CERN 58 - 7

Theoretical Study Division

1st April 1958

The Density Matrix

by

R. Hagedorn

(Lectures given to the experimental physicists at CERN)

GENEVE

Now in proceeding to equ (16) two abstractions have taken place.

i) The invariant formulation of the statement. This has the effect of introducing a matrix instead of the probabilities. This is only a <u>formal mathematical trick</u> which does not add anything physical to the situation.

ii) We have started with a particular situation, namely that, which follows from the chosen example of an intermediate measurement of which we do not know the result, yielding a particular set of probabilities w_b.
It is obvious however, that the equation

$$\langle C \rangle = \sum_b C_{bb} \cdot w_b$$

does not refer to how we got the knowledge of the W_b and what their values are. This equation (and its invariant formulation) is therefore general and one may forget about the special example by which it was introduced.

p. 17 equ β) and γ) : ρ_c means the representation of ρ in the ϕ_c system, that is just the diagonal representation of equ α).

p. 18 line with δ) read: ρ is real, i.e. hermitian; ... (the same correction two lines below)

p. 20 Remark to theorem 4 : The mentioned a priori knowledge reduces the number of measurements only, if one does not count the experiments, which have delivered this a priori knowledge. In the example β) for instance, we know ρ without measuring the polarization of the beam. This implies, however, that we know the applied magnetic field. For this we have to make three measurements. Considering only the beam, we may regard this as an a priori knowledge.

p. 29. The reader should derive the same result by using the Heisenberg picture.

p. 31. 7th line from the bottom read: ... the result $\cos(\vec{P},\vec{Q})$, thus...

Introductory note

These lectures which are now appearing as a report were conceived as an introductory course for experimental physicists at CERN. The notes of the lectures given have been written up in an informal manner, essentially for the benefit of those who attended the course. As so many requests for copies of these lectures have come from outside CERN, it has been decided to make them more generally available. Neither the treatment nor the list of references should be considered in any way as definitive.

When it was decided that this material should be made available to a wider public, most of the lectures had already been duplicated. Therefore it has been felt desirable to add some explanatory notes to the list of corrections.

Explanatory notes and list of corrections

p. 2. 7th line from the bottom : Read : If the vectors are normalized.

p. 5. equ (3) The reader may prove this by putting

$$\Phi(t) = \sum U_{ba}(t)\varphi_b \quad ; \quad \Phi(0) = \Phi_a$$

p. 5. last line : The first sum is over b and c, the second one over c.

p.12. 14th line from below : After nothing insert a semicolon.

p.13. 10th line from above : Read:... its wave number k , ...

p.15. equ (16) : This equation is the invariant formulation of the above equ.

$$\langle C \rangle = \sum_b c_{bb}\, w_b$$

Our partial knowledge of the system is expressed by the probabilities

$$w_b = |U_{ba}(t')|^2 .$$

The Density Matrix

by R. Hagedorn

Contents :

p. 32. 4th line from the bottom read: ... by a little discussion).

p. 38. 4th line from above read: ... formulation (42), which ...

p. 41. above equ. (49) read: The answer is yes. Put ...

1. Representations and Transformation Theory

In this introduction I remind you of a few basic facts of ordinary quantum mechanics.

a) A physical system is said to be in a definite micro-state, if every observable of a complete set of commuting observables has been measured. We represent this state by a state vector in Hilbert space and call it a pure state.
Example : If we know that a hydrogen atom is in energy state belonging to E_n, we do not yet know the whole state unless we determine also the angular momentum j and m . If this is not done, the state may still be any linear combination of angular momentum eigenfunctions belonging to E_n and we do not know which one.

The state is defined by our knowledge and by nothing else.*)

b) To every observable belongs a hermitian operator and a single measurement of this observable always results in giving one of the eigenvalues. After the measurement our knowledge jumps discontinuously because we know something we did not know before. Consequently the state vector also jumps : After the measurement it is an eigenvector of the corresponding operator, but with respect to other operators which commute with that in question, it may still be undetermined and therefore not "pure".

c) Apart from the discontinuous jumping of state vectors because of measurements, there is another movement of the states, which is continuous and follows the equation of motion

$$H\Psi = -\frac{\hbar}{i}\frac{\partial\Psi}{\partial t} \qquad \text{(Schrödinger picture)}$$

The state follows this equation only between the measurements, that is during the time where the system is closed. H is the Hamiltonian of the closed system.

---

*) The meaning is, of course, that one does one's best to get the right knowledge. If, for instance, one cannot get complete knowledge, one should realize that. Otherwise one gets wrong predictions.

d) Whereas the movement according to the equation of motion is completely determined by the initial state (the equation of motion is 1st order in t!), the discontinuous "jumps" connected with measurements are not completely determined. In fact, if the state Ψ is known and we make a measurement to which the operator A belongs, then the rule is :

Take the complete orthonormal set of eigenvectors of A defined by

$A\phi_a = a\phi_a$ and expand the normalized $[(\Psi, \Psi) = 1]$ state

into a series :

$\Psi = \sum c_a \phi_a$; $c_a = (\phi_a, \Psi)$ is the "cosinus" between ϕ_a, Ψ

The scalar product is defined abstractly by

$$(\Psi, \phi) = (\phi, \Psi)^*$$

$(\Psi, \lambda\phi) = \lambda(\Psi, \phi)$ where λ is any complex number (not operator

This definition does not refer to any coordinate system; it is invariant. If the space is a three-dimensional Euclidian (real) space, then this is the normal scalar product, which is - without reference to any coordinates - defined by

$$(\vec{r}_1, \vec{r}_2) = |r_1| \cdot |r_2| \cdot \cos(\vec{r}_1, \vec{r}_2)$$

If the vector is normalized to $(\Psi, \Psi) = 1$, then the scalar product is just the cosinus (in a generalized meaning).

Then $|c_a|^2$ is the probability that the measurement A yields the eigenvalue a. Once the measurement is done, we know that the state is ϕ_a. The state vector has jumped into this new direction and begins immediately to follow again the Schrödinger equation.

e) The eigenvectors of a complete set of commuting hermitian operators form a basis of the Hilbert space and a state is completely defined by its components with respect to these basis vectors, that is by the c_a . Thus

$$\Psi \longrightarrow \begin{pmatrix} c_1 \\ c_2 \\ \vdots \end{pmatrix}$$

which means that the abstract vector Ψ is represented in a certain coordinate system by the components c_a .

The ordinary Schrödinger functions $\Psi(x)$ are such a representation where the subscript of the c is continuous and is called x.

Thus

$$|\Psi(x)|^2 dx$$

is the probability to find in a position (x) - measurement the value between x and x + dx.

f) The mean value of a large number of measurements, all starting from the same state Ψ, is given by

$$\langle A \rangle = \sum_a prob.(a) \cdot a = \sum_a a |c_a|^2 = \sum_a c_a^* a c_a .$$

Because the c_a represent the vector Ψ , this is nothing else than the scalar product (written in components)

$$\langle A \rangle = \sum_a c_a^* a c_a = (\Psi, A\Psi)$$

Indeed : $\Psi = \sum c_a \phi_a$ gives

$$\sum_{a',a} (c_{a'} \phi_{a'}, a c_a \phi_a) = \sum c_{a'}^* a c_a (\phi_{a'} \phi_a) = \sum c_a^* a c_a .$$

The term "expectation value" is common but unjustified, because it is really a mean value which need not coincide with what we <u>expect</u>, namely one of the eigenvalues "a".

g) Commuting operators possess a common basis in Hilbert space (in fact, the basis is completely defined by the complete set of commuting operators) and can be measured simultaneously.

Non-commuting operators cannot be measured simultaneously. Let A, B, C be three non-commuting operators (assume for simplicity that each one forms a complete set) of observables :

$$(1)\quad \begin{cases} A\phi_a = a\phi_a \\ B\varphi_b = b\varphi_b \\ C\psi_c = c\psi_c \end{cases}$$

(We label the eigenvectors by the corresponding eigenvalues rather than writing $A\phi_i = a_i\phi_i$. Our notation is therefore very near to Dirac's.)

Assume A to be measured. Then the system is in state ϕ_a. Which is the probability to find the eigenvalue b in the measurement B ? If this measurement is done immediately after the measurement A, then the state is still

$$\phi(t)_{t=0} = \phi(0) = \phi_a$$

The rule is : Expand ϕ in a series of φ_b's :

$$(2)\quad \phi(0) = \phi_a = \sum \overset{o}{U}_{ba}\varphi_b \quad ; \quad \overset{o}{U}_{ba} = (\varphi_b, \phi_a)$$ is unitary and transforms from φ to ϕ.

The superscript o on $\overset{o}{U}$ means that it is a constant unitary transformation connecting two systems of coordinates : φ and ϕ. According to the above rule d), the wanted probability is $|\overset{o}{U}_{ba}|^2$

If we wait sometime before we measure B, the state ϕ has changed according to Schrödinger's equation :

$$\dot{\phi} = -\frac{i}{\hbar}H\phi \quad ; \quad \phi(0) = \phi_a .$$

This can be taken into account by using a time-dependent $U_{ba}(t)$ which then also obeys Schrödinger's equation :

$$\begin{cases} \dfrac{dU_{ba}}{dt} = -\dfrac{i}{\hbar}(H\cdot U)_{ba}(t) \\ U_{ba}(0) = \overset{\circ}{U}_{ba} \end{cases} \tag{3}$$

$U_{ba}(t)$ are then the components (with respect to the basis φ_b) of that state vector ϕ, which for $t = 0$ coincides with ϕ_a We may call $U_{ba}(t)$ a Schrödinger function (for fixed a).

(Example : $a = E_n$, energy eigenvalue, $b = x$, the position of the particle. Then

$$U_{xn}(t) = \psi_n(x,t)$$

is the ordinary Schrödinger function).

That is : Those unitary matrices $U(t)$ which obey the Schrödinger equation and which for $t = 0$ (the time of the last measurement) coincide with the $\overset{\circ}{U}$, give us the full information on possible results of measurements. The columns of these matrices are - in a general sense - nothing else but Schrödinger functions.

In the case that we make the next measurement at time t (instead of at $t = 0$), the probability is $|U(t)_{ba}|^2$ (for finding the eigenvalue b).

Now another question : What is the probability of measuring c if C is applied? Obviously

prob. (c_t after a) = $|U_{ca}(t)|^2$ This can be written in two ways :

$$\phi(t) = \sum_b U_{ba}(t)\,\overset{\circ}{U}_{cb}\,\psi_c = \sum U_{ca}(t)\,\psi_c \ ;$$

Therefore

$$U_{ca}(t) = \sum_{b} \overset{\circ}{U}_{cb}\, U_{ba}(t)$$

Thus, if B is <u>not</u> measured inbetween A and C, we have

(4) $$\text{prob.}_t\ (\text{c after a}) = \left| U_{ca}(t) \right|^2 = \left| \sum \overset{\circ}{U}_{cb}\, U_{ba}(t) \right|^2$$

If instead B is measured at time t' between A and C, we have

$$\text{prob. (c after b)} = \left| U_{cb}(t-t') \right|^2 \text{ and } \quad \text{prob. (b after a)} = \left| U_{ba}(t') \right|^2$$

Then the probability to measure c after b and b after a is

(5) $$\text{prob. } (c_t \text{ after } b_{t'} \text{ after a}) = \text{prob. } (c_t \text{ after } b_{t'}) \cdot \text{prob.} (b_{t'} \text{ after a}$$

$$= \left| U_{cb}(t-t') \right|^2 \cdot \left| U_{ba}(t') \right|^2$$

If B is measured but we do not recognize the result, then

(6) $$\text{prob.(c after a, B measured)} = \sum_{b} \text{prob.} (c_t \text{ after } b_{t'}) \cdot \text{prob.} (b_{t'} \text{ after}$$

$$= \sum_{b} \left| U_{cb}(t-t') \right|^2 \cdot \left| U_{ba}(t') \right|^2$$

Because $\left| \sum_{b} \overset{\circ}{U}_{cb}\, U_{ba}(t) \right|^2 \neq \sum_{b} \left| U_{cb}(t-t') \right|^2 \cdot \left| U_{ba}(t') \right|^2$

one has to distinguish carefully these two cases. This is not the place for a full discussion, but the reader may try to discuss these questions by means of an idealized experiment and to find out what it really means. He will immediately see that it automatically leads from here into the deepest questions of interpretation of quantum mechanics and the theory of measurements.

h) We may equally well represent the operators by a matrix with respect to a certain basis, say ϕ_a , as we have represented the vectors (states) by their components with respect to this basis :

(7) $$O_{aa'} = (\phi_a , O\phi_{a'})$$

One must clearly distinguish between state vectors and operators on the one hand and their representations, namely components and matrix elements, on the other hand. The latter depend on the chosen basis, the former are defined in an abstract manner and do not depend on the basis.

Especially

(8) $$A_{a'a} = (\phi_{a'} , A\phi_a) = (\phi_{a'} , a\phi_a) = a\delta_{a'a}$$ is diagonal.

Changing the basis by means of a unitary transformation, changes the matrices also

$$O_{aa'} = (\phi_a , O\phi_{a'}) \quad ; \quad \phi_a = \sum \mathring{U}_{ba}\varphi_b$$ or short : $\phi = \mathring{U}\varphi$

gives

$$O_{aa'} = \left(\sum \mathring{U}_{ba}\varphi_b , O\sum \mathring{U}_{b'a'}\varphi_{b'}\right) = \sum_{bb'} \mathring{U}_{b'a'}\mathring{U}^*_{ba} O_{bb'}$$

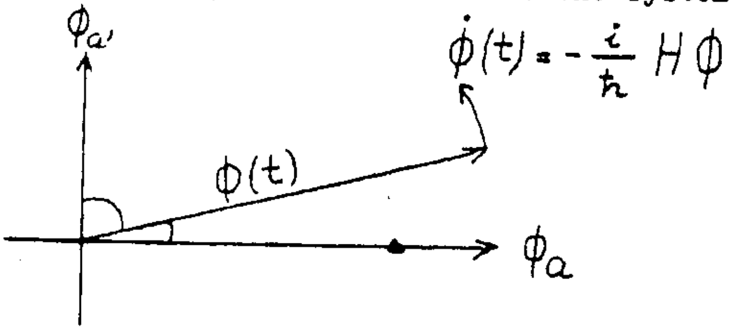
$$= \sum_{bb'} (\mathring{U}^+)_{ab} O_{bb'} \mathring{U}_{b'a'}$$

(9) In short : $O_A = \mathring{U}^+ O_B \mathring{U} = \mathring{U}^{-1} O_B \mathring{U}$

(Here O_A means the matrix representation in the basis ϕ_a of A.)

i) Heisenberg- and Schrödinger Picture

So far we have used the Schrödinger picture, the operators are constant in time and generate by their eigenvectors a constant base in Hilbert space. The state of the system turns around :



and at a given time the cosinus between ψ and the axes ϕ_a give the probability to measure the eigenvalue a of A. The jump of ψ into one of the axes is the more probable the nearer ψ is to this axis.

These statements are the only ones which have experimental meaning. They provide the necessary link between the determined "motion" of the isolated system and the undetermined results of measurements.

Because this is so, we can say that the angles between ϕ and the axes of that operator A which I am going to measure next are the only relevant content of the theory. By definition they do not depend on the description :

$$(\psi, \phi) = (\psi, U^{+} U \phi) = (U\psi, U\phi) = (\psi', \phi') \quad \text{for any unitary } U.$$

Hence we may use even a time-dependent U without changing the physical content of the theory, because if we turn all vectors (including the basis) around, the <u>relative</u> motion of ϕ with respect to ϕ_a (or any other basis) remains unchanged and so does physics.

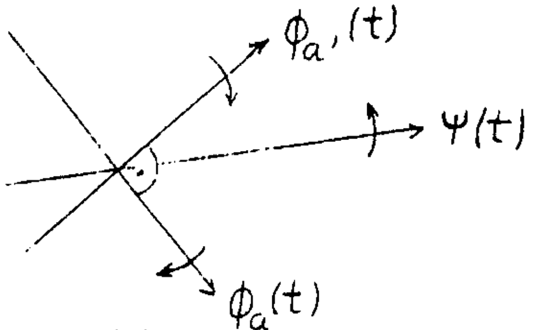
Taking now any unitary V(t) (we call it V instead of U, because it has another significance; see below), we have for all vectors $\Psi, \phi, \varphi \dots$

$$(10) \quad \begin{cases} \Psi' = V(t)\Psi & \text{and for all operators} \\ O'(t) = V(t)\, O\, V^{+}(t), & \text{this holds because of invariance of} \end{cases}$$

the scalar product (expectation value) :

$$(\Psi, O\phi) = (\Psi', O'\phi') = (V\Psi, O'V\phi) = (\Psi, V^{+}O'V\phi)$$

Of course, now the basis vectors are moving : They are the eigenvectors of time-dependent operators. The picture is like this :



Because V(t) may be any time-dependent U-matrix, there are infinitely many ways of turning the Hilbert space around. But three of them have been proved to be especially useful :

a) Schrödinger picture
b) Interaction picture
c) Heisenberg picture.

We get the Heisenberg picture by turning the Hilbert space in such a way that it moves oppositely to ϕ with the same speed - thus ϕ' remains simply constant. To find the V(t) :

$\phi(t) = V(t)\,\phi_0$ where ϕ_0 is constant, yields

$$(11) \quad -\frac{\hbar}{i}\dot{V}\phi_0 = HV\phi_0 \text{ or } -\frac{\hbar}{i}\dot{V} = HV \text{ or } V = e^{-\frac{i}{\hbar}Ht} ;\ V(0) = 1$$

Compare equ. (11) with equ. (3); V and U follow both the same equation, only the initial condition is different. The V in equ. (11), however, is only a special case of that in equ. (3) because θ_{ba} may be the unit matrix. This is the case if after the measurement A one again applies the same measurement at a later time t and asks for the probabilities.

We have now according to (10)

$$\phi_o = V^+\phi(t)$$

$$O(t) = V^+ O V = e^{\frac{i}{\hbar}Ht} O\, e^{-\frac{i}{\hbar}Ht}$$

Now the state $\phi = \phi_o$ remains constant, but the operators and therefore the basis turn around :

(12) $$\frac{dO(t)}{dt} = \frac{i}{\hbar}\left[H,O\right]$$ is now the equation of motion. (Proof?)

<u>The interaction picture</u> is one of the infinitively many inbetween Schrödinger and Heisenberg : Let H equal $H_o + H_1$ where the solution of H_o is known. Then we put

$$V = e^{-\frac{i}{\hbar}H_o t}$$

and find the following interpretation : The axes (operators) move as in the corresponding simple case where only H_o is present. On the other hand, ϕ is not constant; because by this transformation we have only removed that part of its motion which corresponds to H_o. The rest of motion still remains. Indeed : Calling ϕ_i the new vector and ϕ the old (Schrödinger picture), one has

(13) $$\phi(t) = e^{-\frac{i}{\hbar}H_o t}\phi_i(t)$$

From $\frac{d\phi}{dt} = -\frac{i}{\hbar}\left(H_0 + H_i\right)\phi$ follows immediately

$$\left.\begin{aligned} &(14) \quad \frac{d\phi_i(t)}{dt} = -\frac{i}{\hbar} H_i(t)\phi_i(t) \;; \quad H_i(t) = e^{\frac{i}{\hbar}H_0 t} H_i \, e^{-\frac{i}{\hbar}H_0 t} \\ &(15) \quad \frac{dO(t)}{dt} = \frac{i}{\hbar}\left[H_0, O(t)\right] \quad \text{where} \quad O(t) = e^{\frac{i}{\hbar}H_0 t} O \, e^{-\frac{i}{\hbar}H_0 t} \end{aligned}\right\} \begin{array}{l}\text{Inter-}\\ \text{action}\\ \text{picture}\end{array}$$

(The simple proof is left to the reader.)

Here the equation of motion is split into two parts, one for the state and the other for the operators; the relative motion of ψ with respect to the axes remains the same of course.

2) Treatment of States which are not pure

a) General remarks

As said in the introduction, the state represents our knowledge. If this knowledge is an optimum - that is if a complete set of communting observables has been measured - it is a pure state. Practically however this is the exception and the rule is to have only partial knowledge; then the state is not pure.

Three examples :

α) A system in a temperature bath. If we do not count the bath to the system, then we know only macroscopic data : p, V, T or any other set of thermodynamical quantities. We know practically nothing about the state vector ψ and as long as we have the contact of the system with the bath, ψ is not well defined. But even if we remove the system from the bath and isolate it completely, we may say, perhaps, that it is in a certain state ψ. But this statement means absolutely nothing, because if ψ is defined by our knowledge, it is just nonsense to say so. In fact, a macroscopic system of $N = 10^{23}$ particles has so many commuting observables that it is impossible to measure all of them. Even if we measure its energy, then we know only that it is in an energy eigenstate, but we do not know which linear combination it is with respect to all the other constants of motion. But nevertheless we have some knowledge : p, T, V, for example. And this allows us to know at least certain probabilities for finding the system in a pure state; for instance :

$$w(\varepsilon) = \frac{e^{-\frac{\varepsilon}{kT}}}{\sum_{\varepsilon} e^{-\frac{\varepsilon}{kT}}}$$

is the probability of finding the system in an energy eigenstate of energy ε.

This "state" of a system just coming out of a temperature bath is very "impure" so to speak. This may be different in other cases, where a part of the commuting observables have been measured. We may call such a state partially impure and may add a word indicating with respect to which observables it is impure.

β) The following example is of that kind. Consider a monochromatic beam of atoms without interaction. We need not consider then the whole beam as "the system" but rather one single representative atom. We know then its wave number k, its energy (say it is in the ground state), but not its magnetic quantum number if it happens to have a spin $\neq 0$. The state is then impure with respect to spin orientation.

γ) Consider the example of three operators A, B, C (each being a "complete system")and let us start with a pure state ϕ_a. If we wish to know the probability for finding c in the next experiment, we have the unique answer

$$\text{prob.}(c_t \text{ after } a) = \left| U_{ca}(t) \right|^2 \ ;$$

where $\frac{dU}{dt} = -\frac{i}{\hbar} H U$ and $U(0) = \overset{0}{U}_{ca} = (\psi_c, \phi_a)$.

But suppose that before we come to do the experiment C, somebody has done experiment B on our system, but we do not know the result. We only know that before it was in a pure state ϕ_a and that B was measured. Nevertheless we find according to (6) :

$$\text{prob.}(c_t \text{ after } a, \text{ B measured}_{t'}) = \sum_b \left| U_{cb}(t-t') \right|^2 \left| U_{ba}(t') \right|^2$$

In words : Immediately after B has been measured, the system is in a pure state ψ_b for the others. For us, it is no longer in the pure state ϕ_a because B was measured.

But it is not in a pure state ψ_b for us either, since we do not know the result. Nevertheless, after the measurement B (of which we do not know the result) the system is at least in a well defined "mixture state". We know, in fact, with which probability it is in each one of the possible pure states ψ_b : this is $|U_{ba}(t')|^2$. Thus we do not know everything, namely the state of the system (which is one definite of the ψ_b), but on the other hand we are not completely ignorant; we know just the $|U_{ba}(t')|^2$.

b) Introduction of the Density Matrix

Now, what do these examples have in common? We have a well defined partial knowledge of the state. It may be partial with respect to all observables (example a, c) or only with respect to some of them (example b). The question is, how to express this situation in the frame of quantum mechanics. Example c) will very naturally lead to the general description. Obviously, the expectation value of C is now

$$\langle C\rangle = \sum_{c,b} c\cdot|U_{cb}|^2\cdot|U_{ba}|^2 = \sum_{c,b} c\,|U_{cb}(t-t')|^2\cdot w_b(t') \qquad w_b(t') = |U_{ba}(t')|^2 \ .$$

In the following we shall omit the time arguments. We fix the time t' at which B is measured, because any measurement means a "jump" of the state and only between two measurements, when the system is isolated, the state follows the Schrödinger equation. Thus, in order to have definite situations, the time t' has to be fixed, and the ω_b have to be considered as constants. We may put t' = 0, then t - t' = t is the time at which C is measured. U_{cb} always means $U_{cb}(t)$. This is, as we have seen, the Schrödinger function, which (for b fixed) describes the state $\psi(t)$ with respect to the basis of C. Later on we shall come back to the question of time dependence again.

Now

$$\sum_c c\,|U_{cb}|^2 = \sum_c U^*_{cb}\,c\,U_{cb} = \sum_c (U^+)_{bc}\,c\,U_{cb} = C_{bb}$$

is the diagonal element of the matrix representation with respect to the Schrödinger functions U_{cb} . Hence

$$\langle C \rangle = \sum_b C_{bb} \cdot w_b$$

and this may be written also

$$\langle C \rangle = \sum_{bb'} C_{bb'}\,\delta_{b'b} \cdot w_b = \sum_{bb'} C_{bb'}\,\rho_{b'b} = \mathrm{Tr}\,(C\rho)$$

Here $\rho_{b'b} = \delta_{b'b} \cdot w_b$ is diagonal. But this is just due to our special description. The equation

(16) $$\langle c \rangle = \mathrm{Tr}\,(C\rho) = \mathrm{Tr}\,(\rho C)$$ is invariant *) and therefore true for any base.

---

*) $\mathrm{Tr}\,(AB) = \sum_{ik} A_{ik} B_{ki} = \sum_{ik} B_{ki} A_{ik} = \mathrm{Tr}\,(BA)$ and

$\mathrm{Tr}\,(U^{-1} A U) = \sum A_{kl}\, U_{li}\, U^{-1}_{ik} = \sum A_{kl}\,\delta_{lk} = \sum A_{kk} = \mathrm{Tr}\, A$.

c) Properties of the Density Matrix

Let us now play around with this matrix ρ, to find a few theorems. First of all, if we know the state (before measuring C) to be a pure state, say ψ_b, then $\omega_b = 1$, that is

$$\rho = \begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix} \qquad \text{and} \qquad \langle C \rangle = \mathrm{Tr}(\rho C) = C_{bb} = (\psi_b, C\psi_b)$$

That is, the definition of (16) contains the ordinary treatment, where one considers pure states only.

Theorem 1 : $\langle C \rangle = \mathrm{Tr}(\rho C)$ is an invariant (with respect to unitary transformations in Hilbert space) generalization of the ordinary definition of expectation values.

We have found - by chance - that representation in which ρ is diagonal, because we have described ρ with respect to that system in which we knew the probabilities ω_b. In any other system is not diagonal. Let us take again our three operators A, B, C and assume that A has been measured but that we do not know the result. We may, however, know the probability ω_a for the system being in state ϕ_a.

(Whatever our knowledge may be, it can be represented by ω_a. If we know nothing, then all ω_a are equal; if we have complete knowledge, one single ω_a is 1, the others are zero).

We ask now for the expectation of C and will find three different representations of ρ accordingly to whether we express ρ in the basis ϕ_a or ψ_b or ψ_c. We already know that ρ is diagonal in ϕ_a with the ω_a as eigenvalues. For ψ_c as basis we may expect also something simple, because this is the system of eigenvectors of C, the quantity we want to measure. The most general form we will have in the system φ_b, because B may be any operator which does not commute with A and/or C. We have

(17) $$\langle C \rangle = \sum_a \varpi_a C_{aa} = \sum_{abb'} \varpi_a U^*_{ba} U_{b'a} C_{bb'} = \sum_{ac} \varpi_a U^*_{ca} U_{ca} \cdot c$$

Let us analyse these equations in terms of

$$\langle C \rangle = \mathrm{Tr}(\rho C)$$

α) $\mathrm{Tr}(\rho C) = \sum_a \varpi_a C_{aa}$; $\rho_{aa'} = \varpi_a \delta_{aa'}$ is diagonal

β) $\mathrm{Tr}(\rho C) = \sum_{abb'} C_{bb'} U_{b'a} \varpi_a U^*_{ba}$; $\rho_{b'b} = \sum_a U_{b'a} \varpi_a U^*_{ba} = (U \rho_a U^{+})_{b'b}$

γ) $\mathrm{Tr}(\rho C) = \sum_{ac} c\, U^*_{ca} \varpi_a U_{ac} = \sum c\, \delta_{cc'} U^*_{c'a} \varpi_a U_{ac}$

$$\rho_{c'c} = \sum_a U^*_{c'a} \varpi_a U_{ac} = (U \rho_a U^{+})_{c'c}$$

here C is diagonal but ρ is not.

We see :

<u>Theorem 2</u> : The density matrix depends in its representation on <u>two</u> systems of basis vectors in Hilbert space, firstly on that system in which we know the probabilities ϖ_a, secondly on that system which we choose as basis in our Hilbert space. If we choose that special system as basis in which we know the probability distribution, then ρ is diagonal and its eigenvalues are just the probabilities ϖ_a .

Still theorem [illegible]:

Its general form is :

$$\rho_{b'b} = \sum_a U_{b'a} U^{*}_{ba} w_a = (U \rho_d U^{+})_{b'b}$$

where ρ_d means the diagonal form.

A few properties of ρ are the following :

α) By definition $\sum_a w_a = 1$, because w_a are probabilities. Hence

$\sum w_a = \mathrm{Tr}\,\rho = 1$ in any representation. Furthermore $w_a \geqslant 0$, hence $0 \leq w_a \leq 1$

and thus

β) $\sum w_a^2 = \mathrm{Tr}\,\rho^2 \leqslant 1$ in any representation. If ρ is not diagonal this reads

γ) $\mathrm{Tr}\,\rho\rho = \sum_{ik} \rho_{ik}\rho_{ki} = \sum_{ik} |\rho_{ik}|^2 \leqslant 1$ which limits the valu of the elements.

δ) In its diagonal form ρ is real i.e. hermetic; this statement is invariant. Hence

ρ is hermetic : $\rho_{ik} = \rho^{*}_{ki}$, which was already used in γ).

ε) Take that representation in which the operator to measure, C, is diagonal and assume C to have only positive eigenvalues. Then $\langle C \rangle \geqslant 0$ implies that every diagonal element $\rho_{kk} \geqslant 0$. This can be concluded in any representation. Hence

$\rho_{kk} \geqslant 0$ in any representation.

We collect these statements in

Theorem 3 :

$$\left.\begin{aligned} &\mathrm{Tr}\,\rho = 1 \\ &\rho^{+} = \rho \\ &\mathrm{Tr}\,\rho^{2} \leq 1 \\ &\sum_{ik} |\rho_{ik}|^{2} \leq 1 \\ &0 \leq \rho_{kk} \leq 1 \end{aligned}\right\} \text{in any representation}$$

So far, we have written everything with two matrix subscripts, because we supposed that our operators A, B, C form each one a complete set of commuting observables. This is an idealization. In general a state is labelled by many quantum numbers and the transformation matrices are therefore labelled by more than two subscripts, they are multi-dimensional matrices. Accordingly ρ is a multi-dimensional matrix and the foregoing formulae are valid if $i = (i_1 \ldots i_e)$; $k = (k_1 \ldots k_e)$, that is, the subscripts represent each a full set of quantum numbers. It may happen, however, that we know something definite of the state - it may be only partially impure. With respect to the corresponding subscripts, ρ degenerates to a matrix (in diagonal representation) which contains only zeros but for one place where it has a 1. We may then drop this subscript from ρ and the dimension of ρ is reduced by one. Remember, for instance, example 2a, β) where a monochromatic beam of atoms is given. The only quantity which we do not know is the spin component in z -direction, i.e. the polarization. Let j be the spin of the atom, then there are 2j + 1 possible orientations and ρ reduces to a square (two dimensional) matrix of 2j + 1 rows and columns.

Another reduction of ρ takes place if we are not interested in some of the quantum numbers specifying the state and if they are not relevant to the planned experiment. For instance, the considered beam may not be monochromatic, but we are interested only in its polarization. Then the subscript referring to the wave number k drops out and ρ is also in this case a $(2j + 1) \times (2j + 1)$ matrix. This ρ however cannot serve then to calculate expectation values for the momentum.

In general : if ρ is a N by N matrix, then from theorem 3 we conclude :

The N^2 complex matrix elements contain $2N^2$ real numbers. The condition $\rho^+ = \rho$ reduces this to N^2 and the condition $\mathrm{Tr}\,\rho = 1$ to $N^2 - 1$. The inequalities restrict the values of these $N^2 - 1$ parameters but not their total number. Thus we have

Theorem 4 : **If for a given problem the density matrix reduces to a N x N matrix, then ρ contains $N^2 - 1$ independent parameters. Determination of ρ implies therefore in general $N^2 - 1$ measurements. In some cases this number may be reduced by a priori knowledge.**

Example : α) We may pass the beam through an analyser which selects m = 0 only. For the selected beam everything is known. However we have to pay for it in losing part of the previous beam.

β) The atoms may come from a box in a temperature bath and a magnetic field $\vec{\mathcal{H}}$ is applied to the box. Then ρ is determined by measuring T and $\vec{\mathcal{H}}$, since then we know the polarization of the beam because the polarization is to follow Boltzmanns law :

$$\vec{P} = \frac{\mu \vec{\mathcal{H}}}{\frac{1}{2} kT} .$$

* * *

Finally we may ask for the time variation of ρ .

We go back to theorem 2, which states that the most general form of ρ is given by

$$\rho_{b'b} = \sum_{a} U_{b'a}(t)\, U^{*}_{ba}(t)\, w_{a} = \left(U(t)\, \rho_{d}\, U^{+}(t) \right)_{b'b} \tag{18}$$

If we ask now for the time dependence, we must again recall the fact that a physical system is changing in time in two ways :

a) by measurements. These are sudden jumps and one does not know beforehand which will be the jump. This motion to some extent is undetermined.

b) Between measurements, as an isolated system, it moves according to the Schrödinger equation. This is completely determined.

If we now ask for the time variation of the density matrix, this question makes sense only with respect to the second movement. That is, we have to suppose the w_a to be representatives of our knowledge as coming from the last experiment. In so far they have to be considered as constant. However, since that last measurement has been done, the system has changed according to the Schrödinger equation and this is expressed by the time variation of the Schrödinger functions $U_{ba}(t)$. (Remember: $U_{ba}(t)$ is the Schrödinger function representing that state $\phi(t)$ which at $t = 0$ was ϕ_a , if the basis is φ_b .) As we are here in the Schrödinger picture, we see that $\rho_{b'b}(t)$ is timedependent.

Which is its equation of motion? From (18) follows

$$\frac{d\rho_{b'b}}{dt} = \left(\frac{dU}{dt}\, \rho_d\, U^{+} \right)_{b'b} + \left(U \rho_d \frac{dU^{+}}{dt} \right)_{b'b}$$; and from

$$\frac{dU}{dt} = -\frac{i}{\hbar} H U \quad ; \quad \frac{dU^{+}}{dt} = \frac{i}{\hbar} U^{+} H$$ follows then

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}\left[H,\rho\right] \qquad (19)$$ (Schrödinger picture)

Note that this is not the same as for operators in the Heisenberg picture, where one has a + sign /see (15)/. Indeed, this time dependence comes from the fact that the matrix ρ is defined in the Schrödinger picture by Schrödinger functions and, as we have seen, the Schrödinger states move opposite to that direction in which Heisenberg operators move in the Heisenberg picture.

We can integrate (19) at once :

$$\rho(t) = e^{-\frac{i}{\hbar}Ht}\,\rho(0)\,e^{\frac{i}{\hbar}Ht} \qquad (20)$$ (Schrödinger picture)

Now it is easy to go over to the Heisenberg picture :
The expectation of C must be the same in both descriptions, hence

$$\langle C\rangle = \mathrm{Tr}\left(\rho(t)C\right) = \mathrm{Tr}\left(e^{-\frac{i}{\hbar}Ht}\rho(0)\,e^{\frac{i}{\hbar}Ht}\,C\right)$$

$$= \mathrm{Tr}\left(\rho(0)\,e^{\frac{i}{\hbar}Ht}\,C\,e^{-\frac{i}{\hbar}Ht}\right) \quad \text{or}$$

$$\langle C\rangle = \mathrm{Tr}\left(\rho(0)\,C(t)\right)$$ (Heisenberg picture)

Here we have used the fact that Tr(AB) = Tr(BA).

We may therefore state

Theorem 5 : In the Schrödinger picture, one has constant operators ; $\frac{d\rho}{dt} = -\frac{i}{\hbar}\left[H,\rho\right]$

In the Heisenberg picture, one has instead constant ρ ; $\frac{dO}{dt} = \frac{i}{\hbar}\left[H,O\right]$.

Stell Theorem 5

Remembering that for pure states the ϱ-description reduces to the common formulation of quantum mechanics, we may say :

The whole quantum theory may be described by the above equations of motion and the requirement that

$$\langle O \rangle = \mathrm{Tr}(\varrho O) \qquad \text{for any observable } O.$$

Finally, with regard to theorem 4 we may change our interpretation and forget (so to speak) the derivation of the density matrix and its formal definition by theorem 2. Instead, we may define it by expectation values, i.e. by measurements only and say :

Theorem 6 : The density matrix ϱ expresses our knowledge of the physical system in question. It is defined by

$\langle O \rangle = \mathrm{Tr}(\varrho O)$ If it contains $r = N^2 - 1$ independent parameters, then the measurement of r expectation values $\langle O^{(\ell)} \rangle$ is sufficient for its complete determination. Sometimes less is required (see examples following theorem 4). According to the chosen picture it follows the equations of motion given in theorem 5. It is always bound to fulfil the requirements of theorem 3.

3) The Use of the Density Matrix

In the following examples the use of the density matrix is demonstrated. We select only a few very simple cases, and the reader should not think that it is always like that. The general situation does not always follow the ideal lines of theorem 6; rather, the situation is just reversed : we do not determine a density matrix ρ by measuring the N^2-1 independent data and thus have finished our job. The normal situation is that we have some information (a priori or experimental) concerning our system and we want to calculate that density matrix ρ which refers to a certain planned experiment. This calculation is the theoretician's task and can (and will, in general) be a very complicated one. What I mainly can do here is to show how the result of a theoretical investigation, which is expressed in terms of a density matrix, has to be interpreted by the experimentalist. But in some of the following examples we can even calculate the density matrix by using simple arguments.

a) Thermodynamics

Consider a box containing a gas. Assuming the particles to be mass points of mass m without interaction, we may solve the Schrödinger equation for each particle separately and put the solutions together in an appropriate way to have the solution of the whole system. Thereby the energy eigenstates ϕ_E are defined.

We put this box now into a bath of temperature T and assume a terribly small but finite interaction between the bath and the box and also among the particles. This interaction will then cause a thermodynamical equilibrium without visibly changing the spectrum (the set of eigenvalues) of the Hamiltonian H. Also the state vectors are hardly changed, but the system is no longer in a definite state because of the interaction.

In a loose sense one can already say in classical physics that temperature and energy are complementary; the exact measurement of one of these two destroys the knowledge of the other one. This is the same in quantum mechanics.

But we can at least say something, namely : if we remove the system from the bath and then measure its energy, we must find it in a definite state, say ϕ_E, and the probability of finding it just in this state is according to Boltzmann :

(21) $$W_E = \frac{e^{-\frac{E}{kT}}}{\sum_E e^{-\frac{E}{kT}}}$$

This is a situation which calls for a description by the density matrix, since we do not know the state but the whole probability distribution. Using the energy eigenfunctions ϕ_E as basis in our Hilbert space, ρ is diagonal (theorem 2) :

(22) $$\rho_{EE'} = \delta_{EE'} \cdot \frac{e^{-\frac{E}{kT}}}{\sum e^{-\frac{E}{kT}}} \quad ;$$

Calling $\sum_E e^{-\frac{E}{kT}} = Z(T) = \mathrm{Tr}\left(e^{-\frac{H}{kT}}\right)$

we may write it in a general operator notation :

(23) $$\rho = \frac{e^{-\frac{H}{kT}}}{\mathrm{Tr}\left(e^{-\frac{H}{kT}}\right)} .$$

Now the denominator $\mathrm{Tr}\left(e^{-\frac{H}{kT}}\right) = Z(T)$ is a number and here only serves to normalize ρ to have $\mathrm{Tr}\,\rho = 1$. In statistical mechanics, however, this function of T is important because it defines the free energy of the system by :

(24) $$F = -kT \log Z(T) \quad , \text{thus } Z(T) = e^{-\frac{F}{kT}}$$ and then

(25) $$\rho = e^{\frac{F-H}{kT}}$$ (where F is a number times the "unit operator"; H, the Hamiltonian, is an operator.)

Here we have seen that ρ may even be expressed in an abstract manner as operator rather than by a matrix representation.

This operator may now be expanded in a series :

(26) $$\rho = \sum_n \frac{1}{n!}\left(\frac{F-H}{kT}\right)^n = \sum_n \alpha_n H^n$$

and for any operator we find the expectation value :

(27) $$\langle O \rangle_T = \mathrm{Tr}(\rho O) = \mathrm{Tr}\left(\sum_n \alpha_n H^n O\right) = \sum_n \alpha_n \mathrm{Tr}\left(H^n O\right)$$

A simple consequence is that $\langle O \rangle$ equals zero for any operator, for which $\mathrm{Tr}(H^n O)$ equals zero for all n. One intuitively sees that this applies, for instance, to the total angular and linear momentum, in the absence of external fields also to the di-electric and magnetic polarization.

b) Polarization of a Particle Beam

Consider a beam of spin $\frac{1}{2}$ particles and let us be interested only in the polarization. According to theorem 4, we have a 2 x 2 matrix with three independent parameters and consequently need three measurements to determine ϱ. We define a polarization vector $\vec{P}$ by the expectation values of the Pauli spin matrices

(28)
$$\begin{aligned} P_1 &= \langle\sigma_1\rangle = \mathrm{Tr}(\varrho\sigma_1) \\ P_2 &= \langle\sigma_2\rangle = \mathrm{Tr}(\varrho\sigma_2) \\ P_3 &= \langle\sigma_3\rangle = \mathrm{Tr}(\varrho\sigma_3) \end{aligned} \qquad \left| \qquad \begin{aligned} \sigma_1 &= \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} \\ \sigma_2 &= \begin{pmatrix} & -i \\ i & \end{pmatrix} \\ \sigma_3 &= \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \end{aligned} \right.$$

Obviously $P^2 = 1$ for a 100% polarized beam; $P^2 = 0$ for an unpolarized one. We may measure $\vec{P}$ by determining the magnetic dipole moment per cubic centimeter, $\vec{M}$

$$\vec{M} = N \cdot \mu \cdot \vec{P} \qquad (N = \text{Number of particles/cm}^3; \; \mu = \text{magnetic moment})$$

or we may even know $\vec{P}$ a priori if the beam comes out of a box of temperature T in an external magnetic field $\vec{\mathcal{H}}$.

This knowledge of $\vec{P}$ determines ϱ uniquely :

Since ϱ is 2 x 2, it must be a linear combination of the three σ's and the unit matrix :

(29)
$$\varrho = a \cdot 1 + \sum a_i \sigma_i \qquad ;$$

Since $\mathrm{Tr}\,\sigma_i = 0$, we have (from theorem 3) $\mathrm{Tr}\,\varrho = 2a = 1$; thus $a = \frac{1}{2}$.

Furthermore, from $\sigma_1\sigma_2 = i\sigma_3$ (cycl.) and $\sigma_k^2 = 1$ follows $\mathrm{Tr}(\sigma_i\sigma_k) = 2\delta_{ik}$ so that with (28) and (29) we have

$$P_i = \mathrm{Tr}\left[\tfrac{1}{2}\sigma_k + \sigma_k \sum a_i\sigma_i\right] = 2a_k \quad \text{and thus} \quad a_k = \tfrac{1}{2}P_k \, .$$

$$(30) \qquad \rho = \tfrac{1}{2}\left(1 + \sum P_i\sigma_i\right) = \tfrac{1}{2}\left(1 + \vec{P}\vec{\sigma}\right) = \tfrac{1}{2}\begin{pmatrix} 1 + P_3 & P_1 - iP_2 \\ P_1 + iP_2 & 1 - P_3 \end{pmatrix} .$$

Hence ρ is completely determined by $\vec{P}$ and vice versa. Assume now ρ to be known, either by a measurement of $\vec{P}$ or from thermodynamical equilibrium. This knowledge refers to a certain instant and ρ may now vary in time according to the equation of motion. Of course, if the beam is not influenced by any interaction, nothing happens. But suppose the beam to pass through a homegeneous magnetic field. What will be the result of a measurement after that? You of course intuitively know the answer because one may consider the whole beam as a magnetized piece of matter which will follow the classical laws and will therefore show a precession of the polarization vector. But we shall now derive this result from the equation of motion of ρ :

$$(31) \qquad \frac{d\rho}{dt} = -\frac{i}{\hbar}\left[H, \rho\right] \qquad \text{(theorem 5)}$$

The Hamiltonian of a particle in a magnetic field $\vec{\mathcal{H}}$ is

$$(32) \qquad H = -\gamma \frac{\hbar}{2}\vec{\sigma}\cdot\vec{\mathcal{H}}$$

; γ is the gyromagnetic ratio ($\frac{e}{mc}$ for electro

This is the only relevant part of the Hamiltonian, because ρ commutes with the rest of the terms of H. Thus

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}\left(-\frac{\gamma\hbar}{2}\right)\left[\vec{\sigma}\cdot\vec{\mathcal{H}}\, ,\ \tfrac{1}{2}\vec{P}\vec{\sigma}\right]$$

because the rest of ρ commutes with H.

Now simple algebra using $\sigma_1\sigma_2 = i\sigma_3$ (cycl.) and $\sigma_K^2 = 1$ leads to

$$\left\{\begin{aligned} &(\vec{a}\cdot\vec{\sigma})(\vec{b}\cdot\vec{\sigma}) = (\vec{a}\cdot\vec{b}) + i[\vec{a}\times\vec{b}]\cdot\vec{\sigma} \quad \text{for any } \vec{a}\,,\vec{b}\ \text{, thus} \\ &[\vec{a}\cdot\vec{\sigma}\, ,\ \vec{b}\cdot\vec{\sigma}] = 2i(\vec{a}\times\vec{b})\cdot\vec{\sigma} \end{aligned}\right. \tag{33}$$

In our case

$$\frac{d\rho}{dt} = -\frac{\gamma}{2}\left(\vec{\mathcal{H}}\times\vec{P}\right)\cdot\vec{\sigma} = -\frac{\gamma}{2}\sum_i\left(\vec{\mathcal{H}}\times\vec{P}\right)_i\sigma_i \tag{34}$$

We may now calculate $\dfrac{d\vec{P}}{dt}$ from this result :

$$\frac{dP_K}{dt} = \mathrm{Tr}\left(\frac{d\rho}{dt}\sigma_K\right) = -\frac{\gamma}{2}\sum_i(\vec{\mathcal{H}}\times\vec{P})\,\mathrm{Tr}(\sigma_i\sigma_K) = -\gamma\left(\vec{\mathcal{H}}\times\vec{P}\right)_K \qquad \text{or}$$

$$\frac{d\vec{P}}{dt} = -\gamma\left(\vec{\mathcal{H}}\times\vec{P}\right) \tag{35}$$

which is the conjectured classical formula. This tells us how the expectation value of the polarization precesses. Applied to a gas in a temperature bath in equilibrium, we find from the condition $\frac{d\vec{P}}{dt} = 0$ that $\vec{P}$ is parallel to $\vec{\mathcal{H}}$. One can extend this treatment as to obtain also the relaxation effects leading to thermodynamical equilibrium.

Having thus the density matrix (30) at a certain instant t = 0 and knowing the Hamiltonian acting on the system, we may calculate it for any time t > 0. The direction and degree of polarization is then given by

(36) $$\vec{P} = \mathrm{Tr}(\rho\vec{\sigma}) \quad ; \quad |P| = \text{degree of polarization.}$$

Let us now assume that we have a polarization analyser which shows the measured polarization on a scale like this

and points in the direction $\vec{Q}$ (unit vector) relative to which I wish to measure the polarization. Let the scale be normalized so that a completely polarized beam gives ± 1 respectively, if $\vec{Q}$ points parallel or anti-parallel to $\vec{P}$.

Again we intuitively know that if $\vec{P}$ is the polarization and $\vec{Q}$ the direction of the pointer, the instrument will show just the value

$\vec{Q} \cdot \vec{P}$, but we shall derive this result from a consideration on ρ .

We must represent the analyser by an operator $A(\vec{Q})$ so that

$\mathrm{Tr}(\rho A)$ is just the result to read off from the scale.

Since ρ is 2 x 2, A must be 2 x 2 also and has the form

$A(\vec{Q}) = \alpha(\vec{Q}) \cdot 1 + \vec{\beta}(\vec{Q}) \cdot \vec{\sigma} = \alpha + \sum \beta_i \sigma_i$. Then

$$\rho A = \tfrac{1}{2}(1 + \vec{P}\vec{\sigma})(\quad + \vec{\beta}\vec{\sigma}) = \tfrac{1}{2}\left(\alpha + \vec{\beta}\vec{\sigma} + \vec{P}\vec{\sigma} + \vec{P}\vec{\beta} + i(\vec{P} \times \vec{\beta}) \cdot \vec{\sigma}\right)$$

Then

$$\mathrm{Tr}(\rho A) = \alpha(\vec{Q}) + \vec{P} \cdot \vec{\beta}(\vec{Q})$$

If the beam is completely unpolarized ($\vec{P} = 0$), one should measure zero, hence $\alpha = 0$. If it is completely polarized, one should find the result cos $(\vec{P}, \vec{Q})$, thus $\beta(\vec{Q}) = \vec{Q}$. Hence

$A(\vec{Q}) = \vec{Q} \cdot \vec{\sigma}$ and the general case gives

$$\mathrm{Tr}(\rho A) = \vec{P} \cdot \vec{Q} = |P| \cdot \cos(\vec{P}\vec{Q})$$

Then the polarization with respect to the direction $\vec{Q}$ is shown directly on the scale, and the degree of polarization $|P|$ is found by varying the direction $\vec{Q}$ till one finds the maximum value; this value is just $|P|$.

c) The Polarization of Light

The term "a partially polarized beam of light" has a very simple meaning if the light can be decomposed into two oppositely polarized components. But what are these oppositely polarized components in general?

It seems that one of the most elegant and useful descriptions is that by means of the so-called Stokes-parameters.

We consider first a monochromatic polarized plane wave and put the z-axis of the coordinate system in its propagation direction. Then the most general totally (elliptically) polarized wave is

$$(37)\qquad \begin{pmatrix} E_1 \\ E_2 \end{pmatrix} = \begin{pmatrix} \varepsilon_1 \cos \omega t \\ \varepsilon_2 \cos(\omega t + \varphi) \end{pmatrix} = \begin{pmatrix} \frac{\varepsilon_1}{2} e^{i\omega t} \\ \frac{\varepsilon_2}{2} e^{i(\omega t + \varphi)} \end{pmatrix} + \begin{pmatrix} \frac{\varepsilon_1}{2} e^{-i\omega t} \\ \frac{\varepsilon_2}{2} e^{-i(\omega t + \varphi)} \end{pmatrix} \equiv \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} c_1^* \\ c_2^* \end{pmatrix} \equiv c + c^*$$

The three parameters $\varepsilon_1, \varepsilon_2$ and φ determine intensity and state of polarization.

$$(38)\qquad \left\{ \begin{array}{l} I = \frac{1}{2}\left(\varepsilon_1^2 + \varepsilon_2^2\right) \quad \text{is the intensity, whereas} \\ \\ \frac{\varepsilon_1}{\varepsilon_2} \quad \text{and} \quad \varphi \quad \text{determine the polarization.} \end{array} \right.$$

(The reader should recall these things by a title discussion).

We now use the last expression in (37) to introduce

$$(39)\qquad \left.\begin{array}{l} J_\kappa = (c, \sigma_\kappa c) = \sum\limits_{\lambda\mu} c_\lambda^* \sigma_\kappa^{\lambda\mu} c_\mu \\ J_0 = (c, c) = \sum\limits_{\lambda} c_\lambda^* c_\lambda \end{array}\right\} \quad \text{the "Stokes parameters".}$$

Here σ_i are the Pauli matrices (see page 27) and the scalar product is formally the same as if the two-component quantities were spinors. Using $\sigma_0 = 1$ one can write this very short as

$$\mathcal{J}_\mu = ({}^{*}c, \sigma_\mu c) \tag{39a}$$

For this case of a 100 % polarized wave $\mathcal{J}_\mu$ obeys formally the law

$$\mathcal{J}_\mu \mathcal{J}^\mu = \mathcal{J}_1^2 + \mathcal{J}_2^2 + \mathcal{J}_3^2 - \mathcal{J}_0^2 = 0$$

but this does not mean that it transforms like a four-vector under Lorentz-transformations. The components of $\mathcal{J}_\mu$ have nothing to do with space directions!

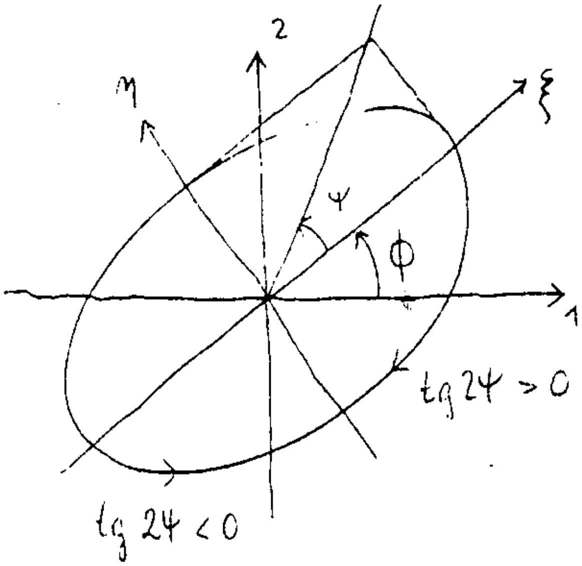
Going back to the common notation, we find

$$\tag{40} \left\{ \begin{aligned} \mathcal{J}_0 &= \tfrac{1}{4}(\varepsilon_1^2 + \varepsilon_2^2) = \tfrac{1}{2} I \\ \mathcal{J}_1 &= c_1^* c_2 + c_2^* c_1 = \tfrac{1}{2}\varepsilon_1 \varepsilon_2 \cos\varphi \\ \mathcal{J}_2 &= -i c_1^* c_2 + i c_2^* c_1 = \tfrac{1}{2}\varepsilon \varepsilon_2 \sin\varphi \\ \mathcal{J}_3 &= c_1^* c_1 - c_2^* c_2 = \tfrac{1}{4}(\varepsilon_1^2 - \varepsilon_2^2) \end{aligned} \right.$$

Note that $\mathcal{J}_\mu$ is not a linear function of the electric field!

$$\mathcal{J}_\mu(\vec{E}' + \vec{E}'') \neq \mathcal{J}_\mu(\vec{E}') + \mathcal{J}_\mu(\vec{E}'')$$

A very evident interpretation is the following :



In general the wave (37) is elliptically polarized and the state of polarization is uniquely defined by two parameters (see (38)) which we can choose as convenient. Two such parameters are the angles ϕ and ψ in the figure. One has the following relations :

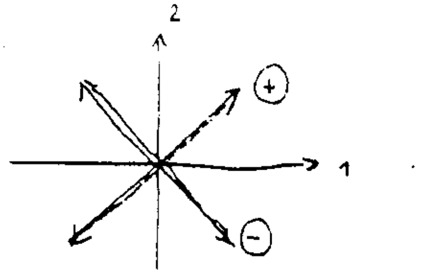
If (37) refers to the xy-syst
($\xi_1 \equiv x$; $\xi_2 \equiv y$) then with
tg $\psi = \eta/\xi$

$$\left\{\begin{aligned} tg\, 2\phi &= \frac{\mathcal{J}_1}{\mathcal{J}_3} \\ tg\, 2\psi &= \frac{\mathcal{J}_2}{\sqrt{\mathcal{J}_1^2+\mathcal{J}_3^2}} \end{aligned}\right\} \longleftarrow \tag{40a}$$

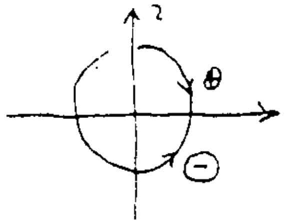
(the proof is left to the reader.)

From (40) and (40a) follows a simple interpretation of the components

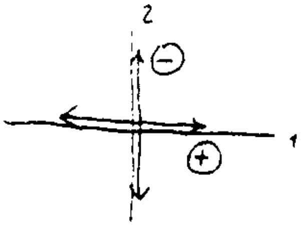
$$\begin{pmatrix}\mathcal{J}_1\\ \mathcal{J}_2\\ \mathcal{J}_3\end{pmatrix} = \begin{pmatrix}\pm 1\\ 0\\ 0\end{pmatrix} \text{ gives } \left.\begin{aligned}\phi &= \pm\frac{\pi}{4}\\ \psi &= 0\end{aligned}\right\} \text{that is}$$



$$\begin{pmatrix}0\\ \pm 1\\ 0\end{pmatrix} \text{ gives } \left.\begin{aligned}&\phi \text{ undetermined}\\ &\psi = \pm\frac{\pi}{4}\end{aligned}\right\}$$



$$\begin{pmatrix}0\\ 0\\ \pm 1\end{pmatrix} \text{ gives } \phi = \psi = 0$$



(The figures are understood as seen from behind the outgoing light : $\vec{k}$ is parallel to z-axis.)

This interpretation also defines the necessary measurements to determine $\mathcal{J}_\mu$ for a given beam.

So far, only completely polarized light was considered and now we turn to the most general case where the light is only partially polarized.

One gets a partially depolarized beam by adding incoherently two beams of opposite polarization. We must first make two things clear :

α) What does incoherent addition mean?

β) What does opposite polarization mean?

α) The concept of incoherent addition involves time and statistics.

If a "physical" source of light radiates, then the atoms are independent of each other and at any instant the wave is a complicated superposition of components of different amplitudes and phases. (We may assume here that the frequency is sharp. The question of frequency spread enters only if one tries to superpose coherently two beams coming from a single source.)

Thus at a certain instant t from all atoms together one has

$$\vec{E} = \begin{pmatrix} \int \varepsilon_1(\alpha) \cos(\omega t + \alpha)\, d\alpha \\ \int \varepsilon_2(\beta) \cos(\omega t + \beta)\, d\beta \end{pmatrix} = \begin{pmatrix} \eta_1 \cos(\omega t + \psi_1) \\ \eta_2 \cos(\omega t + \psi_2) \end{pmatrix}$$

(The reader may prove that it can always be written in the last form.)

Now the essential point is that the momentary total amplitudes η and phases ψ vary rapidly and randomly with time*), since they are the result of superposition of statistically emitted waves. Thus the general form is

$$\vec{E}(t) = \begin{pmatrix} \eta_1(t) \cos(\omega t + \psi_1(t)) \\ \eta_2(t) \cos(\omega t + \psi_2(t)) \end{pmatrix}.$$

---

*) Here our presupposition that ω is sharp becomes inconsistent. If τ is the "correlation time" of these statistical amplitude and phase functions, then the frequency is spread over an interval of the order $\Delta\omega \approx \tau^{-1}$

If two such beams come from different sources, they are statistically independent (the same is true if they come from the same source, but have a difference in optical path greater than the coherence length - for this the frequency spread is essential) and if they are mixed, we say they are added incoherently.

We note further that such a beam is completely depolarized, because η_1/η_2 and $\psi_2-\psi_1 = \Delta\psi$ are the parameters which determine the state of polarization and they vary rapidly and statistically with time. (This applies in the mean over a sufficiently long time. If, however, by some coincidence method one picks out only certain photons, then these may appear completely polarized.)

We wish now to superpose two polarized beams incoherently. Therefore, we must describe a polarized "physical" beam; that is, we must know how such a beam looks after having passed through a polarizing filter. A well defined polarization means :

$$\frac{\eta_1(t)}{\eta_2(t)} = \frac{\eta_1}{\eta_2} = \text{const.} \qquad \text{and}$$

$$\psi_2(t) - \psi_1(t) = \varphi = \text{const.}$$

and an ideal polarizer would just select only those components of the beam, which fulfil these conditions (in practice, intensity and sharpness of polarization are of course complementary). Thus for such a completely polarized beam

$$\left.\begin{aligned} \eta_1(t) &= \eta_1 \cdot f(t) \\ \eta_2(t) &= \eta_2 \cdot f(t) \end{aligned}\right\} \qquad \text{with the same } f(t)$$

$$\psi_2(t) - \psi_1(t) = \varphi \qquad = \text{const.}$$

In the notation of (37) this gives the completely polarized "physical" beam :

$$\vec{E} = \frac{f(t)}{2} e^{i\Psi(t)} \begin{pmatrix} \eta_1 e^{i\omega t} \\ \eta_2 e^{i(\omega t+\varphi)} \end{pmatrix} + \frac{f(t)}{2} e^{-i\Psi(t)} \begin{pmatrix} \eta_1 e^{-i\omega t} \\ \eta_2 e^{-i(\omega t+\varphi)} \end{pmatrix} = c(t) + c^*(t) \tag{41}$$

Assuming two different sources S' and S" (the prime in the following formulae do not mean derivatives), one has at any instant

$$\vec{E} = \vec{E}' + \vec{E}'' = \left[c'(t) + c^{*\prime}(t)\right] + \left[c''(t) + c''^{*}(t)\right]$$

and this becomes an "incoherent" superposition only due to the fact that one observes in general over time intervals, which are long as compared with the "mean fluctuation time" of the c's. From (39a) then follows for any instant

$$J_\mu(t) = J'_\mu(t) + J''_\mu(t) + 2\,\mathrm{Re}\left(c'(t), \sigma_\mu c''(t)\right)$$

The Stokes vector is thus a rapidly varying quantity, but if we go over to the time mean values , we find that

$$\overline{J_\mu} = \overline{J'_\mu} + \overline{J''_\mu} \qquad \text{(incoherent superposition)} \tag{42}$$

because $\overline{J'_\mu}$ and $\overline{J''_\mu}$ involve factors $\frac{1}{T}\int_0^T [f'(t)]^2\,dt$ and $\frac{1}{T}\int_0^{T} [f''(t)]^2\,dt$ whereas in the mixed expression every term contains factors like

$$\frac{1}{T}\int_0^T f'(t)\, f''(t) \cos(\Psi' \pm \Psi'')\,dt = 0 \quad ,$$

which vanish because $f'(t)$ and $f''(t)$ as well as $\psi'(t)$ and $\psi''(t)$ are statistically independent. Thus we have defined what we mean by incoherent addition and have found at the same time a very simple formulation (4[illegible]), which indeed is nothing else than a generalization of the well known fact that the intensities add algebraicly in such a superposition (remember $J_o = \frac{1}{2} I$).

β) The second question is rather different. We have seen that incoherent beams do not interfere, because their intensities simply add. The reason was that their amplitudes and phases are statistically independent. But there is another way to make sure, that two beams connot interfere, even if their amplitudes and phases are not statistically independent : That is the case if they have "opposite polarization". This is well known, for instance, for two perpendicularly linearly polarized waves. We therefore define :

> Two beams are oppositely polarized if and only if they can never interfere whatever their phase relations may be.

We shall thus get the most general description of what "opposite polarization" means if we take two beams which otherwise would have the best chance to interfere and ask for the necessary and sufficient condition that they do <u>not</u> interfere. The best chance for interference is of course present with two beams of the form (41), which are completely correlated, that is, both have the same $f(t)$ and $\psi(t)$ but for a constant phase difference α and a constant difference in amplitude : They are coherent. Two such beams are made by splitting a beam from a single source into two components and letting them have different paths. From (41) we see that they would be described by

$$c'(t) = \frac{f(t)}{2} e^{i\psi(t)} \begin{pmatrix} \varepsilon_1' e^{i\omega t} \\ \varepsilon_2' e^{i(\omega t + \varphi)} \end{pmatrix} ; \quad c''(t) = \frac{f(t)}{2} e^{i\psi(t)} e^{i\alpha} \begin{pmatrix} \varepsilon_1'' e^{i\omega t} \\ \varepsilon_2'' e^{i(\omega t + \varphi'')} \end{pmatrix}$$

with $f(t)$ and $\psi(t)$ being the same in both beams.

The necessary and sufficient condition for non-interference is, that

$$J_0 = J_0' + J_0''$$

that is the mixed term in the intensity vanishes

$$Re\,\overline{(c', \sigma_0 c'')} = \frac{1}{4} \cdot \overline{[f(t)]^2} \cos\alpha \cdot \left(\varepsilon_1' \varepsilon_1'' + \varepsilon_2' \varepsilon_2'' \cos(\varphi'' - \varphi')\right) = 0$$

Thus $\varepsilon_1' \varepsilon_1'' + \varepsilon_2' \varepsilon_2'' \cos(\varphi'' - \varphi') = 0$ If we put $\varphi'' = \varphi'$, we have

$$\varepsilon_1' \varepsilon_1'' + \varepsilon_2' \varepsilon_2'' = 0$$

(The reader may assure himself, that by putting $\varphi'' = \varphi'$ we loose nothing). Now with $\varepsilon_1'' = a\varepsilon_2'$; $\varepsilon_2'' = b\varepsilon_1'$ we get a + b = 0, hence

$$\left\{ \begin{array}{l} \varphi'' = \varphi' \\ \varepsilon_1'' = a\varepsilon_2' \\ \varepsilon_2'' = -a\varepsilon_1' \end{array} \right. \tag{43}$$

are the necessary and sufficient conditions for opposite polarization. From (40) fellows immediately

$$J_0'' = \frac{1}{4}\left(\varepsilon_1''^2 + \varepsilon_2''^2\right) = \frac{1}{4} a^2 \left(\varepsilon_1'^2 + \varepsilon_2'^2\right) = a^2 J_0'$$

$$J_1'' = \frac{1}{2} \varepsilon_1'' \varepsilon_2'' \cos\varphi'' = -\frac{1}{2} a^2 \varepsilon_1' \varepsilon_2' \cos\varphi' = -a^2 J_1'$$

$$J_2'' = \frac{1}{2} \varepsilon_1'' \varepsilon_2'' \sin\varphi = -\frac{1}{2} a^2 \varepsilon_1' \varepsilon_2' \sin\varphi' = -a^2 J_2'$$

$$J_3'' = \frac{1}{4}\left(\varepsilon_1''^2 - \varepsilon_2''^2\right) = \frac{1}{4} a^2 \left(\varepsilon_2'^2 - \varepsilon_1'^2\right) = -a^2 J_3'$$

We get therefore the simple result :

The necessary and sufficient condition for two completely polarized beams J_μ' and J_μ'' having opposite polarization, is (with $0 \leq a^2 \leq 1$)

$$\left(J_0'', J_1'', J_2'', J_3''\right) = \left(a^2 J_0', -a^2 J_1', -a^2 J_2', -a^2 J_3'\right) \tag{44}$$

($a^2 \leq 1$ is no restriction; one can, if necessary, interchange the notation).

The reader may discuss as an exercise what is the opposite to a given elliptical polarization (use (40) and (40a)). Result : the ellipse is turned by 90^0 and the sense of rotation of the electric vector is reversed.

Now, we get a partially depolarized beam by incoherent superposition of two completely polarized beams with opposite polarizations:

$$J_\mu = J'_\mu + J''_\mu \;; \qquad \left.\begin{matrix} J''_0 = a^2 J'_0 \\ \vec{J}'' = -a^2 \vec{J}' \end{matrix}\right\} \text{ hence}$$

(45) $$J_0 = (1+a^2)\, J'_0$$ $$\vec{J} = (1-a^2)\, \vec{J}'$$

For completely polarized beams we had $\vec{J}^2 - J_0^2 = 0$. Here we find

(46) $$\vec{J}^2 - J_0^2 = (1-a^2)\vec{J}'^2 - (1+a^2) J_0'^2 = -4a^2 J_0'^2 \le 0$$

(We have used $\vec{J}'^2 - J_0'^2 = 0$ since that beam is completely polarized). We see that for a partially depolarized beam $\vec{J}^2 - J_0^2$ is always ≤ 0, the equality holding for complete polarization.

We may define the degree of polarization by a number $0 \le K \le 1$

(47) $$K = \frac{I'-I''}{I'+I''} = \frac{1-a^2}{1+a^2} = \frac{\sqrt{J_1^2+J_2^2+J_3^2}}{J_0} \qquad (I' > I'')$$

(48)

We collect these results :

A partially polarized beam of light is obtained by adding incoherently two beams of opposite (pure) polarization. Its stokes vector is :

$$J_\mu = \begin{pmatrix} J_0 \\ J_1 \\ J_2 \\ J_3 \end{pmatrix} \qquad \vec{J}^2 - J_0^2 \le 0 \qquad J_0 = \frac{1}{2} I$$

degree of polarization $K = \dfrac{\sqrt{\vec{J}^2}}{J_0}$

$0 \le K \le 1$

(The interpretation in terms of measurable quantities is given on p. 34

We have shown, that a partially depolarized beam is obtained by adding incoherently two beams of opposite polarization, In order to see that this is already the most general case, we must still show, that any beam can be decomposed uniquely into two incoherent completely polarized beams (with the obvious exception of a completely unpolarized beam).

If a beam is given in the form (48) - i.e. we suppose the components J_μ to be measured (according to the interpretation on p.34) by suitable analyzers - can one decompose it uniquely into two completely polarized beams of opposite polarization ?

The answer is yes. But

$$J_\mu = J'_\mu + J''_\mu = \begin{pmatrix} J_0 \\ J_1 \\ J_2 \\ J_3 \end{pmatrix} \tag{49}$$

The two beams J'_μ and J''_μ shall be of opposite polarization and each one 100% polarized, thus

$\vec{J}'' = -a^2 \vec{J}'$

$J''_0 = a^2 J'_0$. Hence

$$\left\{ \begin{aligned} \vec{J} &= \vec{J}' + \vec{J}'' = (1-a^2)\vec{J}' \\ J_0 &= J'_0 + J''_0 = (1+a^2) J'_0 \end{aligned} \right. \tag{50}$$

From the given J_μ follows the degree of polarization (47)

$$K = \frac{\sqrt{J_1^2+J_2^2+J_3^2}}{J_0} = \frac{1-a^2}{1+a^2}\,\frac{\sqrt{J_1'^2+J_2'^2+J_3'^2}}{J'_0} = \frac{1-a^2}{1+a^2} ,$$

since $\vec{J}'^2 = J_0'^2$ (100% polarization). Thus, we can solve for a^2;

$1 + a^2 = \dfrac{2}{1+K}$; $1 - a^2 = \dfrac{2k}{1+k}$, and with (46) and (47) for $K \neq 0$:

$$J_\mu = J'_\mu + J''_\mu = \frac{1+K}{2}\begin{pmatrix} J_0 \\ J_1/K \\ J_2/K \\ J_3/k \end{pmatrix} + \frac{1-K}{2}\begin{pmatrix} J_0 \\ -J_1/k \\ -J_2/K \\ -J_3/K \end{pmatrix} \; ; \quad K = \frac{\sqrt{J_1^2+J_2^2+J_3^2}}{J_0} \tag{51}$$

Thus we have decomposed the given beam in a unique way into two 100% polarized ones with opposite polarization. The polarization states of these two beams are then interpreted according to the rules of p. 34. For K = 0 the decomposition is no longer unique since $J_1 = J_2 = J_3 = K = 0$ One takes any two oppositely polarized beams of equal intensity.

Finally we go over from this classical description to quantum electrodynamics. It is here not the place to do it rigorously, it can be sketched only (see e.g. Jauch and Rohrlich).

The c's of (37) are now operators :

(52) $$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \longrightarrow \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

where $a_\mu(\vec{k})$ and $a^+_\mu(\vec{k})$ are destruction and creation operators for light quanta of wavevector $\vec{k}$. The coordinates are the same as before, $\vec{k}$ points into the positive z -direction. The commutation relations are

(53) $$\left[a_\mu(\vec{k}) , a_{\mu'}(\vec{k}') \right] = \delta_{\mu\mu'} \, \delta_{kk'}$$

The stokes parameters J_μ are thus replaced by operators :

(54) $$J_\mu \rightarrow \Sigma_\mu = \left(a_1^+ a_2^+ \right) \sigma_\mu \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \equiv \left(a , \sigma_\mu a \right)$$

and the classical stokes parameters are now defined as expectation values :

(55) $$J_\mu = \mathrm{Tr} \left(\rho \, \Sigma_\mu \right)$$

As we may always assume J_μ to be normalized such that $J_0 = 1$, we have just 3 parameters defining the state and degree of polarization of our beam. This is the same situation as for a polarized particle beam, we expect therefore ρ to have essentially the same structure as for particle beams (30).

In fact, from the commutation relations (53) follows that

$$(56)\qquad \begin{cases} [\Sigma_1, \Sigma_2] = 2i\Sigma_3 \quad (cycl.) \; ; \; [\Sigma_i, \Sigma_0] = 0 \\ \Sigma_1^2 + \Sigma_2^2 + \Sigma_3^2 = \Sigma_0(\Sigma_0 + 2) \end{cases}$$

The lowest order representation of these rules is just given by the Pauli σ matrices (p. [illegible]) and $\Sigma_0 = \sigma_0 = 1$

As it must be, we have then

$$(57)\qquad J_0 = Tr(\rho\Sigma_0) = Tr\,\rho = 1 \; ; \; \vec{J} = Tr(\rho\vec{\sigma}) \quad \text{(compare with (28))}$$

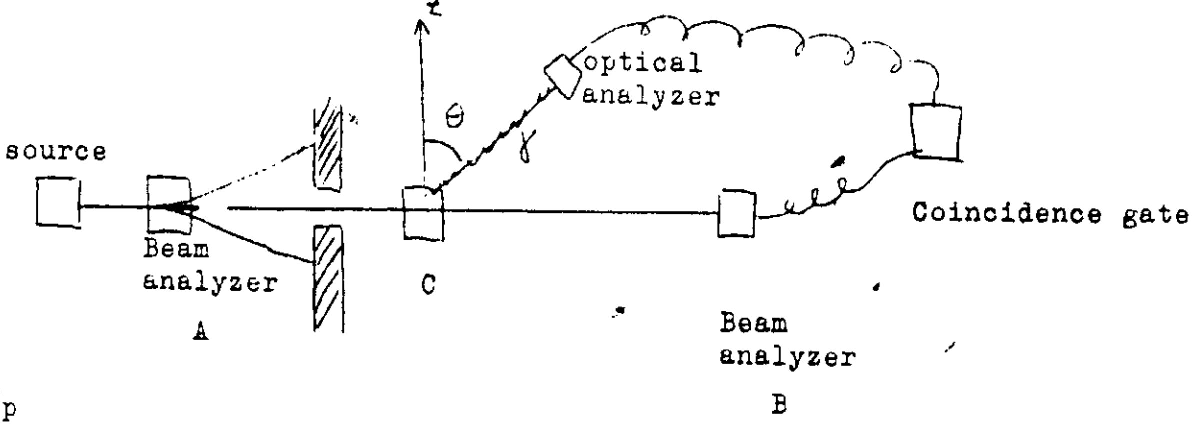
The density matrix is thus in complete analogy to (28) and (30):

$$(58)\qquad \rho = \tfrac{1}{2}\left(1 + \vec{J}\,\vec{\sigma}\right)$$

Note, however, the important difference between $\vec{J}$ and P . For particles the polarization vector $\vec{P}$ gives the degree and direction of the polarization, its components refer directly to space directions. For a light beam $\vec{J}$ defines the polarization according to the interpretation on p. 34 . The components of $\vec{J}$ have nothing to do with space directions !

d) An idealized experiment

In his review article (see references) Fano discusses an idealized experiment, which we shall consider as an illustration of the foregoing formalism.



A beam of excited atoms is analyzed for m (magnetic quantum number) then in a box C it emits radiation which is analized also and finally the be is again analyzed in coincidence with the optical analyzer. Under thes idealized conditions one measures m before and after emission and analyzes the light quantum of this atom.

By a calculation, which goes beyond the scope of this lecture, one can under certain assumptions (transition from $\ell = 1$, $m = 0$) to $\ell' = 2$, $m' = 0, \pm 1$) determine the density matrix for the emitted radiation.

Fano discusses several cases, of which we shall select here two and simplify them even by disregarding the intensities (normalizing them to unity).

α) B is not used, A selects atoms with m = 0. For slow atoms one finds for the density matrix

$$\rho = \frac{1}{2}\left(1 + \frac{\sin^2\theta}{6+\sin^2\theta}\cdot\sigma_3\right) \tag{59}$$

From (58) by comparing components one has immediately

$$\begin{cases} J_0 = 1 & \text{(by definition)} \\ J_1 = 0 \\ J_2 = 0 \\ J_3 = \dfrac{\sin^2\theta}{6+\sin^2\theta} \end{cases} \tag{60}$$

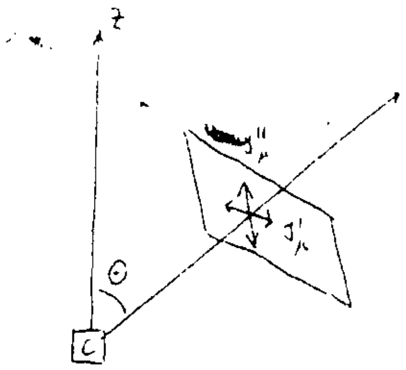
In the direction θ the degree of polarization is thus from (48)

$K = \dfrac{\sin^2\theta}{6+\sin^2\theta}$; hence the light is not 100% polarized (for $\theta = 0$ completely depolarized) and we may ask for a decomposition into 100% polarized beams. From (51) :

$$J'_\mu = \begin{pmatrix} \frac{1+K}{2} \\ 0 \\ 0 \\ \frac{1+K}{2} \end{pmatrix} \qquad J''_\mu = \begin{pmatrix} \frac{1-K}{2} \\ 0 \\ 0 \\ -\frac{1-K}{2} \end{pmatrix} \qquad K = \frac{\sin^2\theta}{6+\sin^2\theta} \tag{61}$$

The discussion follows the scheme of p. 34

For $\theta = 0$; the light is not polarized and may be decomposed arbitrarily. For $\theta \neq 0$, one has linearly polarized components:



β) Let B select in coincidence (with the observed photons) atoms with m' = 0, + 1, - 1. Then ρ is

m' = 0; $\rho = \frac{1}{2}(1 + \sigma_3)$; $J_\mu = (1,001)$; $K = 1$

m' = ± 1; $\rho = \frac{1}{2}\left(1 - \frac{\sin^2\theta \cdot \sigma_3 \mp 2\cos\theta\, \sigma_2}{1 + \cos^2\theta}\right)$

$$J_\mu = \left(1,\ 0,\ \pm \frac{2\cos\theta}{1+\cos^2\theta},\ - \frac{\sin^2\theta}{1+\cos^2\theta}\right);\ K = 1$$

The discussion is left to the reader, see p. 34 .

References

The most recent and general article is by

U. Fano, Rev. mod. phys. 29, 74 (1957)

there one finds some of the things treated here, but many others too, including a large list of further references.

As further application and exercise one may read the article on polarization of fast nucleons by

L. Wolfenstein, Ann. Rev. of Nucl. Science 6, 43 (1956)

which also includes a big list of references.

A more involved use of density matrices is made in many articles by

H.A. Tolhoek and S.R. de Groot, e.g. Physica 15, 833 (1949), Phys. Rev. 83, 189 (1951)

and in the review article on angular correlation by

L.C. Biedenharn and M.E. Rose, Rev. mod. phys. 25, 729 (1953).

A general treatment will presumably be found in the book.

U.Fano and G. Racah, Irreducible tensors (Academic Press Inc. New York) which will appear in 1958).