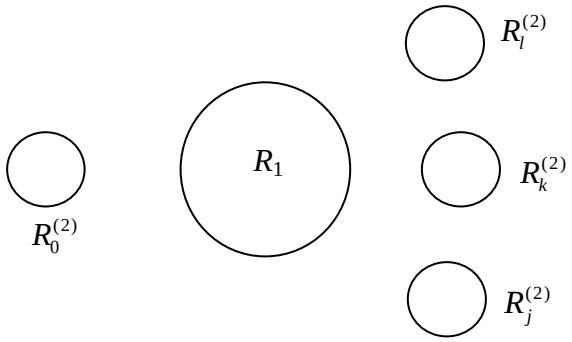

Figure 6.

We now refer to Fig. 6, and demand to know how one can interpret the measurement of A in this case using the ideas of the double solution. That will lead us to a new concept. Since the function Ψ, when it is well-constructed, must always be proportional to ν, one must have $\nu_0^{(1)} \sim \sum_k c_k^0 \varphi_k$ in the initial state of the atom, and one must have $\nu^{(1)} \sim \varphi_j$ in the final state – viz., when the indicator corpuscle arrives in the region $R_l^{(0)}$. Therefore, according to the concept of the double solution, the atomic electron must have a motion in the initial state that conforms to the guidance formula, which is a very complicated motion that has the same phase as the wave $\nu_0^{(1)}$ that is defined by the superposition of the φ_k. However, after the end of the interaction, one must have a motion that is in phase with the single component φ_j, since one will then have $\nu^{(1)} = \varphi_j$. One can then say that during the interaction the motion of the atomic electron progressively "switched" from "unhooking" from the initial superposition of the φ_k to finally "hooking onto" the single component φ_j, and that is precisely because the atomic electron is finally found to be hooked to $\nu^{(1)} = \varphi_j$, while *correlatively* the indicator

corpuscle is found in $R_j^{(2)}$, where there is a possibility of measuring the quantity A whose proper value α_j corresponds to φ_j.

Here then, the corpuscle being studied can be finally found to be hooked to just one of the original components of its wave v as a result of the interaction, all while remaining localized in the same region $R^{(1)}$ as in the beginning, but it is essential for the measurement of A that the regions $R_k^{(2)}$ that relate to the indicator corpuscle be spatially-separated in such a fashion that the triggering of an observable, macroscopic phenomenon will permit one to say that it is the wave $v^{(1)} \sim \varphi_j$ to which the corpuscle under study will finally remain hooked. Naturally, the same thing is not true in the case of a measurement of the first kind, where one performs the measurement by making the corpuscle that is incorporated into its wave pass into a device – such as a mesh + lens – that will isolate the φ_k into spatially-disjoint wave trains. In that case, the corpuscle is, in some way, both the corpuscle under study and the indicator corpuscle, and that is the reason why the separation of the wave trains φ_k will then be necessary. However, that will no longer be true for a measurement of the second kind.

5. The idea of directing. Examination of a remark by Einstein. – While always remaining in the case that was envisioned in the last paragraph, and maintaining the viewpoint of the causal interpretation, we can make the preceding more precise in the following fashion: Starting with the initial position of the two singular regions in $R^{(1)}$ and $R_0^{(2)}$, the interaction evolves in an entirely determinate manner in such a way that in the course of the interaction, the corpuscle under study is progressively directed by the motion that the guidance law imposes upon it to its final state, where it is found to be implanted in the wave $v = \varphi_j$, while the indicator corpuscle is progressively directed to its final state in the same way, where it is implanted on the wave $v = \chi_j$ that is localized in $R_j^{(2)}$. The same interpretation will be valid, *mutatis mutandis*, in the case that was previously studied where the regions $R_k^{(1)}$ were disjoint. One will then arrive at the general idea that for each corpuscle, the streamlines will be animated with a sort of wiggling that is a result of the interaction, and that the corpuscle, which is obliged to constantly follow one of the wiggling streamlines by the guidance law, is therefore progressively directed towards the final state that it will possess at the end of the interaction.

In light of these ideas, it is interesting to examine a point in the present formalism of wave mechanics that Einstein touched upon many times, and which seemed to him to appear particularly difficult to interpret by a causal theory.

Consider a corpuscle whose initial state corresponds to the function $\Psi = c_j\varphi_j$, in which $| c_j | = 1$ and φ_j is one of the proper functions of a quantity A. If A is energy then we can represent the initial state by a wave train R_0 that corresponds to the energy $E_j = \alpha_j$ (mathematically, by a proper differential of the continuous spectrum of energy that corresponds to the central frequency $\nu_j = E_j / h$). If the corpuscle is subjected to a very weak perturbing field of limited duration then its wave Ψ will become $\Psi = \sum_k c_k\varphi_k$ after the end of the perturbation, and since it is very weak, one will have a final value for $| c_j |$

that is very slightly less than 1 and extremely small values for the $| c_k |$ when $k \neq j$. If the corpuscle then passes into a device D the sends the φ_k into spectrally-disjoint regions R_k (such as the mesh-lens device, in the case where A is energy) then the corpuscle will have a probability that is close to 1 of being manifested in the regions R_k with $k \neq j$. Since the α_k can have very different values, one will finally see that the very slight perturbation will make very small probabilities appear for the state of the corpuscle to be subjected to large changes. Einstein considered that one will thus obtain a statistical description that is certainly exact from the standpoint of what takes place, but that it will be, without a doubt, very difficult to substitute a causal description for the individual phenomena for it, which is a description that will nonetheless be necessary in order to avoid unacceptable paradoxes.

Recall this problem from the viewpoint of the theory of the double solution. Consider a wave train R_0 that carries a corpuscle. The corpuscle occupies a certain position in this wave train. It is implanted in a wave $v \sim \varphi_j$ that has objective reality and fills up the region R_0. We suppose that the wave train is a group of waves that is almost monochromatic with energy E_j, and whose streamlines are parallel lines. The wave trains will be directed towards a device D that permits the measurement of energy by separating the wave trains that correspond to the different values of energy. If the wave train does not traverse any perturbing field between its initial position R_0 and its arrival at D then the corpuscle that follows one of the streamlines will arrive at the region D by a uniform, rectilinear motion, and then the action of the device D will impose a more complicated motion upon it that will result in a wave train with energy E_j leaving the device. However, if the wave train traverses a region in which a small perturbing field reigns before arriving at D then the wave v will become proportional to $\sum_k c_k \varphi_k$ with $c_j = 1 - \varepsilon$ and all of the other c_k very small. In the wave train, thus modified, the streamlines that correspond to the guidance formula will be animated with a sort of very small trembling with respect to the rectilinear form that they will preserve in the absence of perturbation. It will then result from this that, according to the position in the wave train, the corpuscle will be sent after traversing the device D either towards the wave train R_j (this will be, by far, the most probable case) or towards the one of the wave trains R_k that corresponds to an energy $E_k \neq E_j$, but this will be a very rare case. We are sure that this will so because we know that there will be an infinitude of corpuscles in R_0 that are distributed with the density $| \varphi_j |^2$, so the motion along these streamlines that is imposed by the guidance formula will finally lead to a proportion that is equal to $| c_k |^2$ in the region R_k. Since the corpuscle will have an energy that is equal to E_k in the region R_k, one can say: The direction that is imposed upon the corpuscle by the necessity that it must always follow a streamline will, when one ignores its initial position in R_0, give it a probability that is almost equal to 1 of having preserved the initial value E_j of its energy at the end of the measuring process, but it will also give it a very small probability of finally possessing an energy E_k that is very different from E_j. One finds a picture that clearly represents the situation that Einstein touched upon in regard to the usual formalism of wave mechanics.

Nevertheless, if one would like to study the basis for the idea of directing that was sketched above then one will have to study the question of the conservation of energy in the theory of the double solution, as well as the analogous problems that we alluded to at

the end of paragraph 3. We will not undertake that study here, since it would be premature, moreover.

6. Conclusions. Pure case and mixture. – The study of measurement that we just carried out in chapter VI and VII shows us that one can envision measurement in microphysics as having a concrete aspect, which is, moreover, obvious, and does not require the overly-abstract formalism that one habitually assigns to it. It is essential to account for the fact that all of the information that we obtain about microphysical reality is triggered by the *local* action of a corpuscle. It is likewise essential to remark that the wave into which a corpuscle is incorporated is always extended over a bounded region of space. It exists only as bounded wave trains, so unbounded monochromatic plane waves in space and time are abstractions, as well as stationary waves that extend to infinity, moreover. As Schrödinger has quite rightly remarked, it is just that bounded character of wave trains that permits one to speak of the beginning and ending of an interaction. It is what permits one to construct devices that have the effect of spatially separating the wave trains that each correspond to a well-defined value of a quantity A, and thus, to perform measurements at the microphysical level. It then seems that this concrete conception of measurement is compatible with a permanent localization and a well-defined motion of the corpuscle, and the arguments of Einstein and Schrödinger, with all of their weight, show clearly that it is necessary that the corpuscle be localized with the wave if one is to avoid certain truly inadmissible consequences of the present interpretation.

To our way of looking at things, the distinction between the pure case and the mixture, which was quite rightly introduced by von Neumann, takes on a concrete significance that does not appear in the exact, but overly-abstract, formalism that was discussed in chapters II and III.

One has a "pure case" when the wave v (viz., the exterior part of the wave u) of a corpuscle is formed by a superposition of *interfering* components, and the corpuscle follows one of the streamlines that results from that superposition. On the contrary, one has a "mixture" when the various components of the initial wave v cease to interfere at the end of the interaction, either by a sequence of spatial separations of the wave trains or by a sequence of directing phenomena, in the sense of that we gave that word in the last paragraphs. The corpuscle is then found to be hooked to just one of these components, and until we know that, our ignorance will be represented by a mixture.

It is upon assuming that viewpoint – namely, carefully distinguishing the objective wave u from the subjective and predictive wave Ψ, and thus establishing a clear distinction between objective and subjective – that the formalism of von Neumann's theory seems to us to have been included and interpreted.

We shall, moreover, arrive at an analogous conclusion in the last chapter, which will be concerned with von Neumann's thermodynamics.

———

CHAPTER VIII.

GLIMPSE OF VON NEUMANN'S THERMODYNAMICS.

———

1. Introduction to von Neumann's formalism in thermodynamics. – First recall some points regarding classical statistical thermodynamics. Boltzmann established the celebrated relation:

(1) $$S = k \log P$$

between the entropy of a physical system and the probability of the state of that system, where k – viz., Boltzmann's constant – has the value 1.37×10^{-16} in C.G.S. units and the absolute temperature scale. The relation (1) has been confirmed by an immense volume of verifications that were deduced from its consequences.

If we consider a set of N systems that are distributed into a certain number of states, in the classical sense of the word, in such a way that there will be n_i states in the state i $\left(\text{with } \sum_i n_i = N \right)$ then one will easily find that the probability of that distribution will be:

(2) $$P = \frac{N!}{n_1! n_2! \cdots}.$$

Since N and the n_i are assumed to be large, Stirling's formula permits one to set, very approximately, $N! = N^N e^{-N}$ and $n_i! = n_i^{n_i} e^{-n_i}$, and one easily gets:

(3) $$\log P = \log N! - \sum_i \log n_i ! \approx N \log N - \sum_i n_i \log n_i .$$

Set $p_i = n_i / N$, where p_i is the statistical weight of the state i in the distribution. Since $\sum_i p_i = 1$, we will have:

(4) $$\log P = - N \sum_i p_i \log p_i + N \log N - N \sum_i p_i \log N = - N \sum_i p_i \log p_i .$$

As a result, from formula (1):

(5) $$S = - kN \sum_i p_i \log p_i ,$$

which is a classical formula from statistical thermodynamics.

If we would like to now construct quantum thermodynamics then we must modify the definition of entropy by replacing the classical concept of the state of a system with a wave function.

Recall the algorithms of chapter II. When the various states of N systems that we are considering are defined by wave functions φ_1, …, φ_i, … that define a complete, orthonormal system of basis functions (for example, the proper functions of a measurable quantity), one can convert von Neumann's statistical matrix P into its diagonal form by taking the φ_i to be basis functions; one will then have:

(6) $$P_{kl} = p_k \, \delta_{kl} \qquad \left(\sum_k p_k = 1 \right).$$

Moreover, the matrix log P, whose elements are $(\log P)_{kl} = \delta_{kl} \log p_k$, will also have a diagonal form. It is then natural to convert the Boltzmann formula by defining the entropy upon starting with the statistical matrix P and setting:

(7) $$S = - kN \, \mathrm{Tr}(P \log P),$$

because that expression, which has a value that is independent of the chosen system of basis functions, by reason of the invariance of the trace, will be expressed by:

(8) $$S = - kN \sum_k p_k \log p_k$$

in the basis system where the matrices P and log P are diagonal, in such a way that we come back to the old formula (5).

We shall seek to determine the maximum entropy when one supposes that one has been given the number of systems N and the value E of their total energy.

First, recall the calculation of that maximum in classical statistical thermodynamics. One can write $\delta \log P = 0$, with the conditions $\delta N = 0$ and $\delta E = 0$, which leads us to introduce the Lagrange multipliers α and β, and to write:

(9) $$\delta(\log P - \alpha N - \beta E) = 0.$$

Namely, from (4):

(10) $$- N \sum_i \delta p_i \, [\log p_i + 1 + \alpha + \beta E_i] = 0$$

for any variation of p_i such that $\sum_i p_i$ remains equal to 1 $\left(\text{because } \delta E = \sum_i E_i \delta n_i = N \sum_i E_i \delta p_i \right)$. One deduces from this that:

(11) $$p_i = e^{-\alpha - \beta E}.$$

This is the classical Boltzmann-Gibbs law, which can also be written:

$$p_i = \frac{e^{-\beta E_i}}{\sum_k e^{-\beta E_k}}, \tag{12}$$

when one takes the condition $\sum_i p_i = 1$ into account. If one compares this expression with the one in the theory of perfect gases then one can see that $\beta = 1 / kT$, where T is the absolute temperature of a set of N systems, which is assumed to be well-defined; for example, by contact with a thermostat.

One then easily finds that entropy of the most probable distribution (which is almost always realized) will be:

$$S = -kN \sum_i p_i \log p_i = kN \left[\log \sum_i e^{-\frac{E_i}{kT}} + \frac{1}{kT} \frac{\sum_i E_i e^{-\frac{E_i}{kT}}}{\sum_i e^{-\frac{E_i}{kT}}} \right]. \tag{13}$$

Set $Z = \sum_i e^{-\frac{E_i}{kT}}$; this is Planck's "state sum." We have:

$$\begin{cases} S = kN \left[\log Z - \beta \dfrac{\partial \log Z}{\partial \beta} \right], \\ E = -N \dfrac{\partial \log Z}{\partial \beta}, \\ F = E - TS = -kNT \log Z \end{cases} \tag{14}$$

as expressions for entropy, energy, and free energy.

In von Neumann's quantum thermodynamics, the calculations are completely parallel to the preceding one. One must express the idea that the entropy (7) is a maximum under the conditions:

$$\mathrm{Tr}\, P = 1, \qquad E = NE - N\, \mathrm{Tr}(PH) = \text{const.}, \tag{15}$$

where H is the Hamiltonian matrix of any of the N systems. One must then write:

$$\delta \sum_k P_{kk} \log P_{kk} = 0, \quad \text{with} \quad \delta \sum_k P_{kk} = 0, \quad \delta \sum_{k,l} P_{kl} H_{lk} = 0, \tag{16}$$

so, upon introducing Lagrange multipliers α and β:

(17) $$\delta\sum_k P_{kk}\log P_{kk} + \alpha\delta\sum_k P_{kk} + \beta\delta\sum_{k,l} P_{kl}H_{lk} = 0,$$

which one can also write:

(18) $$\sum_k \delta P_{kk}[\log P_{kk} + 1 + \alpha + \beta H_{kk}] + \beta\sum_{k\neq l}\delta P_{kl}H_{lk} = 0,$$

which is a relation that must be verified for any value of δP_{kk} . It is then necessary that the system must be in an energy proper state ($H_{kl} = 0$ for $k \neq l$) and that:

(19) $$P_{kk} = e^{-\alpha-\beta H_{kk}},$$

moreover, with $e^{-\alpha}\sum_k e^{-\beta H_{kk}} = 1$, since $\sum_k P_{kk} = 1$. Thus, upon setting $Z(\beta) = \mathrm{Tr}\, e^{-\beta H}$, one will get:

(20) $$P = \frac{e^{-\beta H}}{\mathrm{Tr}\, e^{-\beta H}} = \frac{e^{-\beta H}}{Z(\beta)}.$$

As in the classical theory, one proves that $\beta = 1 / kT$, and one finds for the most probable state (while taking into account the fact that $\frac{\partial \log Z}{\partial \beta} = \frac{1}{Z}\frac{\partial Z}{\partial \beta} = -\frac{1}{Z}\mathrm{Tr}\, H e^{-\beta H}$):

(21) $$\begin{cases} S = \dfrac{k(N)}{Z(\beta)}\mathrm{Tr}[e^{-\beta H}(\beta H + \log Z)] = kN\left[\log Z + \dfrac{\mathrm{Tr}\,\beta H e^{-\beta H}}{Z}\right] \\ \quad = kN\left[\log Z - \beta\dfrac{\partial \log Z}{\partial \beta}\right], \\ E = N\,\mathrm{Tr}(PH) = -N\dfrac{\partial \log Z}{\partial \beta}, \\ F = E - TS = -kNT\log Z. \end{cases}$$

One thus recovers formulas (14) of classical statistical thermodynamics, but with a different definition of Z. Formula (20) teaches us that the statistical weight of the quantum state Ψ_k with energy $E_k = H_{kk}$ in the mixture will be $\frac{e^{-\beta E_k}}{\sum_k e^{-\beta E_k}}$, which indeed brings us back to the canonical Boltzmann-Gibbs law.

2. Reversible and irreversible evolution. – The preceding considerations led von Neumann to distinguish the two types of evolution at the microphysical scale – namely, reversible and irreversible evolutions – that would result from the measurement.

The reversible evolution of a system or a set of systems is represented by an entirely determinate evolution of the wave function of the system or all of the wave functions of the systems of the set. If one is dealing with a pure case, and if $\Psi(0)$ is the initial form of the wave function then that evolution will follow the equation:

$$\frac{h}{2\pi i}\frac{\partial \Psi}{\partial t} = H\Psi, \tag{22}$$

where H is the Hamiltonian operator of the system, which will be independent of time if the system is isolated. One will then have:

$$\Psi(t) = e^{\frac{2\pi i}{h}iH}\ \Psi(0), \tag{23}$$

with $e^{\frac{2\pi i}{h}iH} = \sum_n \frac{1}{n!}\left(\frac{2\pi i}{n!}tH\right)^n$. The pure initial case will then remain a pure case. The operator $e^{\frac{2\pi i}{h}iH}$ has $e^{-\frac{2\pi i}{h}iH}$ for its adjoint, as one will easily verify, in such a way that its inverse will coincide with its adjoint. It is then a unitary operator that preserves the traces of matrices, so the entropy S of a set of N systems in the state Ψ, which is equal to $-kN\ \mathrm{Tr}(P \log P)$, will thus remain invariant in the course of the evolution – viz., the process will be reversible.

Now, consider, no longer a pure case, but a mixture of pure cases. Each of the functions $\Psi^{(k)}(t)$ of the pure cases will evolve according to the wave equation (22), where H is the Hamiltonian operator of each of the identical systems that are considered. Each of these evolutions is entirely determinate, and therefore represented by:

$$\Psi^{(k)}(t) = e^{\frac{2\pi i}{h}iH}\ \Psi^{(k)}(0)\ ; \tag{24}$$

i.e., by a unitary transformation of $\Psi^{(k)}(0)$. The evolution of the statistical matrix will then be given by:

$$P(t) = \sum p_k P_k(t)\,, \tag{25}$$

or by:

$$(P(t))_{lm} = \sum p_k c_l^{(k)}(t) c_m^{(k)*}(t)\,, \tag{26}$$

if $\Psi^{(k)}(t) = \sum_l c_l^{(k)}(t)\,\varphi_l$.

The trace of P, which is equal to $\sum_k p_k \sum_l \left| c_l^{(k)}(t) \right|^2$, is preserved by the unitary transformation (25) of the $\Psi^{(k)}$, which preserves the norm of $\Psi^{(k)}$; i.e., $\sum_l \left| c_l^{(k)}(t) \right|^2$. Therefore, here again, the entropy S of the mixture will remain invariant, just as it does in the pure case.

According to von Neumann, the irreversible transformations will correspond to processes that are not subject to determinism that will be produced at the moment of the measurement interactions. The interaction of the system or corpuscle 1 under study by a measuring apparatus 2 (or, in our presentation of things, by an indicator corpuscle 2) will correspond to a determinate and reversible evolution of the global system 1 + 2 up to the point at which there is a macroscopic confirmation of the individual state of the system 2 by the observer that would interrupt that evolution by a process that is neither reversible nor even causal, in the present interpretation.

If the initial state of system 1 is a pure case then all of the p_i will be zero, except for just one, which is equal to 1. The entropy S of the system will then be zero, and it will remain so as long as the system is isolated and evolves reversibly. If the interaction that follows the measurement with the system 2 then transforms the state of system 1 into a mixture then all of the p_k become less than 1, and the entropy of system 1 will obviously become positive.

The measurement process is therefore irreversible and is accompanied by an increase of entropy. The chain of reversible evolution is broken, and one can no longer reassemble the state that preceded the measurement from the one followed it by any means.

Von Neumann also showed, by a very long argument, that if the initial state is already a mixture then any measurement that effectively modifies that mixture will have the effect of increasing entropy.

Von Neumann's conclusion is that any measurement will increase entropy, and will in turn have an irreversible character. That irreversibility is obviously linked to the destruction of phase differences by the measurement that was already pointed out and to the impossibility of reassembling the state that existed before the measurement from the state that exists after the measurement, which results from it.

3. How the theory of the double solution must interpret the irreversibility that results from measurement processes. – Since the increase of entropy that is provoked by a measurement process is linked to irreversibility, what is the origin of that irreversibility, when it cannot result from indeterminism in a causal theory? It seems that in the context of the ideas that we have discussed, that irreversibility must be interpreted in the following fashion: After the measurement, either the initial wave u will have been fragmented into spatially-separated portions R_k , and the corpuscle will be finally found in one of these wave trains, or the corpuscle that unhooked form the initial wave u is finally hooked onto one of the component φ_j by a process like "directing." In one case, as in the other one, there will no longer be any interference between the components φ_k and v after the measurement, and the phase difference between these components will no longer intervene. If one assumes that viewpoint then it will clearly appear that von

Neumann's increase in entropy after the measurement is not at all linked to the perception of the result of the measurement by the observer, but to the *objective* fact that the corpuscle will, as a result of the action of the measuring device, be finally found to be hooked to just one component φ_k in the superposition that constitutes the initial pure case.

Our ideas regarding the relationship between the wave *u* and the wave Ψ permit us to make this more precise. In order to represent the probability state after the measurement, physicists will be led to construct the function Ψ_k that corresponds to the "hooking" of the corpuscle to each of the components φ_k of the initial wave. If the result of the measurement is not known then in order to represent the probability state, one must envision a "mixture" of the set of Ψ_k with statistical weights $p_k = | c_k |^2$ that are equal to the squares of the moduli of the coefficient of the φ_k in the original wave. However, once the observation of an observable, macroscopic phenomenon has permitted us to know the result of the measuring process, it can no longer preserve that one Ψ_k, which will again constitute a pure case. The phase relations between the φ_k will then disappear, and one can no longer reassemble the initial Ψ from the final Ψ. That is why in von Neumann's theory, which involves the subjective wave function Ψ exclusively, the increase in entropy by the measurement will correspond to a loss of "information" about the phase differences when we pass from the initial pure case to the mixture that represents the probability state after the measurement when one does not known the result. That concept is in accord with the well-known idea in cybernetics that information corresponds to entropy, with the opposite sign (i.e., negentropy), in such a way that a diminishing of information is the same thing as an increase in entropy (cf., Shannon, Léon Brillouin).

Despite the great interest in the links that can be made between information and entropy (with the sign changed), it seems certain that the entropy of a physical system corresponds to an objective situation, and cannot be defined by starting with the information (which can be more or less exact) of observers. The increase of entropy that is provoked by the measurement must, in our opinion, be attached, not to the fact that the observer is aware of the result of the measurement, but to an *objective* process that provokes the cessation of the interference between the components of the initial wave *v*. In the theory of the double solution, where the evolution of the wave *u* is determined completely, the phases of the components φ_k of *v* keep a well-defined value after the end of the measurement process, but since the corpuscle is no longer hooked to one of the components φ_j , and any observation that provides us with knowledge of microphysical phenomena will imply the localized action of a corpuscle, there is no longer any possibility of knowing the phases of the other components besides φ_j after the measurement.

It will be very interesting to analyze the interpretation of the increase in entropy that is provoked by a measurement in detail using the theory of the double solution. Such an analysis will probably permit the best understanding of the true meaning of von Neumann's thermodynamics.

APPENDIX

STUDY OF THE PASSAGE FROM CLASSICAL MECHANICS
TO WAVE MECHANICS IN A PARTICULAR EXAMPLE.

In this Appendix, we shall study an important precise experimental case that shows quite well how the passage from classical mechanics to wave mechanics is introduced in the present interpretation.

We envision an "electron gun" that is basically defined by an incandescent plate *P* that emits electrons, which is followed by a grille *G* that carries an electrostatic potential that is much higher than that of the plate.

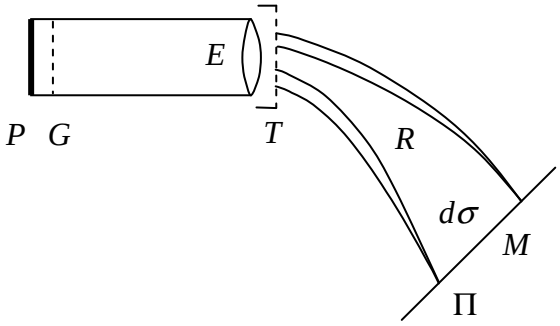

Figure 7.

A flux of electrons leaves the mouth *E* of the electron gun with roughly the same energy *W* and forms a parallel sheaf. In wave mechanics, this sheaf will be associated with a wave train that has a cross section that is equal to that of the mouth *E* and can be roughly assimilated into a piece of a monochromatic plane wave with a frequency $\nu = W / h$ and wave length $\lambda = h / p$.

In the region *R*, the electrons pass through an electrostatic field that is created by human means, and which, in turn, will vary only slightly from the scale of the wave length (which is of order at most 10^{-8} cm). We know that it will then result that wave that is associated with one of the electrons will propagate according to the laws of geometrical optics, which will permit us to define ray-trajectories in a classical fashion. After traversing the static field, the electrons arrive at a photographic plate Π, where their arrivals will be inscribed by successive local actions.

We shall then assume that we have a "mesh" *T* at our disposal, which we can possibly place at the mouth of the electron gun. This mesh is pierced with equal holes that are extremely close and distributed regularly over its surface, where the diameter of the holes is so small that they seem almost point-like to us, but also sufficiently large that they are bigger than the wave length λ.

We shall analyze the production of the phenomena that are observed on the photographic plate in the case where there is no mesh and then in the case where the mesh is placed on the mouth by successively adopting the viewpoint of classical

mechanics, then the interpretation of wave mechanics by the double solution, and finally that of the usual interpretation of wave mechanics. That study will be very instructive.

1. Viewpoint of classical mechanics.

a. The mesh is not in place. – A trajectory will possibly pass through each point of the mouth of the electron gun that will be normal to the plane of that mouth. In the region *R*, the static field will curve the trajectories, which will have the effect of increasing their density in certain regions and rarefying it in other regions. It will then result that the relative of number of trajectories that traverse equal areas $d\sigma$ in the plane Π will vary from one point of that plane to another, so there will be variations in the photographic impression on the plate. In Hamilton-Jacobi theory, the trajectories will be rays of the propagation of a fictitious wave in the geometrical optics approximation, and it will then result that the density of the trajectories that pierce an area $d\sigma$ of the plane Π around a point *M* must vary proportionally to the square of the amplitude $a(M)$ of the Hamilton-Jacobi wave at the point *M*, if one nonetheless assumes the very natural hypothesis that all of the trajectories that leave the mouth of the electron gun are equally probable. Indeed, experiment shows that the variations in the photographic impressions will be proportional to $a^2(M)$ at the different points *M* of the plate.

b. The mesh is in place. – Certain electronic trajectories will be stopped by the solid parts of the mesh, while other ones will pass through the holes in it. Since these holes are very close and uniformly distributed over the surface of the mesh, we will have a very dense set of bundles of electronic trajectories that escape from the holes of the mesh. Each of these bundles can be considered to a bundle of rays of the Hamilton-Jacobi wave. It will further result that the density of the trajectories that pierce the plane Π of the plate will be, in the mean, proportional to $a^2(M)$, and as a result, the same thing will be true for the photographic impression. There is therefore no essential difference between case *a* and case *b*, as far as the interpretation of the distribution of electron impacts on the photographic plate Π is concerned.

2. Viewpoint of the interpretation of wave mechanics by the theory of the double solution.

a. The mesh is not in place. – In the theory of the double solution, the electron is a very localized accident in the structure of the objective wave *u*, which is assumed to be everywhere proportional the wave Ψ before that accident. Each electron that escapes from the electron gun will thus be incorporated into a wave train *v* whose transverse dimensions are macroscopic, since they are equal to the dimensions of the mouth. However, the electron has a position and a well-defined trajectory at the wave front, and since that trajectory is defined by the guidance formula, and the common phase of the waves *v* and Ψ is equal to the Jacobi function, one will see (since $\mathbf{v} = -\frac{1}{m}\operatorname{grad} S$) that the possible trajectories of the electron will again coincide with the rays of the Hamilton-

Jacobi wave. Therefore, although the electron is no longer conceived as an isolated material point here, but as a local accident that is incorporated into a wave, the electron trajectories will be the same as in classical mechanics, and the interpretation of the variations in the photographic impressions on the surface of the plate Π will remain the same as it was in 1.*a*.

b. The mesh is in place. – Here, when an electron leaves the gun, we must say that a small wave train *v* carries the electron through one of the holes in the mesh. Since these holes have dimensions that are very small at our scale, but very large with respect to the wave length, we can consider the wave train that leaves one of the holes as coinciding with a small piece of the Hamilton-Jacobi wave, and since the holes on the surface of the mesh are very numerous and regularly distributed, the guidance formula will further show us that the electron trajectories will coincide, as in 1, with a very dense set of bundles of rays of the Hamilton-Jacobi wave. The interpretation of the distribution of the photographic impressions on the plate Π will thus once more be exactly the same as it was in 1.*b*.

3. Viewpoint of the usual interpretation of wave mechanics.

a. The mesh is in place. – Contrary to what we did in paragraphs 1 and 2, we shall commence with the case in which the mesh is in place. Indeed, it is by studying that case that we shall bring about the agreement between classical mechanics and wave mechanics, according to a method that is frequently discussed in the usual treatments.

If the mesh involves only one hole then the wave Ψ of an electron once it leaves the gun will be reduced to a small wave train whose transverse dimensions will be negligible at our scale. This small wave train will slide along one of the ray-trajectories of the Hamilton-Jacobi theory. The usual interpretation will then tell us that the electron is localized in the wave train, which is distributed statistically over all of its extent. However, since the wave train is roughly point-like at our scale, everything happens *practically* as if the electron were a material point that describes one of the trajectories that were predicted by classical mechanics. That is why we can bring about agreement between classical mechanics and wave mechanics in the case where the propagation of the wave conforms to geometrical optics, and the trajectory is defined approximately at our scale by the hole that is pierced in the mesh that covers the mouth of the gun. However, this agreement will conceal a great difference in principle, since in the usual interpretation of wave mechanics the corpuscle is no longer rigorously localized in the wave train. The trajectory is no longer defined rigorously, but only a very fine bundle of Hamilton-Jacobi rays-trajectories that constitutes a sort of very thin tube in which the corpuscle is present *without being localized* and which, at our scale, appears to be a line with no thickness; it is a sort of "pseudo-trajectory."

Now, pass to the case in which the mesh is pierced with a very large number of small holes that are distributed regularly over its surface. When the electron leaves the gun, its wave Ψ will be composed of a very large number of small separated wave trains that will define a very dense sheaf of pseudo-trajectories. It is easy to account for the fact that the distribution on the photographic plate of the intersections of the pseudo-trajectories with

the plane Π will always be given by the function $a^2(M)$. With the usual interpretation, the electron will not describe any of these pseudo-trajectories; it will be statistically distributed with an equal probability over every set of these pseudo-trajectories. This concept is somewhat strange. What is more, at the moment when the local photographic impression is produced, the electron chooses *one* of these possible pseudo-trajectories, in some way, in order to be localized at the point where it pierces the plane Π. Despite the strangeness of this concept, it nonetheless permits one to understand the origin of the distribution in $a^2(M)$ of photographic impressions on the plate better. Indeed, that distribution comes from the distribution of intersections of the pseudo-trajectories with the plane Π.

b. The mesh is not in place. – This is the most interesting case. When the electron escapes from the gun, it is then associated with a wave train Ψ whose transverse section has macroscopic dimensions, namely, those of the mouth. The usual interpretation of wave mechanics tells us that the electron will be statistically spread over all of the volume of *macroscopic* dimensions that is occupied by the wave train. There is thus no longer any question of defining the trajectories, nor even the pseudo-trajectories. Once again, we are led to say that at the moment when the electron produces a photographic impression at a point of Π, will briefly choose the very small region where it is localized, a conclusion that, as Einstein has often emphasized, is in contradiction with the validity of the usual notions of space and time, even at the macroscopic scale. Moreover, in order for there to be agreement with experiment, we are always obliged to assume that the distribution of photographic localizations on the plate Π is given by the function $a^2(M)$, but here that assertion cannot be justified at all by considering the intersection of the trajectories or pseudo-trajectories with the photographic plate, since there are neither trajectories nor pseudo-trajectories, now; it must then be a purely arbitrary postulate.

Conclusions. – From this detailed analysis of the experimental device that was studied, one can infer the following conclusions: Perhaps it does not prove the falsehood of the present interpretation of wave mechanics, but it at least shows, without a doubt, that the law of distribution of the corpuscular localizations in $a^2 = |\Psi|^2$ has a perfectly-intelligible origin in the interpretation of wave mechanics by the double solution, as it does in classical mechanics, while it takes on the aspect of an entirely arbitrary postulate in the present interpretation, and despite everything, this injects a certain element of doubt into the validity of that present interpretation.

———

BIBLIOGRAPHY.

[1] J. VON NEUMANN, *Les fondaments mathématiques de la théorie des quanta*, transl. PROCA, Alcan, Paris, 1946.

[2] F. LONDON and F. BAUER, *La théorie de l'observation en Mécanique quantique* (Actualités scientifique et industrielles, no. 775, Hermann, Paris, 1939).

[3] LOUIS DE BROGLIE, *Une tentative d'interprétation causale et non linéaire de la Mécanique undulatoire: la théorie de la double solution*, Gauthier-Villars, Paris, 1956.

[4] J.-P. VIGIER, *Structure des micro-objets dans l'interprétation causale de la théorie des quanta*, Gauthier-Villars, Paris, 1956.

[5] A, SOMMERFELD, *Vorlesungen über theoretische Physik*, Bd. VI, *Partielle Differentialgleichungen der Physik*, Dieterich'sche Vorlagsbuchhandlung, Wiesbaden, pp. 183, *et seq.*

[6] Th. VOGEL, *Physique mathématique classique*, Collection Armand Colin, Paris, 1956, pp. 97, *et seq.*

[7] J.-L. DESTOUCHES, *La quantification en théorie fonctionnelle des corpuscles*, Gauthier-Villars, Paris, 1956.