Stochastics, 1975 Vol. 1, pp. 315–345

Printed in Great Britain

Optimal Multiple Quantum Statistical Hypothesis Testing

V. P. BELAVKIN

M.I.E.M., B. Vusovski Street 3/12, Moskow 109028, U.S.S.R.

This paper is concerned with the problem of optimal M-alternative determination of quantum statistical states. A review of newest achievement of solving this problem is given. A notion of an effective decision Hilbert space is introduced and necessary and sufficient conditions for optimality of multiple quantum hypothesis testing in this space are formulated. The general solution is found for the case of a two-dimensional decision space. Another problem solved is that of discrimination of quantum pure non-orthogonal states. The result is represented in explicit analytical form for an "equidiagonal" case, which is quite general. In particular, we find explicit solutions of optimal discrimination problem of homogeneous and equiangle sets of pure states. These results are used for the M-ary detection problem in solving for the quantum coherent non-orthogonal signals. It is proved that the simplex signals are optimal also in quantum case. The optimal estimates of phase and amplitude of quantum coherent signals are found. For decision operators a notion of Π-representation is introduced to get a general quasi-classical (optimal in quasi-classical limit) M-ary detection procedure of stochastic fields and particles, which submits to Bose-Einstein statistics. An optimal solution of problem of non-coherent detection of quantum stochastic (including optical) signals are found in the extreme quantum limit (weak noise and signals with unknown phase).

1. INTRODUCTION

The solution of the quantum states testing problem of interest in itself enables one to estimate the potential possibilities of optical-digital communication systems and to formulate the fundamental limitations of their performance caused by quantum-mechanical nature of electromagnetic field. However until recently, there were very few cases, for which the exact solution had been obtained.

The quantum generalization of the problem of selection among a few probability distributions $P_1(dy), \ldots, P_M(dy)$ on the space of observed data y, or, as one also calls it, the problem of multiple statistical hypothesis testing, has been formulated by Helstrom [1]. Let M hypotheses be associated with M possible statistical states of observable quantum system (for example, the quantum electromagnetic field on aperture). In the general case the quantum statistical states are described by the density operators ρ acting in some complex Hilbert space $\mathscr{H}$ called the observation space. One should remember that the (Hermitian) density operators are non-negative definite: $\rho \geqq 0$, have the unit trace: $\mathrm{Tr}\rho = 1$, and correspond to the probability distributions in non-quantum (classical case).

The mathematical solution of the hypothesis testing problem associated with quantum states $\rho_1, \ldots, \rho_M$ reduces to finding Hermitian "decision" operators $\Pi_1, \ldots, \Pi_M$ by analogy to classical statistical decision functions. The operators $\{\Pi_j, j = \overline{1, M}\}$ define the conditional probabilities $\Pr\{j|k\}$ of decisions:

$$\Pr\{j|k\} = \mathrm{Tr}\Pi_j\rho_k \equiv P_{jk}, \tag{1.1}$$

and satisfy therefore the conditions $\Pi_j \geqq 0$ for all $j = \overline{1, M}$ and $\sum_{j=1}^{M}\Pi_j = \mathbf{1}$, where $\mathbf{1}$ is the identity operator. Unlike [1], according to a modern approach [2–6], we shall not require the conditions of orthogonality $\Pi_k\Pi_l = 0$ for all $k \neq l$ and commutativity $\Pi_k\Pi_l = \Pi_l\Pi_k$. It corresponds to the admission of randomized decision rules, based on the date of quantum indirect measurements, or quasi-measurements, at the expense of an observation over another accessible system. Such extension of quantum measurements and decision rules enables one to estimate noncommuting observables of the received quantum signal [7].

According to the general Bayes approach. to select among the sets of decision operators $\{\Pi_j\}$ an optimal set $\{\Pi_j^0\}$ one has to give a matrix of costs $\{C_{jk}\}$ and prior probabilities $\{\pi_k\}$ and to minimise the average cost (or risk)

$$C = \sum_{j,k=1}^{M} P_{jk}C_{jk}\pi_k = \mathrm{Tr}\sum_{j,k=1}^{M} \Pi_j C_{jk}\rho_k\pi_k. \tag{1.2}$$

For the case $M = 2$ the solution of this minimization problem has been found by Helstrom [1]. The binary detection of pure quantum states has been also considered in [8]. For M-alternative situations ($M > 2$) there have been no exact solutions until recently except for the special commutative case; $\rho_k\rho_l = \rho_l\rho_k$ for all k, l. This commutative case is of some practical interest [9], but is degenerative from the point of view of quantum theory: as because of commutative all density operators $\{\rho_k\}$ have a common diagonal representation, in which they are described by non-quantum probability distributions $\{P_k(dy)\}$, and their optimal discrimination is carried out by observing y according to the ordinary rules of the classical statistical decision theory.

This paper contains more general and exact results. We treat here several general cases of optimization of many alternative quantum statistical hypothesis testing which are exactly solved. In the next paragraph, necessary notions will be introduced and conditions of optimality of multiple testing will be formulated and analysed. Then a general quantum optimization problem will be solved for the case of two-dimensional decision space. In the fourth and fifth paragraph we shall consider and solve a problem of optimal discrimination of several not necessary orthogonal and linearly independent quantum pure states for the minimum average error probability criterion. The last paragraph is concerned with the general problem of M-ary detection of optical stochastic signals. The solution of this problem is found in two asymptotic cases: in quasi-classical approximation (powerful signals) and in extreme quantum limit (weak signals).

2. THE CONDITIONS OF OPTIMALITY AND SUFFICIENT DECISION SPACES

The necessary and sufficient conditions for the set of decision operators Π_j^0 to be an optimum solution of the multiple quantum hypothesis testing problem have been recently reported in [2, 3, 10, 11]. Let us write these conditions in the form

$$(A_j - \Lambda)\,\Pi_j = 0, \qquad \Pi_j(A_j - \Lambda) = 0, \tag{2.1}$$

$$A_j - \Lambda \geqq 0 \qquad \text{for all } j = \overline{1, M}, \tag{2.2}$$

where $A_j = \sum_{k=1}^{M} C_{jk}\rho_k\pi_k$ are Hermitian operators of posterior risk. Being defined by the summation of the Eqs. (2.1) extended over j with the condition $\sum_{j=1}^{M}\Pi_j = \mathbf{1}$ the operator Λ is Hermitian:

$$\Lambda = \sum_{j=1}^{M} A_j\Pi_j^0 = \sum_{j=1}^{M} \Pi_j^0 A_j, \tag{2.3}$$

where $\{\Pi_j^0\}$ is a solution of these equations.

The sufficiency of the Eqs. (2.1), (2.2) is almost obvious: if operators $\{\Pi_j^0\}$ satisfy the Eqs. (2.1) then a corresponding average cost (1.2) equals $\mathrm{Tr}\Lambda$, and for any set $\{\Pi_j\}$, for which $(A_j - \Lambda)\Pi_j \neq 0$, at least for one operator Π_j, the difference of costs $C = \mathrm{Tr}\sum_j A_j\Pi_j$, $C^0 = \mathrm{Tr}\Lambda$ satisfies

$$C - C^0 = \mathrm{Tr}\left(\sum_{j=1}^{M} A_j\Pi_j - \Lambda\right) = \sum_{j=1}^{M}\mathrm{Tr}\,(A_j - \Lambda)\,\Pi_j \tag{2.4}$$

and is positive, if $A_j - \Lambda \geqq 0$, $\Pi_j \geqq 0$.

The reduction of the equations $(A_j - \Lambda)\,\Pi_j = 0$ as necessary condition of optimality may be seen for instance in [3], and proof of necessity of inequalities (2.2) may be seen in [11].

Let us consider some set of Hermitian operators $A_1, \ldots, A_M$ and an operator lower bound $\Lambda \leqq A_j$, $j = \overline{1, M}$. It is obvious that for any set of bounded operators $\{A_j\}$, $\|A_j\| < \infty$ there is at least one such operator Λ. It is easy to prove that the operator Λ has the maximum trace $\mathrm{Tr}\Lambda = \max_{\Lambda \leq \{A_j\}}$ if and only if it satisfies the Eqs. (2.1) for some resolution of the identity $\Sigma\Pi_j = \mathbf{1}$. Defined by the Eqs. (2.1) a maximum lower bound $\Lambda \leqq A_j$, $\mathrm{Tr}\Lambda = \max$ is a quantum analog of the infimum of the set $\{A_j\}$: $\Lambda = \inf\{A_j\}$. When all operators A_j commute, this quantum infimum has the usual associative property: $\inf\{A_j, \ldots, A_M\} = \inf\{\inf\{A_1, \ldots, A_{M-1}\}\ A_M\}$, but in the noncommutative case it is invalid. Let us prove that the quantum infimum (2.3) is unique even if the set $\{\Pi_j\}$ satisfying the conditions (2.1) is not unique. Let $\Sigma\Pi_j^0 = \mathbf{1}$, $\Sigma\Pi_j^1 = \mathbf{1}$ be two different resolutions of identity, which define two Hermitian operators $\Lambda^0 = \Sigma A_j\Pi_j^0$, $\Lambda^1 = \Sigma A_j\Pi_j^1$. As soon as the trace of a product of any two non-negative operators is non-negative and equals to zero if and only if this product is zero operator, the inequalities

$$\mathrm{Tr}(\Lambda^1 - \Lambda^0) = \sum_j \mathrm{Tr}(A_j - \Lambda^0)\,\Pi_j^1 \geqq 0,$$
$$\mathrm{Tr}(\Lambda^0 - \Lambda^1) = \sum_j \mathrm{Tr}(A_j - \Lambda^1)\,\Pi_j^0 \geqq 0$$

are valid, from which $(A_j - \Lambda^0)\,\Pi_j^1 = 0$, $(A_j - \Lambda^1)\,\Pi_j^0 = 0$. Extending the summation over j, we obtain $\Sigma A_j\Pi_j^1 = \Lambda^0$, $\Sigma A_j\Pi_j^0 = \Lambda^1$, or $\Lambda^0 = \Lambda^1 \equiv \Lambda$.

Let B_j denote the difference $A_j - \Lambda$ for every $j = \overline{1, M}$, range B_j denote the subspace $B_j\mathscr{H}$, and ker B_j denote the kernel of the operator B_j: $B_j \ker B_j = \varnothing$ which is an orthogonal complement to the range B_j for any Hermitian operator B_j. According to Eqs. (2.1) the range B_j is a subset of the kernel of corresponding decision operator Π_j^0: range $B_j \subseteq \ker \Pi_j^0$. Hence it follows from the condition $\Sigma\Pi_j = 1$ that the intersection $\bigcap$ range B_j of all ranges of B_j is the empty subspace $\varnothing$. Let us prove the following theorem using notation $\bigcup$ range B_j for algebraic sum of all range B_j.

THEOREM 1 *It is sufficient for solving the problem of optimal quantum hypothesis testing to find a set of non-negative Hermitian operators* $\{D_j, j = \overline{1, M}\}$ *acting in any subspace* $\mathscr{U} \subseteq \mathscr{H}$ *for which* $\bigcup_j$ range $B_j \subseteq \mathscr{U}$ *and satisfying equations*

$$B_jD_j = 0, \quad \sum_{j=1}^{M} D_j = I,$$

where 0 *is the zero operator, and* I *is the identity in* $\mathscr{U}$.

Proof The case $\mathscr{U} = \mathscr{H}$ is trivial, and we shall suppose that $\mathscr{U} \subset \mathscr{H}$, i.e. the orthogonal complement $\bar{\mathscr{U}}$ is not empty. Let $\bar{0}$, $\bar{I}$ be the zero and identity acting in $\bar{\mathscr{U}}$. The direct sum $I \oplus \bar{0}$ is then the projector from $\mathscr{H}$ to $\mathscr{U}$ and direct sum $I \oplus \bar{I}$ is the identity $\mathbf{1}$ acting in $\mathscr{H} = \mathscr{U} \oplus \bar{\mathscr{U}}$. As soon as every

range B_j is insertion of $\mathscr{U}$, the operator projections $(I \oplus \bar{0})B_j(I \oplus \bar{0})$ coincide with B_j. Taking this into account and using the cyclic invariance of the trace, let us rewrite the risk (1.2) as follows:

$$C = \mathrm{Tr}\Lambda + \mathrm{Tr}\sum_{j=1}^{M} B_j(I \oplus \bar{0})\,\Pi_j(I \oplus \bar{0}) = C^0 + \mathrm{Tr}\sum_{j=1}^{M} B_jD_j. \tag{2.5}$$

Hence it follows that the risk (2.5) for any resolution of the identity $I \oplus \bar{I} = \Sigma\Pi_j$ depends only on the projections $D_j = (I \oplus \bar{0})\,\Pi_j(I \oplus \bar{0})$ of the decision operators D_j to the subspace $\mathscr{U}$. These operators D_j induced in $\mathscr{U}$ by Π_j are non-negative, satisfy the constraint $\Sigma D_j = I$, and we shall call them decision operators also. The equations $B_jD_j = 0$ follows from the equations $B_j\Pi_j = 0$ and the identity $B_j = B_j(I \oplus \bar{0})$. It is not difficult to imbed any resolution of identity $\sum_{j=1}^{M} D_j = I$ in the resolution $\sum_{j=1}^{M} \Pi_j = I \oplus \bar{I}$ of complete Hilbert space $\mathscr{H} = \mathscr{U} \oplus \bar{\mathscr{U}}$ by operators Π_j, whose projections into $\mathscr{U}$ coincide with D_j. For instance, it is sufficient to take $\Pi_j = D_j \oplus \bar{0}$, $j = \overline{1, M-1}$, $\Pi_M = D_M \oplus \bar{I}$. The theorem is proved.

Let us call the subspace $\mathscr{V} = \bigcup_j$ range B_j the effective decision space, and any subspace $\mathscr{U} \subseteq \mathscr{H}$ containing $\mathscr{V}$ let's call the sufficient decision space. In accordance with the proved theorem it is not necessary to constrain oneself by some observation Hilbert space $\mathscr{H}$, and one ought to restrict oneself by effective Hilbert subspace $\mathscr{V}$ or by a sufficient decision space $\mathscr{U} \supseteq \mathscr{V}$ as simple as possible. Acting in $\mathscr{U}$ operators D_j play the role of a sufficient statistic, even if they do not commute: it is possible by Naimark's theorem [12] to imbed any non-orthogonal resolution of the identity $I = \Sigma D_j$ in some orthogonal resolution of supplemented in a corresponding manner Hilbert space $\mathscr{H} \supset \mathscr{U}$ which describes some quantum-mechanical measurement procedure.

The theorem 1 simplifies very much the general solution of the problem of multiple quantum hypothesis testing in those cases, when operators of the posterior risk are represented in the form $A_j = A - R_j$, where the set of Hermitian operators R_j, $j = \overline{1, M}$ have the sum of ranges $\mathscr{U} \equiv \bigcup_j$ range R_j with a finite dimension†. Since range $(Q+R) \subseteq$ range $Q \bigcup$ range R, and range $(RQ) \subseteq$ range R for any operators Q, R, the range of difference $K \equiv A - \Lambda = \Sigma_j R_j \Pi_j^0$ is a subspace of $\mathscr{U} = \bigcup_j$ range R_j, and range $B_j \subseteq \mathscr{U}$ for all $B_j \equiv A_j - \Lambda = K - R_j$. Therefore the finite-dimensional subspace $\mathscr{U} = \bigcup_j$ range R_j is a sufficient decision space, and according to the theorem 1 the problem (2.1)–(2.2) reduces to the problem of finding a Hermitian

†The following consideration is also useful in a more general case, when $\mathscr{U}$ is an infinite orthogonal sum $\oplus\ \mathscr{U}^{(n)}$ of finite dimensional subspaces $\mathscr{U}^{(n)}$ which are invariant for all R_j (see §6, i.3.2). The optimization problem reduces to a set of independent problems (2.6) for every family $\{R_j^{(n)}, j = \overline{1, M}\}$ of operators $R_j^{(n)}$ induced in $\mathscr{U}^{(n)}$ by R_j.

operator $K \equiv \sup\{R_j\}$ and a resolution of the identity $I = \sum_{j=1}^{M} D_j$ in the finite dimensional subspace $\mathscr{U}$ which satisfy the conditions

$$K - R_j \geqq 0, \qquad (K - R_j)\, D_j = 0. \tag{2.6}$$

It is obvious that the operators R_j can always be made non-negative by a suitable choice of the operator A defining the resolution of posterior risks $A_j = A - R_j$. From the inequalities $K \geqq R_j$ for any operators $R_j \geqq 0$, $j = \overline{1, M}$ it follows that range $K \supseteq$ range R_j for all j. Hence the operator

$$K = \sum_{j=1}^{M} R_j D_j = \sum_{j=1}^{M} D_j R_j$$

satisfying inequalities $K \geqq R_j \geqq 0$, $j = \overline{1, M}$ is strictly positive in $\mathscr{U} = \bigcup_j$ range R_j, i.e. range $K = \mathscr{U}$. Let us prove the following theorem.

THEOREM 2 *For any Hermitian operators $\{R_j\}$ and a Hermitian operator K not degenerated in $\mathscr{U} = \bigcup_j$ range R_j the equations $(K - R_j)\, D_j = 0$ restrict the rank of decision operators D_j:*

$$\text{rank } D_j \leqq \text{rank } R_j. \tag{2.7}$$

Proof is obtained by using the well known [13] inequality rank $(AB) \leqq$ rank A for ranks of products of any two matrixes:

$$\text{rank } (R_j D_j) \leqq \text{rank } R_j, \quad \text{rank } (K D_j) = \text{rank } D_j$$

(the sign of equality here takes place since the matrix K is not degenerate). The inequalities (2.7) are valid as $R_j D_j = K D_j$.

Let us remark that every lth hypothesis with $R_l \leqq R_k$ for some k can be disregarded as for any decision $\{D_j\}$ there is a decision $\{D'_j\}$ with $D'_l = 0$, which is not worse. Such decision is $D'_j = D_j, j \neq l, k$, $D'_l = 0$, $D'_k = D_l + D_k$, the difference of corresponding risks is non-positive: $C' - C = \text{Tr}\,(R_l - R_k)\, D_l \leqq 0$. In particular, if $R_j \leqq R_k$ for all j, the optimal decision is trivial: $D_j^0 = 0, j \neq k$, $D_k^0 = I$.

Further we shall suppose with no loss of generality that the density operators $\{\rho_k\}$ multiplied by the prior probabilities $\{\pi_k\}$ can be represented as

$$\pi_k \rho_k = R_0 + R_k, \tag{2.8}$$

where R_0 is some operator in $\mathscr{H}$, and $C_{jk} = 1 - \delta_{jk}$. The risk (1.2) gives then the probability of error P_e, the lower bound of which is defined by the trace of quantum supremum $K = \sup\{R_j\}$ in subspace $\mathscr{U} = \bigcup$ range R_j satisfying the condition (2.6):

$$P_e^0 = \min P_e = 1 - \text{Tr}_{\mathscr{H}} R_0 - \text{Tr}_{\mathscr{U}} K. \tag{2.9}$$

We shall call this trace $\varkappa = \min_{\{K \geqq R_j\}} \text{Tr}\, K$ the quality of the optimal decision in sufficient decision space $\mathscr{U}$.

3. OPTIMAL MULTIPLE HYPOTHESIS TESTING IN TWO-DIMENSIONAL DECISION SPACE

We shall begin with the simplest case when the dimension of the complex vector space $\mathscr{U} = \bigcup$ range R_j equals 2.† We shall identify then the operators $\{R_j\}$ in the optimization problem (2.6) with Hermitian matrices which can be considered without loss of generality to be non-negative. Any 2×2-matrix R can be decomposed by the Pauli matrices

$$\imath = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix};$$

$$R = \begin{pmatrix} \tau + z & x - iy \\ x + iy & \tau - z \end{pmatrix} = \tau\imath + x\sigma_x + y\sigma_y + z\sigma_z \equiv \tau + \hat{r}, \tag{3.1}$$

where τ, x, y, z are real when R is Hermitian, and $\hat{r} \equiv x\sigma_x + y\sigma_y + z\sigma_z$ is an operator vector, represented by the vector $\mathbf{r} = (x, y, z)$ of real three-dimensional space $\mathscr{R}$. The product of $\hat{r}, \hat{s}$, in terms of $\mathbf{r} \in \mathscr{R}$, $\mathbf{s} \in \mathscr{R}$ is $\hat{r}\hat{s} = (\mathbf{r}, \mathbf{s}) + i[\mathbf{r} \widehat{,} \mathbf{s}]$, where $(\mathbf{r}, \mathbf{s})$, $[\mathbf{r}, \mathbf{s}]$ are the scalar and vector products of $\mathbf{r}, \mathbf{s}$. It is well known that Tr $R = 2\tau$, Det $R = \tau^2 - |\mathbf{r}|^2$, $(|\mathbf{r}| \equiv \sqrt{(\mathbf{r}, \mathbf{r})})$, and condition of non-negativity $R \geqq 0$ has the form $\tau \geqq |\mathbf{r}|$. Obviously, rank $R = 2$ when $\tau > |\mathbf{r}|$, rank $R = 1$ when $\tau = |\mathbf{r}| \neq 0$, and rank $R = 0$ when $\tau = 0$.

The operators $R_k = (\pi_k + \hat{p}_k)/2$ with $\pi_k \geqq |\mathbf{p}_k|$, $\sum_{k=1}^{M} \pi_k = 1$ can be interpreted as density operators associated with quantum statistical hypotheses to be tested, normalized to prior probabilities: $\pi_k = \text{Tr}\, R_k$, and represented by vectors $\mathbf{p}_k \in \mathscr{R}$ called polarization vectors. For instance, such a problem arises when one wants to discriminate a photon polarization or the spin of an electron. We shall suppose that the polarizations $\{\mathbf{p}_j\}$ satisfy inequalities

$$|\mathbf{p}_k - \mathbf{p}_l| > |\pi_k - \pi_l| \tag{3.2}$$

for all $k \neq l$, in contrary case $|\pi_k - \pi_l| \geqq |\mathbf{p}_k - \mathbf{p}_l|$ there is a predominance among kth and lth hypotheses: either $R_k > R_l$ or $R_k = R_l$ or $R_k < R_l$, and one of these hypotheses can be disregarded.

The decision operators D_j in Pauli representation $D_j = \delta_j + \hat{d}_j$ are described by non-negative numbers $\delta_j \geqq 0$ and by vectors $\mathbf{d}_j \in \mathscr{R}$ $(|\mathbf{d}_j| \leqq \delta_j)$, and a resolution of the identity $\Sigma D_j = \imath$ takes the form

$$\sum_{j=1}^{M} \delta_j = 1, \qquad \sum_{j=1}^{M} \mathbf{d}_j = 0. \tag{3.3}$$

Solving the problem of optimal discrimination of polarizations $\mathbf{p}_k$ with probabilities π_k reduces to finding a real number $\varkappa$ and a vector $\mathbf{q} \in \mathscr{R}$ defining

†Note that according to the Pauli principle such a decision space proves to be sufficient for detection of particles with half-integral spin in some fixed state.

the operator $K = (\varkappa + \hat{q})/2$ satisfying conditions (2.6) for some set $D_j^0 = \delta_j^0 + \hat{d}_j^0, j = \overline{1, M}$.

THEOREM 3 *It is necessary and sufficient for solving the problem of optimal discrimination of the polarizations,* $\{\mathbf{p}_k\}$, *satisfying with probabilities* $\{\pi_k\}$ *the conditions* (3.2) *to find a set of non-negative numbers* $\mu_k \geqq 0$, $k = \overline{1, M}$, *satisfying the inequalities*

$$\left|\sum_{k=1}^{M} (\mathbf{p}_j - \mathbf{p}_k)\mu_k\right| + \sum_{k=1}^{M} (\pi_j - \pi_k)\,\mu_k \geqq 1, \quad j = \overline{1, M}, \tag{3.4}$$

where the sign of equality holds at least for those j, where $\mu_j \neq 0$. *The optimal decision operators are represented by* $\delta_j^0 = |\mathbf{d}_j^0|$,

$$\mathbf{d}_j^0 = \mu_j(\mathbf{p}_j - \mathbf{q}), \text{ where } \mathbf{q} = \sum_{k=1}^{M} \mu_k \mathbf{p}_k / \sum_{k=1}^{M} \mu_k, \tag{3.5}$$

and the minimum probability of error equals

$$P_e^0 = 1 - \varkappa, \text{ where } \varkappa = (1 + \sum_{k=1}^{M} \mu_k \pi_k) / \sum_{k=1}^{M} \mu_k. \tag{3.6}$$

Proof The first condition of optimality (2.6) $K - R_j \geqq 0$ in terms of $\varkappa$, $\mathbf{q}$, π_j, $\mathbf{p}_j$ has the form

$$\varkappa \geqq |\mathbf{p}_j - \mathbf{q}| + \pi_j. \tag{3.7}$$

The second equations $(K - R_j)D_j = 0$ can be written in the form of equations for scalar and real vector part of product $(\varkappa - \pi_j + \hat{q} - \hat{p}_j)(\delta_j + \hat{d}_j)$:

$$(\varkappa - \pi_j)\,\mathbf{d}_j + (\mathbf{q} - \mathbf{p}_j)\,\delta_j = 0, \quad (\varkappa - \pi_j)\,\delta_j + (\mathbf{q} - \mathbf{p}_j, \mathbf{d}_j) = 0 \tag{3.8}$$

(the imaginary vector equation $i[\mathbf{q} - \mathbf{p}_j, \mathbf{d}_j] = 0$ is the consequent of real vector Eq. (3.8) from which either $\mathbf{d}_j$ is collinear to $\mathbf{p}_j - \mathbf{q}$ for any $\delta_j \geqq |\mathbf{d}_j|$, or $\mathbf{d}_j = 0$). As soon as $\varkappa - \pi_j \neq 0$, the Eqs. (3.8) are equivalent to the equations

$$\mathbf{d}_j = \delta_j(\mathbf{p}_j - \mathbf{q})/(\varkappa - \pi_j), \quad ((\varkappa - \pi_j)^2 - |\mathbf{p}_j - \mathbf{q}|^2)\,\delta_j = 0 \tag{3.9}$$

in the contrary case $\varkappa = \pi_k$ for some k from (3.7) with $j = k$ we obtain $\mathbf{q} = \mathbf{p}_k$, and inequalities (3.7), (3.2) are incompatible. Optimal decision vectors $\mathbf{d}_j^0$ can be written as (3.5), where $\mu_j = \delta_j/(\varkappa - \pi_j)$ is non-negative according to $\delta_j \geqq 0$, $\varkappa > \pi_j$ (3.7), and $\mathbf{q}$ is defined by the set $\{\mu_j\}$ in accordance with $\Sigma \mathbf{d}_j = 0$. It follows from the second Eq. (3.9) that the inequalities (3.7) are the equalities for those j, where $\delta_j = \mu_j(\varkappa - \pi_j) \neq 0$. Multiplying (3.7) by the sum $\sum_{k=1}^{M} \mu_k$ and defining $\varkappa$ from the condition $\Sigma \delta_j = 1$ for $\delta_j^0 = \mu_j(\varkappa - \pi_j) = \mu_j|\mathbf{p}_j - \mathbf{q}| = |\mathbf{d}_j^0|$: $\varkappa = (1 + \Sigma \mu_k \pi_k)/\Sigma \mu_k$, we obtain the conditions (3.4) for defining $\{\mu_j\}$. Since Tr $K = \varkappa$ the minimum probability

of error (2.8) equals (3.6) in the case $R_0 = 0$ to be considered. The theorem is proved.

Let us remark that the equalities (3.4) must be fulfilled at least for two indexes $j = k, l$, since there is no set $\{\mu_j\}$ with $\mu_j \neq 0$ only for one index j, which satisfies the jth equation. For any pair $\mathbf{p}_k$, $\mathbf{p}_l$ with (3.2) there is always unique solution $\mu_j = 0$, for all $j \neq k, l$, $\mu_k > 0$, $\mu_l > 0$ of pair of kth and lth equations (3.4):

$$\mu_k = (|\mathbf{p}_l - \mathbf{p}_k| + \pi_l - \pi_k)^{-1}, \quad \mu_l = (|\mathbf{p}_k - \mathbf{p}_l| + \pi_k - \pi_l)^{-1},$$

but such set $\{\mu_j\}$ can not satisfy other inequalities (3.4) with $j \neq k, l$. If there is such a pair $\mathbf{p}_k$, $\mathbf{p}_l$ for which all inequalities (3.4) are valid with $\mu_j \neq 0$ only for $j = k, l$, the optimal decision vectors (3.5) are

$$\mathbf{d}_j^0 = 0, j \neq k, l, \ \mathbf{d}_k^0 = (\mathbf{p}_k - \mathbf{p}_l)/2|\mathbf{p}_k - \mathbf{p}_l|, \ d_l^0 = (\mathbf{p}_l - \mathbf{p}_k)/2|\mathbf{p}_l - \mathbf{p}_k|.$$

The optimal decision operators $D_j^0 = |\mathbf{d}_j^0| + \hat{d}_j^0$ are orthogonal, and corresponds with the probability of error

$$P_e^0 = 1 - (\pi_k + \pi_l)/2 - |\mathbf{p}_k - \mathbf{p}_l|/2.$$

In general case the optimal operators D_j^0 are non zero for more than two numbers j, and define a non-orthogonal resolution of the identity in two dimensional space $\mathscr{U}$. We shall not try to find any general analytical solution of the system of the Eqs. (3.5) with $\mu_j \neq 0$ for more than two numbers j, but we shall give its geometrical interpretation.

Let us consider the four-dimensional space points of which $r = (\tau, x, y, z) = (\tau, \mathbf{r})$ represent Hermitian operators (3.1). Any non-negative operator is associated with a point inside of the "light" cone $\tau = |\mathbf{r}|$. In these terms there is no prior predominance among kth and lth hypotheses with $R_k = (\pi_k + \hat{p}_k)/2$, $R_l = (\pi_l + \hat{p}_l)/2$ if and only if the interval $p_k - p_l = (\pi_k - \pi_l, \mathbf{p}_k - \mathbf{p}_l)$ is a "space-like" interval (3.2). In accordance with (3.7) the point $q = (\varkappa, \mathbf{q})$ representing the operator $K = \sup\{R_j\}$ is the top of a four-dimensional right-angled cone

$$\mathscr{C}(q) = \{r = (\tau, \mathbf{r}) : \tau - \varkappa + |\mathbf{r} - \mathbf{q}| = 0\} \tag{3.10}$$

covering all points $p_j = (\pi_j, \mathbf{p}_j)$ and containing a subset $\{p_{j_\alpha}\} \subset \{p_j\}$ of boundary points p_{j_α}, satisfying the equalities (3.7) (see Figure 1). On the other hand there are admitted only the points q with the space projections $\mathbf{q}$ belonging to the convex covering of the boundary subset space projections $\mathbf{p}_{j_\alpha}$, $\alpha = \overline{0, s}$ $(s \leqq M-1)$:

$$\sum_{\alpha=0}^{s} \mathbf{p}_{j_\alpha} \lambda_{j_\alpha} = \mathbf{q}, \quad \sum_{\alpha=0}^{s} \lambda_{j_\alpha} = 1, \tag{3.11}$$

where according to (3.5) $\lambda_{j_\alpha} = \mu_{j_\alpha} / \sum_{\alpha=0}^{s} \mu_{j_\alpha} \geqq 0$ ($\mu_j = 0$ if p_j is covered by the cone (3.10): $\pi_j + |\mathbf{p}_j - \mathbf{q}| < \varkappa$). We shall say that the subset $\{p_{j_\alpha}\}$ has the

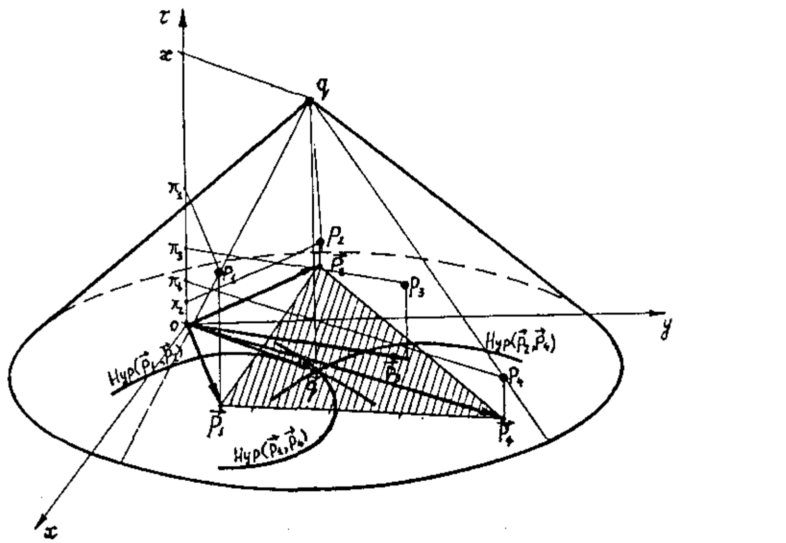

FIGURE 1 Optimal discrimination of four polarizations $\mathbf{p}_k$, $k = 1, 2, 3, 4$, being in the plane (x, y) with prior probabilities π_k. The point $q = (\varkappa, \mathbf{q})$ in the space (τ, x, y) is the top of cone (3.10) covering all points $p_k = (\pi_k, \mathbf{p}_k)$ and osculating with three points p_1, p_2, p_4 (the fourth point p_3 is lying inside of the cone). The convex covering of osculating points (triangle $(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_4)$) must contain the space projection $\mathbf{q}$. The vector $\mathbf{q}$ called projection the centre of $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_4$ is a common point of hyperbols (3.15) marked as $\mathrm{Hyp}(\mathbf{p}_k, \mathbf{p}_l)$ with $k, l = 1, 2, 4$.

top q, if the points p_{j_α} are situated on a cone: $p_{j_\alpha} \in \mathscr{C}(q)$ with the top q, space projection $\mathbf{q}$ of which is a point of the convex covering of $\{\mathbf{p}_{j_\alpha}\}$. In these terms the theorem 3 can be reformulated as follows.

THEOREM 3 *It is necessary and sufficient for solving the problem of optimal discrimination of points* $p_k = (\pi_k, \mathbf{p}_k)$, $k = \overline{1, M}$ *separated by space-like intervals* (3.2) *to find a subset* $\{p_{j_\alpha}\}$ *with the top* q, *the cone of which covers the other points of the set* $\{p_k\}$, *that is to find a subset of vectors* $\mathbf{p}_{j_\alpha} \in \{\mathbf{p}_k\}$; $\alpha = \overline{0, s}$ *whose convex covering contains a vector* $\mathbf{q}$ *with respect of which the sum* $|\mathbf{p}_{j_\alpha} - \mathbf{q}| + \pi_{j_\alpha}$ *is a constant* $\varkappa$:

$$|\mathbf{p}_{j_\alpha} - \mathbf{q}| + \pi_{j_\alpha} = \varkappa, \quad \alpha = \overline{0, s} \tag{3.12}$$

and $|\mathbf{p}_j - \mathbf{q}| + \pi_j \leqq \varkappa$ *for all other indexes* $j \,\bar{\in}\, \{j_\alpha\}$. *Optimal decision operators are represented by the cone points* $d_j^0 = (\delta_j^0, \mathbf{d}_j^0)$: $\delta_j^0 = |\mathbf{d}_j^0|$, *with the space vectors*

$$\mathbf{d}_j^0 = \lambda_j(\mathbf{p}_j - \mathbf{q}) \Big/ \sum_{j=1}^{M} \lambda_j |\mathbf{p}_j - \mathbf{q}|, \tag{3.13}$$

where $\lambda_j = 0$ *if* $j \,\bar{\in}\, \{j_\alpha\}$, *and* $\{\lambda_{j_\alpha}, \alpha = \overline{0, s}\}$ *is any non-negative solution of the system of Eqs.* (3.11). *The minimum probability of error is*

$$P_e^0 = 1 - \sum_{j=1}^{M} (\pi_j + |\mathbf{p}_j - \mathbf{q}|)\lambda_j. \tag{3.14}$$

Let us remark that any pair of space separated points p_k, p_l defines by two Eqs. (3.12) $j_\alpha = k, l$ with respect to $\mathbf{q}$ the set of points $\mathbf{q} \in \mathscr{R}$, the difference of distances from which to $\mathbf{p}_k, \mathbf{p}_l$ is constant:

$$|\mathbf{q} - \mathbf{p}_k| - |\mathbf{q} - \mathbf{p}_l| = \pi_l - \pi_k. \tag{3.15}$$

Such points are placed in one of the two branches of the hyperboloid of revolution with the focii $\mathbf{p}_k$ and $\mathbf{p}_l$, and with the excentricity, $\varepsilon = |\mathbf{p}_k - \mathbf{p}_l| / |\pi_k - \pi_l| > 1$. When $\pi_k = \pi_l$ the hyperboloid (3.15) degenerates into a plane which is normal to the segment $\mathbf{p}_k\lambda_k + \mathbf{p}_l\lambda_l$ $(\lambda > 0, \lambda_k + \lambda_l = 1)$ in the point $(\mathbf{p}_k + \mathbf{p}_l)/2$, and when $\pi_k \neq \pi_l$ that branch is chosen, in the cavity of which the focus with the maximum of π_k, π_l lies. It is obvious that if a subset $\{p_{j_\alpha}\}$ has the top q, the space projection $\mathbf{q}$ is a common point of the hyperboloids (3.15) associated with all pairs of this set $\{p_{j_\alpha}\}$, which belongs to the convex covering of $\{\mathbf{p}_{j_\alpha}\}$ (see Figure 1). We shall call this point $\mathbf{q}$ the center of $\{\mathbf{p}_{j_\alpha}\}$. From the uniqueness of the quantum supremum proved in the previous paragraph it follows that the top of $\{p_{j_\alpha}\}$ representing $K = \sup\{R_{j_\alpha}\}$ is unique, and the center of $\{\mathbf{p}_{j_\alpha}\}$ is also unique. It is easy to prove that for any vector $\mathbf{q}$ of convex covering of $\{\mathbf{p}_{j_\alpha}\}$ the system of linear Eqs. (3.11) has a unique solution $\{\lambda_{j_\alpha}\}$, if and only if the vectors $\mathbf{p}_{j_\alpha} - \mathbf{p}_{j_0}$, $\alpha = \overline{1, s}$ are linearly independent. The convex covering of such set $\{\mathbf{p}_{j_\alpha}\}$ of $s+1$ points $\mathbf{p}_{j_\alpha}$ is called s-simplex (s = 1-segment, s = 2-triangle, s = 3-tetrahedron). It is well known that every s-dimensional side (s-side) of m-simplex $(m \geqq s)$ is also simplex. We shall call the subset generating the simplex convex covering the simplex subset.

THEOREM 4 *The problem of optimal discrimination of polarizations* $\{\mathbf{p}_k, k = \overline{1, M}\}$ *has always the solution described by the simplex set* $\{\mathbf{d}_{j_\alpha}^0, \alpha = \overline{0, s}\}$ $(s+1 \leqq M)$ *of non-zero vectors* (3.13), *which corresponds to the simplex subset* $\{\mathbf{p}_{j_\alpha}\} \subseteq \{\mathbf{p}_k\}$ *having center* $\mathbf{q}$ *with maximum sum*

$$\pi_{j_\alpha} + |\mathbf{p}_{j_\alpha} - \mathbf{q}| = \max.$$

This solution is unique if and only if the s-simplex generated by the subset $\{\mathbf{p}_{j_\alpha}\}$ *is an s-side of the convex covering of all vectors* $\mathbf{p}_{j_0}, \ldots, \mathbf{p}_{j_{m-1}}$ *having the common center* $\mathbf{q}$.

Proof According to the theorem 3 the solving of the problem under consideration reduces to finding a cone (3.10) covering all points $\{p_k\}$ and having the projection of top $\mathbf{q}$ inside of convex covering of projections $\mathbf{p}_{j_\alpha}$ of osculating points p_{j_α}. It is obvious that such cone exists always. Let $m \leqq M$ be the full number of osculating points p_{j_α}: $\alpha = \overline{0, m-1}$. If the subset $\mathbf{p}_{j_\alpha}$, $\alpha = \overline{0, s}$ $(s = m-1)$ is a simplex set, Theorem 4 is obvious. In non simplex case the convex covering of $\{\mathbf{p}_{j_\alpha}\}$ can be divided into a number of simplexes with common top $\mathbf{p}_j$ by the diagonal compunctal planes $(\mathbf{p}_{j_0}, \mathbf{p}_{j_\alpha}, \mathbf{p}_{j_\beta})$, or diagonal compunctal lines $(\mathbf{p}_{j_0}, \mathbf{p}_{j_\alpha})$, when all vectors $\mathbf{p}_{j_\alpha}$ are complanar. Hence the center $\mathbf{q}$ is an inner point of one of s-simplex $(s+1 < m \leqq M)$ with the tops $\mathbf{p}_{j_\alpha}$, $\alpha = \overline{0, s}$ of osculating point projections, defining the unique positive solution $\{\lambda_{j_\alpha}\}$ of the system (3.11). The set $\{\mathbf{d}^0_{j\alpha}\}$ of non-zero vectors (3.13) is a simplex set if and only if the subset $\{\mathbf{p}_{j_\alpha}\}$ is simplex, and defines an optimal decision with maximum quality (3.12), or minimum probability of error (3.14). When the center $\mathbf{q}$ is an inner point of non-simplex convex covering of osculating point projections, the simplex optimal decision is not unique (the simplex division is not unique), and there is non simplex optimal decisions. But if the point $\mathbf{q}$ is a boundary point of convex covering, i.e. an inner point of s-side, the optimal decision can be unique, if the side is s-simplex.

CONSEQUENCE *It is always sufficient for solving the problem of optimal multiple hypothesis testing with two-dimensional space $\mathscr{U}$ to confine oneself by consideration of $s+1 \leqq 4$ decisions $j_0, \ldots, j_s$ associated with simplex subsets of hypothesis $\mathbf{p}_{j_0}, \ldots, \mathbf{p}_{j_s}$. Every such decision procedure can be always realized as a quantum-mechanical measurement described by orthogonal resolution of observation space $\mathscr{H} = \mathscr{U} \otimes \mathscr{U}$.*

Indeed, there is no simplex subset $\mathbf{p}_{j_0}, \ldots, \mathbf{p}_{j_s}$ in three-dimensional space $\mathscr{R}$ with $s > 3$, and for any M an optimal resolution of identity in two-dimensional decision space $\mathscr{U}$ exists always by $s+1 \leqq 4$ non-zero decision operators $D^0_{j_\alpha} = \delta^0_{j_\alpha} + \hat{d}^0_{j_\alpha}$ having the unite rank. It is well known [10] that every non-orthogonal resolution of identity by operators $D_{j_0}, \ldots, D_{j_s}$ with unite rank can be imbeded in an orthogonal one in $s+1$-dimensional space $\mathscr{H} \supset \mathscr{U}$. Hence it is possible to confine with four-dimensional observation space $\mathscr{H}$ which can always be represented as tensor square of two-dimensional $\mathscr{U}$: $\mathscr{H} = \mathscr{U} \otimes \mathscr{U}$ corresponding with a composition of two "spin" systems.

Let us note that the optimal decision can be also degenerate (in the sense that some hypotheses p_j correspond with $D_j = 0$) when set $\mathbf{p}_1, \ldots, \mathbf{p}_M$ is a simplex set ($M \leqq 4$), for instance, $M = 3$, $\pi_1 = \pi_2 = \pi_3$, and the vectors $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ form an obtuse triangle. The analogous case also takes place in the problem of estimation of parameters of quantum Gaussian signal [14].

Let us consider two particular cases.

1. *Optimal discrimination of pure polarizations.* Here the polarizations are

normalized to prior probabilities: $|\mathbf{p}_k| = \pi_k$. The representing points $p_k = (\pi_k, \mathbf{p}_k)$ are the points of cone $\tau = |\mathbf{r}|$. The Eq. (3.12) defining the subset of points p_{j_α} osculating with the covering cone (3.10) have the form

$$|\mathbf{p}_{j_\alpha} - \mathbf{q}| + |\mathbf{p}_{j_\alpha}| = \varkappa. \tag{3.16}$$

With respect to $\mathbf{p}_{j_\alpha}$ the Eq. (3.16) is the equation of an ellipsoid of revolution with the focuses $0, \mathbf{q}$ and excentricity $\varepsilon = |\mathbf{q}|/\varkappa < 1$. According to (3.7) all other points $\mathbf{p}_j \bar{\in} \{\mathbf{p}_{j_\alpha}\}$ are to lie inside of this ellipsoid. Hence the problem of optimal discrimination of pure polarizations reduces to finding an ellipsoid circumscribed over the points $\{\mathbf{p}_k\}$ and with the focuses $0, \mathbf{q}$, where $\mathbf{q}$ is an inner point of convex covering of osculating points $\{\mathbf{p}_{j_\alpha}\}$. The quality $\varkappa$ of the optimal decision is the length of big axis of this ellipsoid.

2. *Optimal discrimination of equiprobable polarizations.* The prior probabilities are $\pi_k = 1/M$, and the representing points are the points of hyper-plane $\tau = 1/M$. The density operators $\rho_k = (1+\mathbf{r}_k)/2$ with the unit trace are represented by renormalized vectors $\mathbf{r}_k = M\mathbf{p}_k$. The crossing of covering cone (3.10) with the hyper-plane $\tau = 1/M$ is the sphere $|\mathbf{p}-\mathbf{q}| = \varkappa - 1/M$, or in terms of renormalized vectors the sphere $|\mathbf{r}-\mathbf{s}| = R$ $(\mathbf{s} = M\mathbf{q}$, $R = M\varkappa - 1)$. Hence the problem of optimal discrimination of equiprobable polarizations reduces to finding a sphere circumscribed over all points $\mathbf{r}_k$: $|\mathbf{r}_k - \mathbf{s}| \leqq R$ with a radius R and center $\mathbf{s}$ belonging to the convex covering of osculating points $\mathbf{r}_{j_\alpha}$: $|\mathbf{r}_{j_\alpha} - \mathbf{s}| = R$. The radius $R = M\varkappa - 1$ defines the minimal probability of error (3.6):

$$P^0_e = 1 - (1+R)/M. \tag{3.17}$$

($R \leqq 1$ as soon as $|\mathbf{r}_k| \leqq 1$ for all k). The minimum of the probability (3.17) is obtained when $R = 1$: $P^0_e = 1 - 2/M$. It corresponds to the pure equiprobable case $|\mathbf{r}_k| = 1$, when there is at least one simplex subset $\{\mathbf{r}_{j_\alpha}\}$ for which the center $\mathbf{s} = 0$ is an inner point of the simplex.

4. OPTIMAL DISCRIMINATION OF QUANTUM PURE STATES

We shall use the Dirac notations for investigation of optimal multiple quantum pure states discrimination problem in the general case. Let $|\psi_k\rangle$, $k = \overline{1, M}$ be a set of vectors of complex Hilbert space $\mathscr{H}$ normalised to prior probabilities: $\langle\psi_k|\psi_k\rangle = \pi_k$ with given matrix of scalar products

$$Q_{kl} = \langle\psi_k|\psi_l\rangle, \tag{4.1}$$

($\mathrm{Tr}\, Q = 1$.) It is necessary to test the multiple hypothesis associated with pure quantum states $\rho_k = |\psi_k\rangle\langle\psi_k|/\pi_k$ minimising the average probability of error. Let $r \leqq M$ be the rank of matrix $Q = \|Q_{kl}\|$. As soon as the operators $R_k = |\psi_k\rangle\langle\psi_k|$ have the sum of ranges $\mathscr{U} = \bigcup_k$ range R_k consisting only of vectors $|\varphi\rangle = \sum_{k=1}^{M} \lambda_k |\psi_k\rangle$ with dimension r, it is possible to confine oneself to a r-dimensional sufficient subspace $\mathscr{U} \subset \mathscr{H}$. In order to use the result of Theorem 2 let us prove the following lemma.

LEMMA *Every Hermitian operator satisfying inequalities*

$$K - |\psi_j\rangle\langle\psi_j| \geqq 0, \qquad j = \overline{1, M} \tag{4.2}$$

is strictly positive in the subspace $\mathscr{U}$ generated by $\{|\psi_j\rangle\}$.

Proof Let us assume the opposite: The space $\mathscr{U}$ contains such a vector $|\chi\rangle$, for which $\langle\chi|K|\chi\rangle = 0$. Then from (4.2) we obtain $-|\langle\chi|\psi_j\rangle|^2 \geqq 0$, i.e. $\langle\chi|\psi_j\rangle = 0, j = \overline{1, M}$. This means that $|\chi\rangle$ is orthogonal to any vector $|\varphi\rangle \in \mathscr{U}$: $|\varphi\rangle = \sum_{j=1}^{M} \lambda_j |\psi_j\rangle$. This is possible only for zero vector of $\mathscr{U}$: $|\chi\rangle = 0$.

Thus the condition of non-negativity of K in the space $\mathscr{U}$ is carried out, and in accordance with Theorem 2, the rank of optimal decision operators D_j^0 is not more than 1.

THEOREM 5 *It is necessary and sufficient for solving the problem of optimal discrimination of pure quantum states $|\psi_k\rangle \in \mathscr{H}$, $k = \overline{1, M}$ to find a set of non-negative numbers μ_k, $k = \overline{1, M}$ satisfying the inequalities*

$$\langle\psi_j|\Big(\sum_{k=1}^{M} \mu_k |\psi_k\rangle\langle\psi_k|\Big)^{-1/2}|\psi_j\rangle \geqq 1 \tag{4.3}$$

where the sign of equality takes place at least for those $j_\alpha \in \{\overline{1, M}\}$, $\alpha = \overline{1, m}$ ($r \leqq m \leqq M$) where $\mu_{j_\alpha} \neq 0$:

$$\langle\psi_{j_\alpha}|\Big(\sum_{k=1}^{M} \mu_k |\psi_k\rangle\langle\psi_k|\Big)^{-1/2}|\psi_{j_\alpha}\rangle = 1, \qquad \alpha = \overline{1, m}. \tag{4.4}$$

The optimal decision operators D_j^0 have the form

$$D_j^0 = \mu_j K^{-1}|\psi_j\rangle\langle\psi_j|K^{-1}, \quad K = \Big(\sum_{k=1}^{M} \mu_k |\psi_k\rangle\langle\psi_k|\Big)^{1/2}, \tag{4.5}$$

and the minimum probability of error is

$$P_e^0 = 1 - \sum_{k=1}^{M} \mu_k. \tag{4.6}$$

... D^0 satisfying the conditions of optimality (2.6) either

are equal to zero or look like $D_j^0 = |\chi_j\rangle\langle\chi_j|$ in accordance with the proved inequality rank $D_j^0 \leqq 1$. Here the vectors $|\chi_j\rangle$ must satisfy the equations

$$(K - |\psi_j\rangle\langle\psi_j|)|\chi_j\rangle = 0.$$

Multiplying this equation by K^{-1}, we obtain

$$|\chi_j\rangle = \langle\psi_j|\chi_j\rangle K^{-1}|\psi_j\rangle. \tag{4.7}$$

The operator K is defined by the condition of completeness $\sum_{j=1}^{M} |\chi_j\rangle\langle\chi_j| = 1$ as follows:

$$K^2 = \sum_{j=1}^{M} \mu_j |\psi_j\rangle\langle\psi_j| \tag{4.8}$$

where $\mu_j = |\langle\psi_j|\chi_j\rangle|^2$. Hence we obtain (4.5). The coefficients $\mu_j \geqq 0$ define the minimum probability of error $P_e = 1 - \sum_{j=1}^{M} |\langle\psi_j|\chi_j\rangle|^2$ by the simple formula (4.6). Multiplying (4.7) by $\langle\psi_j|$, we obtain the following equations for defining the numbers μ_j:

$$(1 - \langle\psi_j|K^{-1}|\psi_j\rangle)\mu_j^{1/2} = 0.$$

Hence either $\mu_j = 0$, or the Eqs. (4.4) are satisfied. As soon as the operator inequality $|\varphi\rangle\langle\varphi| \leqq 1$ is valid if and only if $\langle\varphi|\varphi\rangle \leqq 1$, the other condition of optimality (4.2) can be rewritten with respect to $|\varphi\rangle = K^{-1/2}|\psi_j\rangle$ as the scalar inequalities (4.3): $\langle\psi_j|K^{-1}|\psi_j\rangle \leqq 1$. Hence the necessary and sufficient conditions for optimality (4.2), (4.7) are equivalent to the conditions (4.3) (4.4) defining the numbers $\{\mu_k\}$. The theorem is proved.

The system of Eqs. (4.4) is strongly non-linear, and it is difficult to find some analytical methods of solving it in general cases. When rank $Q = 2$ we have "pure" case of the general two-dimensional case which is completely described in previous paragraph. In general case if the system (4.4) with $m = M$ has only one non-zero solution $\mu_j > 0$, $j = \overline{1, M}$, then it is optimal. But such solution exists always if and only if $r = M$. If the system with $m = M > r$ has no solution, the number m of the equations must be diminished by putting $\mu_j = 0$ for some j. An example of such situation with $r = 2$ is given in previous paragraph by any set of polarizations $\mathbf{p}_k$ normalised to the prior probabilities, which has not a center. Looking over every possible case of equalities $\mu_j = 0$ we have to find the solution with minimum of probability of error (4.6), that is with $\sum_{\alpha=1}^{M} \mu_{j_\alpha} = \max$. It is obvious that the number m cannot be less than the dimension r of the space $\mathscr{U}$, otherwise the set $\{|\chi_{j_\alpha}\rangle, \alpha = \overline{1, m}\}$ could not be complete. With $m = r$ this set is orthogonal, and with $m > r$ is not orthogonal (any overcomplete set is not orthogonal). It was proved in previous paragraph that when $r = 2$ we always have an optimal set of decision operators (or vectors $|\chi_{j_\alpha}\rangle$) which consists of not more than four non-zero elements: $m \leqq 2^2$. It may be supposed that it is enough to consider r^2 non-zero numbers μ_{j_α}.

Let us note that if the set $|\psi_k\rangle$ disintegrates into a number of mutually orthogonal subsets $\{|\psi_{k\alpha}\rangle, \alpha = \overline{1, M_k}\}, \sum M_k = M: \langle\psi_{k\alpha}|\psi_{l\beta}\rangle = 0$ when $k \neq l$, the matrix (4.7) acquires quasi-diagonal form $K = \|\delta_{kl}K_l\|$ and the search of the numbers $\{\mu_{k\alpha}\}$ is reduced to solution of independent systems of equations $\langle\psi_{l\alpha}|K^{-1}|\psi_{l\alpha}\rangle = 1$ (if $\mu_{l\alpha} \neq 0$) of dimension $m_l \leqq M_l$. For completely orthogonal sets $M_k = 1$, we obtain $\mu_k = \pi_k$. The minimum probability of error is equal to zero if and only if the set $\{|\psi_k\rangle\}$ is orthogonal, and is not equal to zero even if the non-orthogonal set $\{|\psi_k\rangle\}$ admits an orthogonalization (i.e. if rank $Q = M$). In this fact the fundamental difference of the problem of discrimination of quantum states from the corresponding classical problem, is shown. The last problem always leads in the case of pure states (i.e. non-fluctuating signals) to the singular solution $P_e^0 = 0$.

Linearly independent non-orthogonal states. Let r = rank Q be M. It is convenient to put the results in this case into matrix form in representation of any basis set $\{\langle\varphi_k|, k = \overline{1, M}\}$, received from $\{|\psi_k\rangle\}$ by using some procedure of orthogonalization. Introducing matrix denotations

$$\Psi = \|\langle\varphi_k|\psi_l\rangle\|, \quad \mathrm{X} = \|\langle\varphi_k|\chi_l\rangle\|, \quad \mu = \|\mu_k\delta_{kl}\|$$

and taking into account that in this representation $K^2 = \Psi\mu\Psi^+$, we get from (4.7), (4.8)

$$\mathrm{X} = \Psi(\mu Q)^{-1/2}\mu^{1/2}, \text{ where } Q = \Psi^+\Psi. \tag{4.9}$$

Using the matrix identity $f(\Psi\mu\Psi^+)\Psi = \Psi f(\mu\Psi^+\Psi)$, which is true for any sufficiently general function f, and matrix Ψ it is easy to check that the optimal decision matrix (4.9) really defines a resolution of the matrix identity $I = \mathrm{XX}^+$ and moreover it is unitary: $\mathrm{X}^+\mathrm{X} \equiv \|\langle\chi_k|\chi_l\rangle\| = \|\delta_{kl}\|$. Hence it is possible to introduce an observable in M-dimensional space $\mathscr{U}$:

$$N = \mathrm{X}\nu\mathrm{X}^+ = \Psi(\mu Q)^{-1/2}\nu\mu(Q\mu)^{-1/2}\Psi^+ \tag{4.10}$$

($\nu = \|k\delta_{kl}\|$), measurement of which gives the number of the unknown hypothesis k with the minimum probability of error $P_e^0 = 1 - \mathrm{Tr}\,\mu$. Using the matrix function diag $A = \|A_{kl}\delta_{kl}\|$ which diagonalizes the matrix A and applying it to $\Psi^+K^{-1}\Psi = Q(\mu Q)^{-1/2}$, we can put down the system of Eqs. (4.4) as a single matrix equation

$$\mathrm{diag}\,(Q(\mu Q)^{-1/2}) = I \tag{4.11}$$

with respect to diagonal matrix μ. It is easy to prove that this equation always has a unique solution, i.e. the optimal discrimination of the linearly independent vectors $|\psi_k\rangle$ is unique.

Coherent non-orthogonal signals. Non-orthogonal sets of coherent states $|\alpha_k\rangle$ are most interesting from the point of view of applications. The sets $\{|\alpha_k\rangle\}$ are defined by the sets $\{\alpha_k\}$ of vectors α_k of some complex unitary space $\mathscr{L}$ with the scalar product $(\alpha_k, \alpha_l) \equiv \alpha_k^+\alpha_l$. The scalar product $\langle\alpha_k|\alpha_l\rangle$ in the Hilbert space $\mathscr{H}$ is defined by the scalar product $\alpha_k^+\alpha_l$ in the signal space $\mathscr{L}$ according with the following well known formula

$$\langle\alpha_k|\alpha_l\rangle = \exp\{-\alpha_k^+\alpha_k/2 + \alpha_k^+\alpha_l - \alpha_l^+\alpha_l/2\}. \tag{4.12}$$

Since $\langle\alpha_k|\alpha_k\rangle = 1$, the coherent states normalised to the prior probabilities are

$$|\psi_k\rangle = \pi_k^{1/2}|\alpha_k\rangle \tag{4.13}$$

We shall suppose that the signals α_k are represented by complex-valued functions $\alpha_k(t)$, $t \in [0, T]$ in the space $\mathscr{L}^2(0, T)$ so that $\alpha_k^+\alpha_l = \int_0^T \alpha_k^*(t)\cdot \alpha_l(t)\,dt$. It is known [15], that any finite set of coherent states is linearly independent. Its optimal discrimination with given prior probabilities π_k is defined by (4.9)–(4.11), where $Q = \|\pi_k^{1/2}\langle\psi_k|\psi_l\rangle\pi_l^{1/2}\|$. Let us note that in accordance with (4.12) the vector states $|\alpha_k\rangle$ are non-orthogonal even if all the signals α_k are orthogonal: $\alpha_k^+\alpha_l = 0$ $(k \neq l)$.

It is not difficult to make up a computer program to search for the maximum of the sum $\sum_{\alpha=1}^m \mu_{j_\alpha}$ by one of gradient methods with the side conditions (4.4) (or, if rank $Q = M$ with the conditions (4.11)) quickly leading to the aim in every concrete case when M is not large.

5. OPTIMAL DISCRIMINATION OF PURE STATES IN SPECIAL CASE

In the classical theory a simplifying assumption which is often accepted in order to find an analytical solution of the problem of optimal detection of non-orthogonal signals is the assumption of equal prior probabilities $\pi_k = \pi_l$. In the case of quantum non-orthogonal pure states $|\psi_k\rangle$ this assumption of equality of the diagonal elements of the matrix (4.1) ($Q_{kk} = Q_{ll} = M^{-1}$) does not give in general any simplification. Another assumption of the equality of the diagonal elements of the matrix $Q^{1/2}$:

$$(Q^{1/2})_{kk} = (Q^{1/2})_{ll} \text{ for all } k, l \tag{5.1}$$

is useful. In this special equidiagonal case with any r = rank Q and $M \geqq r$ the system of Eqs. (4.4) has the solution $\mu_k = \mu_l \equiv \mu$ where μ is a constant

$\mu > 0$. Indeed as soon as the identity

$$\langle\psi_k|(\sum_{j=1}^{M}|\psi_j\rangle\langle\psi_j|)^{-1/2}|\psi_l\rangle = (Q^{1/2})_{kl}$$

(which can easily be checked by squaring) holds, the substitution $\mu_k = \mu$ satisfies the system of Eqs. (4.4) if and only if $(Q^{1/2})_{kk} = (Q^{1/2})_{ll}$. We have hen

$$\mu^{1/2} = (Q^{1/2})_{kk} = (Q^{1/2})_{ll} = M^{-1}\,\mathrm{Tr}\,Q^{1/2}. \qquad (5.2$$

The substitution $\mu_k = \mu$ in (4.7), (4.8) gives the following optimal set of decision vectors in the space $\mathscr{U}$:

$$|\chi_j\rangle = (\sum_{k=1}^{M}|\psi_k\rangle\langle\psi_k|)^{-1/2}|\psi_j\rangle, \quad j = \overline{1, M}, \qquad (5.3)$$

or $|\chi_j\rangle = \sum_{k=1}^{M}|\psi_k\rangle(Q^{-1/2})_{kj}$, if rank $(Q) = M$. The minimum probability of error $P_e^0 = 1-\mu M$ of discrimination of pure states in the equidiagonal case (5.1) can be represented in the form:

$$P_e^0 = 1-(M^{-1}\,\mathrm{Tr}\,(MQ)^{1/2})^2. \qquad (5.4)$$

Let us remark that the case (5.1) is general enough, as no additional restrictions are placed on the non-diagonal elements of the matrix $Q^{1/2}$. Let us consider the next particular case.

Homogeneous sets of the pure states. We shall call a set of equiprobable states $|\psi_k\rangle$ homogeneous if the scalar products (4.1) are invariant under translation

$$Q_{k,l} = Q_{k+j,l+j} \equiv Q_{k-l}$$

where j is any integer. Homogeneous sets which contain only a finite number M of non-coinciding states satisfy the condition of cyclicity $|\psi_k\rangle = |\psi_{k+nM}\rangle$ for any integer n. It is obvious that any finite set of states having the matrix of scalar products depending only on the difference of indices: $\langle\psi_k|\psi_l\rangle = Q_{k-l}$ and satisfying the condition $Q_j = Q_{j+M}$ can be interpreted as a homogeneous set by cyclic continuation. For the homogeneous case, the spectral Fourier representation gives:

$$Q_j = \sum_{k=1}^{M} q_k M^{-1}\,e^{2\pi ikj/M}, \qquad M < \infty$$

$$Q_j = \int_{-\pi}^{\pi} q(\varphi)(2\pi)^{-1}\,e^{i\varphi j}\,d\varphi, \qquad M = \infty. \qquad (5.5)$$

Since the matrix $Q^{1/2}$ (and every other function of matrix Q also) is homogeneous and has the spectral representation (5.5) with the spectrum $q^{1/2}$, it is equidiagonal (5.1), and the solution of the problem of optimal discrimination of homogeneous states takes the form (5.3), (5.4) For example, let us consider the next homogeneous set of non-orthogonal states which has a maximal degree of symmetry.

Equiangle and simplex states. Let $|\psi_k\rangle$ be equiprobable states $\langle\psi_k|\psi_k\rangle = M^{-1}$, with equal angles $\langle\psi_k|\psi_l\rangle M = \Delta$ $(k \neq l)$. Having represented the matrix Q in the form

$$Q = M^{-1}(1-\Delta)\left(I+\frac{\Delta}{1-\Delta}\varkappa\varkappa^+\right),$$

$$I = \begin{bmatrix} 1\cdots 0 \\ 0\cdot 1\cdot 0 \\ \vdots \quad \vdots \\ 0\cdots 1 \end{bmatrix}, \varkappa = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \varkappa^+ = (1\ldots 1) \qquad (5.6)$$

and calculating the matrix $Q^{1/2}$ by the well known formula

$$f(I+\tau\varkappa\varkappa^+) = f(1)I+\frac{1}{\varkappa^+\varkappa}(f(1+\tau\varkappa^+\varkappa)-f(1))\varkappa\varkappa^+ \qquad (5.7)$$

we obtain from (5.4) the minimum probability of error

$$P_e^0 = (1-M^{-1})M^{-1}(\sqrt{[1+(M-1)\Delta]}-\sqrt{[1-\Delta]})^2. \qquad (5.8)$$

The same expression for this special case was found by Yuen, Kennedy and Lax [2]. The expression (5.8) has a sense with $(1-M)^{-1} \leq \Delta \leq 1$. That coincides with the condition of non-negative definiteness of the matrix (5.6). It follows from (5.8) that $P_e^0 \to \Delta$ with $M \to \infty$. With $\Delta = (1-M)^{-1}$ the rank of the matrix Q is equal $M-1$, the vectors $|\psi_k\rangle$ form a regular simplex. The minimum probability of error of the optimal discrimination of the simplex set $\{|\psi_k\rangle\}$ has the simple sight: $P_e^0 = M^{-1}$ (as it follows from (5.8)) and with $M \to \infty$ tends to zero (the simplex vectors with $M \to \infty$ are being orthogonalized). With $\Delta = 1$ rank $Q = 1$, all vectors coincide: $|\psi_k\rangle = |\psi_l\rangle$ and are completely indiscriminate, that corresponds with the prior probability of error $P_e^0 = 1-M^{-1}$. With $(1-M)^{-1}<\Delta<1$ the matrix $Q^{1/2}$ is strictly positive defined, and the inverse matrix $Q^{-1/2}$ may be found by the formula (5.7). This matrix defines the optimal decision set (4.9) of the orthogonal vectors $|\chi_j\rangle$ having a form

$$|\chi_j\rangle = (1-\Delta)^{-1/2}\left(M^{1/2}|\psi_j\rangle-\left(1-\sqrt{\frac{1-\Delta}{1+(M-1)\Delta}}\right)M^{-1}\sum_k M^{1/2}|\psi_k\rangle\right). \qquad (5.9)$$

Homogeneous sets of coherent signals. In order for the coherent states $|\alpha_k\rangle$ to form a homogeneous set it is sufficient for them to be equiprobable and defined by a homogeneous set of complex amplitudes α_k. Indeed, if the matrix

$S_{kl} = \alpha_k^+\alpha_l$ is an invariant under translation: $S_{kl} = S_{k+j,l+j} \equiv S_{k-l}$ (in the case of finite M it is cyclic: $S_j = S_{j-M}$) then the scalar products (4.12) are also invariant which in the case π_k = const is also equivalent to the invariancy $Q_{kl} = Q_{k-l}$. However, for the homogeneousness of the set $\{|\alpha_k\rangle\}$ the formulated condition is not necessary. In the single mode case, besides the condition of equiprobability, the following fact is necessary and sufficient: equidistant disposition of the representing points α_k on the complex plane α, either on the circle of any radius with zero center $\alpha = 0$; or on any right line containing the center $\alpha = 0$ (infinite number of hypotheses). Let us consider the following examples.

1) *Equiangle and optimal signals*. Let the equiprobable signals α_k, $k = \overline{1, M}$ form in $\mathscr{L}$ an equiangle set:

$$\alpha_k^+\alpha_l = \varepsilon(\delta_{kl}+\delta(1-\delta_{kl})) = \varepsilon(1-\delta)\left(\delta_{kl}+\frac{\delta}{1-\delta}\right). \tag{5.10}$$

Substituting (5.10) into (4.12) we obtain that the matrix $Q = M^{-1}\|\langle\alpha_k|\alpha_l\rangle\|$ has the same structure (5.6), i.e. the set of states $|\psi_k\rangle = M^{-1/2}|\alpha_k\rangle$ is equiangle with the angle $\Delta = \exp\{\varepsilon(\delta-1)\}$. Hence, the minimum probability of error (5.8) of M-ary detection of the coherent signals $\alpha_k = \alpha_k(t)$ with the finite equal energy $\varepsilon = \alpha_k^+\alpha_k$ differs from zero even if they are orthogonal (i.e. $\delta = 0$). Moreover, the minimum of the probability of error with the condition of the energy limitation $\alpha_k^+\alpha_k \leqq \varepsilon$ is obtained not on the orthogonal signals, but on the signals forming the regular simplex in $M-1$-dimensional subspace $\mathscr{L}$ for which $\delta = (1-M)^{-1}$. This follows from the monotonously non-decreasing dependence P_e^0 from Δ (5.8) and from the condition of non-negativity of the matrix (5.10): $\varepsilon \geqq 0$, $(1-M)^{-1} \leqq \delta \leqq 1$. Let us remark that for the optimal "simplex" signals α_k the states $|\alpha_k\rangle$ remain linearly independent. Therefore optimal detection of these signals is described (in the case $(1-M)^{-1} \leqq \delta < 1$) by the orthogonal decision set (5.9), where $M^{1/2}|\psi_k\rangle = |\alpha_k\rangle$ and $\Delta = \exp\{\varepsilon(\delta-1)\}$. When $\varepsilon \ll 1$ (weak signals) for the minimum probability of error (5.8) the next asymptotic formula takes place:

$$P_e^0 = (1-M^{-1})(1-2\sqrt{[(1-\delta)\varepsilon M^{-1}]})+O(\varepsilon). \tag{5.11}$$

2) *Optimal estimation of phase*. Let $|\psi_k\rangle$ be $M^{-1/2}|\alpha_k\rangle$, where $\alpha_k = e^{-2\pi ik/M}\varepsilon^{1/2}\psi(t)$, $\int_0^T|\psi(t)|^2\,dt = 1$ (single mode impulse-code phase modulation). Taking into account (4.12) we obtain in this case

$$Q_{kl} = M^{-1}\exp\{\varepsilon(e^{2\pi i(k-l)/M}-1)\} = Q_{k-l}. \tag{5.12}$$

Hence the set of states $M^{-1/2}|\alpha_k\rangle$ is homogeneous. Optimal detection of such signals is described by the formulas (5.3), (5.4), and the matrixes

$Q^{1/2}$, $Q^{-1/2}$, can easily be calculated in the discrete Fourier representation (5.5). It is easy to write a simple analytical formula for the cases $M = 2, 3, 4, \ldots$.

Let us consider a continuous analog of this problem when $M = \infty$, namely the problem of estimation of the phase x of a regular signal $\alpha_x = e^{-ix}\varepsilon^{1/2}\psi(t)$. The coherent states with uniformly distributed phase x over the interval $[-\pi, \pi]$ $|\psi_x\rangle = (2\pi)^{-1/2}|\alpha_x\rangle$ have the continuous matrix of scalar products

$$\langle\psi_x|\psi_y\rangle = (2\pi)^{-1}\exp\{\varepsilon(e^{i(x-y)}-1)\}.$$

This matrix has the eigen-vectors $\varphi_n(x) = (2\pi)^{-1/2}e^{-ixn}$ and eigen-values $q_n = \varepsilon^n e^{-\varepsilon}/n!$, $n = \overline{0, \infty}$. The optimal orthogonal set of vectors (5.3) is complete with the measure dx on $[-\pi, \pi]$. The phase decision vectors can be represented as

$$|\chi_x\rangle = (2\pi)^{-1/2}\sum_{n=0}^{\infty} e^{ixn}|n\rangle,$$

where $|n\rangle$ are the eigen-vectors of the operator of number of particles a^+a ($a = \int_0^T\psi(t)^*a(t)\,dt$ is the well known operator of annihilation in the state $\psi(t)$). Optimal estimation of the phase is realised by measurement of the operator

$$\hat{x} \equiv \int_{-\pi}^{\pi}|\chi_x\rangle x\langle\chi_x|\,dx = \sum_{n\neq m}^{\infty}|n\rangle i(m-n)^{-1}\langle m| =: \mathrm{Arg}\,(a):,$$

which it is natural to call as phase operator (the symbol $:\ldots:$ is normal ordering symbol [16]).

3) *Optimal estimation of amplitude*. Let $\{|\psi_k\rangle\}$ be an infinite set of equally probable coherent states $|\alpha_k\rangle$, where $\alpha_k = k\Delta\psi(t)$, $k = \overline{0, \infty}$, $\int|\psi(t)|^2\,dt = 1$, and Δ is the step of quantization (single mode impulse-code amplitude modulation). In this case we obtain:

$$\langle\alpha_k|\alpha_l\rangle = \exp\{-\Delta^2(k-l)^2/2\}. \tag{5.13}$$

In order to find the solution, (5.3), (5.4), of the optimal discrimination problem of the homogeneous set $\{|\alpha_k\rangle\}$ it is necessary to extract the root of the matrix (5.13) by using the continuous Fourier transform (5.5). With $\Delta \ll 1$ the optimal discrimination problem is reduced to the problem of optimal measurement of amplitude observable $\hat{x} = \mathrm{Re}\,(\psi^+a)$ ($\psi^+a = \int_0^T\psi^*(t)\,a(t)\,dt$). The result x of this measurement gives the number k of the hypothesis with minimum error (here $k = [x/\Delta]$, $[x/\Delta]$ is integer of x/Δ). In limit $\Delta \to 0$ this problem turns into the problem of optimal estimation of amplitude $x = \psi^+\alpha$ of the coherent quantum signal by the criterion of maximum likelihood [7]. The optimal discrimination in this case is obtained by exact measurement of operator of amplitude $\mathrm{Re}\,(\psi^+a)$.

6. PROBLEM OF M-ARY DETECTION OF STOCHASTIC QUANTUM SIGNALS

1) Stochastic boson signals are described by density operators which can be represented by mixture of coherent projectors $|\alpha\rangle\langle\alpha|$:

$$\rho(s) = \int_\Omega |\alpha(s, \omega)\rangle\langle\alpha(s, \omega)| P(d\omega). \tag{6.1}$$

Here $\alpha(s, \omega)$ is a complex amplitude in the signal observation space $\mathscr{L}$ depending on the useful signal s and on random factors $\omega \in \Omega$ caused by the channel noise or by an uncertainty about some parameters of the signal α. The representation (6.1) is called a Glauber P-representation. Two points of view are possible: 1) the distributions P are assumed to be the usual probability distributions [16], but then not each density operator is P-representable; and 2) singular distributions are admitted [15] from such a class of generalised functions in which every density operator could be represented as (6.1). Then not each distribution can be given a probability interpretation. If Π_j are any decision operators then the probability $\Pr(j|k) = \mathrm{Tr}(\Pi_j \rho(s_k))$ of jth decision with the condition that the signals s_k have taken place in P-representation can be calculated by using the "diagonal" elements $\langle\alpha|\Pi_j|\alpha\rangle$:

$$\Pr(j|k) = \int \langle\alpha(s_k, \omega)|\Pi_j|\alpha(s_k, \omega)\rangle\, P(d\omega), \tag{6.2}$$

where integral has generalised sense when $P(d\omega)$ is singular distribution.

The Glauber P-representation is not very convenient for solving the quantum problem of optimal detection because not every non-negative test function $\varphi(\alpha)$ defined by the formula $\varphi(\alpha) = \langle\alpha|\Pi|\alpha\rangle$ corresponds to a non-negative operator Π. The "Π-representation" conjugated to the Glauber P-representation

$$\Pi_j = \int_{\mathscr{L}} \Pi_j(\alpha)|\alpha\rangle\langle\alpha| d\mu(\alpha) \tag{6.3}$$

proves to be more convenient than the Glauber P-representation. Here μ is Lebesgue measure on the complex signal space $\mathscr{L}$. It is convenient to normalize it in a manner so that the integral $\int|\alpha\rangle\langle\alpha|\, d\mu(\alpha)$ is the resolution of identity $\mathbf{1}$ in Hilbert space $\mathscr{H}$. This measure is defined on each finite dimensional r-subspace in such a way: $d\mu(\alpha) = \prod_{i=1}^r (\pi)^{-1}\, d\,\mathrm{Re}\,\alpha_i\; d\,\mathrm{Im}\,\alpha_i$. The function $\Pi_j(\alpha)$ defines the probabilities (6.2) for every state $\rho(s_k)$ by integration in $\mathscr{L}$:

$$\Pr(j|k) = \int \Pi_j(\alpha)\langle\alpha|\rho(s_k)|\alpha\rangle\, d\mu(\alpha), \tag{6.4}$$

where the entry $\Pi_j(\alpha)$ meaning an abbreviated entry representing the real $\mathrm{Re}\,\alpha$ and the imaginary $\mathrm{Im}\,\alpha$ parts of α, or α, α^*. It is obvious that the formula (6.3) compares every set of regular functions $\Pi_j(\alpha) \equiv \Pi_j(\alpha, \alpha^*)$ on $\mathscr{L} \oplus \mathscr{L}^*$ satisfying the conditions $\Pi_j(\alpha) \geq 0$ $\sum_{j=1}^M \Pi_j(\alpha) = \mathbf{1}$ with some set of decision operators, Π_j. But like the P-representation not all decision operators

Π_j can be represented by regular non-negative functions $\Pi_j(\alpha)$, and in the general case the functions $\Pi_j(\alpha)$ must be treated as generalized ones, i.e. as the linear continuous functionals on the space of test functions being represented as $p(\alpha) = \langle\alpha|\rho|\alpha\rangle$. Every test function $p(\alpha) \equiv p(\alpha, \alpha^*)$ associated with some density operator ρ is positive, normalized: $\int p(\alpha)\, d\mu(\alpha) = 1$ and can be interpreted as probability density of the results of the coherent measurement [5] described by the non-orthogonal resolution of the identity $\mathbf{1} = \int|\alpha\rangle\langle\alpha|\, d\mu(\alpha)$ in all coherent projectors $|\alpha\rangle\langle\alpha|$. In accordance with the fundamental properties of the probabilities (6.4) the generalized functions $\Pi_j(\alpha)$ must be normalized, $\sum_{j=1}^M \Pi_j(\alpha) = 1$, and non-negative in the sense that every test function $p(\alpha)$, for which associated with its operator ρ is non-negative definite the value of the functional $\int \Pi_j(\alpha) p(\alpha)\, d\mu(\alpha)$ is non-negative. In the case when the functions $\Pi_j(\alpha)$, $j = \overline{1, M}$ are regular non-negative functions, they can be interpreted as the conditional probabilities $\Pr(j|\alpha)$, or ordinary random rules with respect to the results of coherent measurement. The decision operators Π_j represented as (6.3) by the regular decision functions $\Pi_j(\alpha)$ we shall call Π-representable.

2) Using the idea of Π-representation a universal suboptimal algorithm of M-ary detection of a quantum stochastic boson signals can be given. Let $\rho(s_k)$, $k = \overline{1, M}$ be quantum statistical hypotheses to be tested, i.e. density operators $\rho_k \equiv \rho(s_k)$, which are described by the density functions

$$p_k(\alpha) = \langle\alpha|\rho_k|\alpha\rangle, \qquad k = \overline{1, M}$$

in Π-representation. In accordance with (6.4) the average probability of decision error $P_e = 1 - \sum_{k=1}^M \pi_k \Pr(j|k = j)$ with given prior probabilities π_k is represented as follows

$$P_e = 1 - \sum_{j=1}^M \pi_j \int \Pi_j(\alpha)\, p_j(\alpha)\, d\mu(\alpha). \tag{6.5}$$

Having confined oneself by the class of Π-representable decision operators it is easy to find a minimum of the probability (6.5). Indeed, if the functions $\Pi_j(\alpha)$ are regular, then the integrals in (6.5) have the ordinary sense and minimum with the conditions $\Pi_j(\alpha) \geq 0$, $\sum_{j=1}^M \Pi_j(\alpha) = 1$ is realised on the characteristic functions of the regions

$$E_j = \{\alpha : \pi_j p_j(\alpha) = \max_k \pi_k p_k(\alpha)\} \tag{6.6}$$

of the complex space $\mathscr{L}$, where the posterior probabilities are maximal:

$$\Pi_j(\alpha) = 1 \text{ when } \alpha \in E_j,\ \Pi_j(\alpha) = 0 \text{ when } \alpha \,\bar{\in}\, E_j. \tag{6.7}$$

The problem of finding the splitting of the signal space $\mathscr{L} = \sum_{j=1}^M E_j$

minimising the corresponding with (6.7) probability of error

$$P_e = 1 - \sum_{j=1}^{M} \pi_j \int_{E_j} p_j(\alpha)\, d\mu(\alpha),$$

i.e. the regions (6.6), is a typical problem of the classical theory of M-ary detection. The decison operators

$$\Pi_j = \int_{E_j} |\alpha\rangle\langle\alpha|\, d\mu(\alpha) \quad (6.9)$$

corresponding with every such split are non-orthogonal, and describe the ordinary non-randomised decision rule (6.7) with respect to the results α of the coherent measurement $|\alpha\rangle\langle\alpha|$, or the indirect measurement of the non-commuting operators of creation and annihilation [7]. In the classical limit, the suboptimal operators (6.9) are orthogonal and absolutely optimal. That is why the universal sub-optimal operators (6.9) can be called quasi-classical.

3) Let us consider some typical problems of the detection of quantum stochastic signals $\alpha(s, \omega)$. The densities $p_k(\alpha) = \langle\alpha|\rho(s_k)|\alpha\rangle$ describing in the Π-representation a quantum boson channel, can be calculated by the integration

$$p_k(\alpha) = \int_\Omega \exp\{-|\alpha - \alpha(s_k, \omega)|^2\}\, P(d\omega) \quad (6.10)$$

of the "elementary" Gaussian conditional densities $|\langle\alpha|\alpha(s, \omega)\rangle|^2 = \exp\{-|\alpha-\alpha(s, \omega)|^2\}$ according to (6.1). For simplicity we shall consider the densities (6.10) only on some finite-dimensional signal subspace $\mathscr{L}$, on which every signal $\alpha(t)$ can be represented by the set of complex numbers $\alpha = \{\alpha_i, i = \overline{1, r}\}$ with respect to any orthogonal basis $\{\varphi_i(t), i = \overline{1, r}\}$: $\alpha(t) = \sum_{i=1}^{r} \varphi_i(t)^* \alpha_i \equiv \varphi(t)^+\alpha$. It proves to be sufficient for solving the problems of detection of a finite number of signals $\{s_k(t), k = \overline{1, M}\}$ in linear communication channels $\alpha(s, \omega) = A(\omega)s + n(\omega)$, because the dimension of the space generated by this signals does not exceed M.

1) *Phase-coherent detection of Gaussian signals.* Let the quantum channel be linear $\alpha(s, \omega) = (1+m(\omega))s + n(\omega)$, where $m(\omega), n(\omega)$ are additive and multiplicative complex Gaussian noises. Gaussian quantum-statistical hypotheses corresponding with the signals s_k are described by the Gaussian normal-ordered operator expressions of the operators of equation $a^* = \{a_i^*, i = \overline{1, r}\}$ and annihilation $a = \{a_i, i = \overline{1, r}\}$:

$$\rho_k = |L_k^{-1}| : \exp\{-(a-s_k)^+ L_k^{-1}(a-s_k)\} : \quad (6.11)$$

Here $s_k = \{s_{ki}, i = \overline{1, r}\}$ is complex vector (column) representing the regular useful part of the observed signal α in r-dimensional space $\mathscr{L}$, L_k is a total correlation matrix in $\mathscr{L}$, L_k^{-1} is the inverse matrix, $|L^{-1}| \equiv \det L^{-1}$, and the symbol : : is the normal ordering symbol meaning that the operators a^* act from the left hand, and a act from the right hand. For every hypothesis k the matrix L_k is positive definite and is equal to the sum $I + N + S_k$ of the identity matrix I which corresponds with the quantum zero noise, the intensity of which is taken by the unity, of the correlation matrix N of the additive complex Gaussian noise $n(t, \omega)$, thermal one for example, and of the non-negative definite matrix S_k, defining the intensity of the stochastic useful part $m(\omega)s_k(t)$ of the observed signal, containing the multiplicative complex Gaussian noise $m(\omega)$ caused by Rayleigh fading for example. When $m(\omega)$ is one complex Gaussian number (with uniformly distributed phase), the matrixes S_k have the unit rank: $S_k = \sigma^2 s_k s_k^+$ (σ^2 is the intensity of m: $\sigma^2 = \langle|m|^2\rangle$), and such quantum channel is the analog of the Rice channel.

The solution of the general quantum Gaussian problem of the optimal M-ary detection is unknown. In the case when $\alpha \simeq s_k$, that is the intensities of the additive and multiplicative noises may be disregarded in comparison with the quantum zero noise: $N \ll I$, $S_k \ll I$, every density operator (6.11) coincides in zero approximation with the coherent projector

$$\rho_k = : \exp\{-(a-s_k)^+(a-s_k)\} : = |s_k\rangle\langle s_k|. \quad (6.12)$$

As an approximate optimal solution, the solution found in fourth paragraph for the case of completely coherent signals $\alpha_k = s_k$ can be used. When at least one of the two conditions $N \gg I$, or $S_k \gg I$ is fulfilled, the optimal solution is the quasi-classical solution (6.6), (6.9), where

$$p_k(\alpha) = |L_k^{-1}| \exp\{-(\alpha - s_k)^+ L_k^{-1}(\alpha - s_k)\}. \quad (6.13)$$

The latter follows from the well known formula [15], [16]:

$$\langle\alpha| : f(a^*, a) : |\alpha\rangle = f(\alpha^*, \alpha),$$

which is true for any normal ordered operator $\rho = :f(a^*, a):$. In the case when the observed signal does not contain a multiplicative noise ($S_k = 0$) and the signal to noise ratio $s_k^+ L^{-1} s$ is constant for all hypotheses k, the optimal regions (6.6) are

$$E_j = \{\alpha : \mathrm{Re}\,(s_j^+ L^{-1}\alpha) = \max_k \mathrm{Re}\,(s_k^+ L^{-1}\alpha)\}. \quad (6.14)$$

It may be assumed that this classical solution is close to the optimal one not only in quasi-classical region but in the essentially quantum one, at least for the multi-threshold detection $M \gg r$. The foundation of such an assumption is the optimality of the coherent measurement and linear estimation in quantum linear channels containing Gaussian additive noises. The most simple proof of this optimality is given in [11] for the criterion of maximum likelihood.

2) *Detection of the signals with unknown phase.* If the Gaussian channel described in previous item is not phase-coherent, then the phase θ of the signal $s(t)$ must be considered as an unknown random parameter: $\theta = \theta(\omega)$, having a uniform distribution. Placing in (6.11) $s_k = \lambda_k^{1/2} e^{-i\theta}\psi_k$ ($\lambda_k = s_k^+ L_k^{-1} s$ is the energy of the kth signal), expending the exponent: $\exp\{2\lambda_k^{1/2} \operatorname{Re}(e^{i\theta}\psi_k^+ L_k^{-1} a)\}$: in a power series of $\lambda_k^{1/2}$, and averaging by θ, we obtain

$$\rho_k = \sum_{n=0}^{\infty} \frac{1}{(n!)^2} |L_k^{-1}| (a^+ L_k^{-1}\psi_k)^n :$$

$$\exp\{-\lambda_k \psi_k^+ L_k^{-1}\psi_k - a^+ L_k^{-1} a\} : (\psi_k^+ L_k^{-1} a)^n \lambda_k^n, \tag{6.15}$$

where $L_k = L + S_k$, $L = I + N$.

A great number of examples of quantum signals described by density operators of Poisson kind (6.15) are given by the quantum optics and by the quantum theory of optical coherence [15, 16]. If the observed signal has a regular part, and multiplicative noise is absent $S_k = 0$, then the corresponding signals are called amplitude-coherent.

The solution of the problem of detection of signals with unknown phase, which is important for optical communication, is not in principal difficult, in the case when all signals are orthogonal $\psi_k^+\psi_l = \delta_{kl}$ and the matrixes L_k, $k = \overline{1, M}$ commute. In such cases the operators ρ_k commute and the solution of considered problem can be found by the classical methods of the theory of statistical decision, because there are the obvious sufficient statistics—these are the numbers of occupation in the orthogonal modes ψ_k. Such commutative cases of particular practical interest are elaborately detailed in [9]. In the general non-orthogonal (that is non-commuting) case the problem of optimal M-ary detection has an infinitely dimensional decision space $\mathscr{H} = \oplus_{n=0}^{\infty} \mathscr{L}^{(n)}$, where $\mathscr{L}^{(n)}$ is the symmetrized nth power of $\mathscr{L}$, generated by n-particle states $(a^+ L^{-1}\psi)^n|0\rangle$. But since every $\mathscr{L}^{(n)}$ is an invariant subspace for density operators (6.15) (i.e. the set $\{\rho_k\}$ is reducible: $\pi_k\rho_k = \oplus_n R_k^{(n)}$) this problem can be reduced to an infinite set of independent optimization problems having the orthogonal decision spaces $\mathscr{L}^{(n)}$ with the dimensions $n!(r-1)!/(n+r-1)!$ In accordance with Theorem 2 we obtain the inequality rank $D_j^{(n)} \leq$ rank $R_j^{(n)}$ for the decision operators $D_j^{(n)}$ in $\mathscr{L}^{(n)}$ defining an optimal resolutions of identities $I^{(n)} = \sum_j D_j^{(n)}$ (complete decision is the infinite direct sum: $\pi_j = \oplus_n D_j^{(n)}$).

Optimal detection of nonorthogonal signals with random phase and zero additive noise. If $L_k = I + \sigma_k^2\psi_k\psi_k^+ (N = 0)$ then all operators (6.15) (or $R_k = \pi_k\rho_k$) are reducible to the operators in $\mathscr{L}^{(n)}$, $n = 0, \ldots, \infty$ with rank $R_k^{(n)} = I$: $R_k^{(n)} = \pi_k^{(n)} (\alpha^+\psi_k)^n|0\rangle\langle 0| (\psi_k^+ a)^n/n!$, where

$$\pi_k^{(n)} = \pi_k(I - x_k) \exp\{(x_k - I)\lambda_k\} \sum_{m=0}^{n} (n!/(m!)^2 (n-m)!)\lambda_k^m x_k^{n-m},$$

$x_k = \sigma_k^2/(I + \sigma_k^2)$. Hence, the optimal detection problem reduces to the problem of optimal discrimination of pure n-particle states

$$|\psi_k^{(n)}\rangle = (\pi_k^{(n)}/n!)^{1/2} (a^+\psi_k)^n|0\rangle,\ k = \overline{I,M}$$

with matrix of scaler products $Q_{kl}^{(n)} = (\pi_k^{(n)}\pi_l^{(n)})^{1/2} S_{kl}^n$, where $S_{kl} = \psi_k^+\psi_l$ The solution is given by Theorem 5, the total probability of error for optimal detection being $P_e^0 = I - \Sigma_n \Sigma_k \mu_k^{(n)}$. If the signals are equiprobable: $\pi_k = M^{-1}$, with equal energies $\lambda_k = \lambda$, $\sigma_k^2 = \sigma^2$, and satisfy the condition of cyclicity: $S_{kl} = S_{k-l} = S_{k-l+M}$, the states $|\psi_k^{(n)}\rangle$ are cyclically invariant, and the minimum probability of error is $P_e^0 = I - \Sigma_n (M^{-1} \operatorname{Tr} S_n^{1/2})^2 P_n$, where $S_n = \|S_{k-l}^n\|$,

$$P_n = (I - x) \exp\{(x - I)\lambda\} \sum_{m=0}^{n} (n!/(m!)^2 (n-m)!)\lambda^m x^{n-m}$$

We can point out that at optical frequencies the thermal noise may be disregarded in comparison with quantum zero one: $N \ll I$. But in general case when $N \neq 0$ the quasi-classical solution (6.6), (6.9) can be used, where

$$p_k(\alpha) = |L_k^{-1}| \exp\{-(\lambda_k\psi_k^+ L_k^{-1}\psi_k + \alpha^+ L_k^{-1}\alpha)\} \sum_{n=0}^{\infty} \frac{1}{(n!)^2} \lambda_k^n |\psi_k^+ L_k^{-1}\alpha|^{2n}. \tag{6.16}$$

This solution is close to the optimal if at least one of the inequalities is valid: $N \gg I$, $S_k \gg I$, $\lambda_k \gg 1$. If the signals are equally probable, and the regular parts have equal energies $\lambda_k = \lambda$ and $S_k = \sigma^2\psi_k\psi_k^+$, then the optimal regions (6.6) look like

$$E_j = \{\alpha : |\psi_j^+ L^{-1}\alpha| = \max_k |\psi_k^+ L^{-1}\alpha|\}, \tag{6.17}$$

where $L = I + N$ and sub-optimal decision rule is the choice of number j for which $|\psi_j^+ L^{-1}\alpha| = \max$ with respect to the result α of coherent measurement. This solution is close to optimal not only in quasi-classical region, but in the essentially quantum one $L_k \approx I$ for multi-threshold detection, when the detection problem can be replaced by the estimation one solved in the next item.

Optimal estimation of mode of quantum signal with known energy and unknown phase. Let us consider an infinite analog of this detection problem: the problem of optimal estimation of complex vector-amplitude $s = \lambda^{1/2} e^{-i\theta}\psi \in \mathscr{L}$ with fixed $\lambda = s^+L^{-1}s$, random θ, Gaussian noises $m, n(\alpha = (\lambda^{1/2}e^{-i\theta} + m)\psi + n)$, and uniform prior distribution in the projection space of $\psi\psi^+$ (*i.e.* maximum likelihood measurement of unknown mode ψ). According to (6.15) the density operator $\rho(\psi)$ associated with hypothesis ψ has the following normal ordering form

$$\rho(\psi) = (I-x)|L^{-1}|e^{(x-1)\lambda} \sum_n \frac{1}{(n!)^2} (1-x)^{2n}\lambda^n :(a^+L^{-1}\psi\psi^+L^{-1}a)^n$$
$$\times \exp\{-a^+(L^{-1}-xL^{-1}\psi\psi^+L^1)a\}:$$

In accordance with §2 to prove the optimality of coherent measurement for maximal likelihood estimation of ψ it is sufficient to find such a Hermitian operator $K \geqq \rho(\psi)$, that $K|\alpha\rangle = \rho(\psi)|\alpha\rangle$ for coherent states $|\alpha\rangle$ with amplitudes $\alpha = \varepsilon^{1/2}e^{-i\theta}\psi$. It is easy to prove that from the matrix inequality $L \geqslant \psi\psi^+$ follows the operator inequality: $(a^+L^{-1}a)^n \exp\{-(I-x)a^+L^{-1}a\}: \geqslant :(a^+L^{-1}\psi\psi^+L^{-1}a)^n \exp\{-a^+(L^{-1}-xL^{-1}\psi\psi^+L^{-1})a\}:$

Since $\psi^+L^{-1}\psi = I$, the sign of equality holds on the sub-space of coherent vectors $|\varepsilon^{1/2}e^{-i\theta}\psi\rangle$ of mode ψ. Hence, the coherent measurement is optimal, and $K = (I-x)|L^{-1}|e^{(x-1)\lambda}\sum_n (n!)^{-2} (I-x)^{2n}\lambda^n: (a^+L^{-1}a)^n \exp\{-(I-x) a^+L^{-1}a\}:$.

3) *Detection of weak non-coherent signals.* There is a special case of particular interest when the regular part of the signals are absent $\lambda_k = 0$. The density operators (6.15) describing such signals are Gaussian with zero expectations

$$\rho_k = |I+S_k|^{-1}: \exp\{-a^+(I+S_k)^{-1}a\}: \tag{6.18}$$

(the additive part N is not being considered here). If all the signals are very weak: $S_k \ll I$, then the problem of optimal detection can be simplified by replacing the operators (6.18) by coinciding with them in the first order of S_k operators

$$\rho_k = (1+\mathrm{Tr}\, S_k)^{-1}(|0\rangle\langle 0| + a^+|0\rangle S_k\langle 0|a), \tag{6.19}$$

where $a^+|0\rangle S_k\langle 0|a = \sum_{i,j=1}^r a_i^*|0\rangle S_k^{ij}\langle 0|a_j$. The density operators (6.19) are concentrated on the direct sum of the vacuum $\mathscr{L}^{(0)}$ and the one-particle $\mathscr{L}^{(1)}$ subspaces that obviously correspond to neglecting the probability of detection of more than one particle (foton) during the observing time interval T. The corresponding decision is $D_j^{(0)} + D_j^{(1)}$, where the decision in $\mathscr{L}^{(0)}$ is trivial: $D_j^{(0)} = \delta_{jk}$ with $k: \pi_k/(1+\mathrm{Tr}\, S_k) \geqq \pi_j/(1+\mathrm{Tr}\, S_j)$ for all $j = \overline{1, M}$.

So, the problem of optimization of detection of weak non-coherent signals in the infinitely dimensional Hilbert space $\mathscr{H} = \oplus \mathscr{L}^{(n)}$ is reduced in first approximation to the considerably simpler problem of finding the optimal resolution of the identity $I = \sum_{j=1}^M D_j$ in the subspace $\mathscr{L}^{(1)}$, i.e. in the space $\mathscr{L}$, generated by the matrices $\{S_k, k = \overline{1, M}\}$. Using the results of previous paragraphs, let us consider the following two cases. For simplicity we suppose that the trace $\mathrm{Tr}\, S_k$ does not depend on k: $\mathrm{Tr}\, S_k = \varepsilon$ and the hypotheses are equally probable: $\pi_k = M^{-1}$. The general one-particle case (6.19) can be also reduced to it by a proper choice of the prior probabilities:

$$\pi_k = (1+\mathrm{Tr}\, S_k)M^{-1}(1+\varepsilon)^{-1}, \text{ where } \varepsilon = M^{-1}\sum_{k=1}^M \mathrm{Tr}\, S_k.$$

Optimal M-ary detection of weak two-mode signals. If two communicating orthogonal modes $\varphi_1(t), \varphi_2(t)$ are only used (for instance, with polarization modulation) then the matrixes S_k, $k = \overline{1, M}$ are all concentrated in two-dimensional subspace generated by the linear combinations $\lambda_1\varphi_1(t) + \lambda_2\varphi_2(t)$. The solution is given by the formulas (3.13), (3.14), where $\pi_k = \varepsilon_k/M(1+\varepsilon)$, $\mathbf{p}_k = \mathbf{e}_k/M(1+\varepsilon)$, in accordance with the decompositions $S_k = (\varepsilon_k + \hat{e}_k)/2$ of the 2×2-matrixes S_k in Pauli matrixes (3.1). The minimum probability of error (2.9) is

$$P_e^0 = 1 - M^{-1}(1+\varepsilon)^{-1}(1 + \sum_{j=1}^M \lambda_j(\varepsilon_j + |\mathbf{e}_j - \mathbf{e}|)),$$

where $\mathbf{e} = \sum\lambda_j\mathbf{e}_j$ is the centre of a simplex subset $\{\mathbf{e}_{j_\alpha}\}$, satisfying the condition of maximality $\varepsilon_{j_\alpha} + |\mathbf{e}_{j_\alpha} - \mathbf{e}| = \max$.

Let us note that the one particle approximation (6.19) is although fit for the partially coherent signals with unknown phase described by the operators (6.15). For instance, when all signals are amplitude-coherent and weak: $L_k = L$, $\lambda_k \ll 1$, we have the special case of (6.19): $S_k = N + \lambda_k\psi_k\psi_k^+$.

Optimal detection of weak amplitude-coherent signals. This case in first approximation can be represented by the pure states $a^+\psi_k|0\rangle$. Having chosen in the proper way the prior probabilities: $\pi_k = (1+\varepsilon_k)/M(1+\varepsilon)$, where $\varepsilon_k = \mathrm{Tr}\, N + \lambda_k$, we obtain the case (2.8), where the operators

$$R_k = M^{-1}(1+\varepsilon)^{-1}\lambda_k a^+\psi_k|0\rangle\langle 0|\psi_k^+ a \tag{6.20}$$

have the rank being equal to unity. The solution is given by Theorem 5, where $|\psi_k\rangle = \lambda_k^{1/2}a^+\psi_k|0\rangle$. The minimum probability of error is

$$P_e^0 = 1 - M^{-1}(1+\varepsilon)^{-1}(1 + \mathrm{Tr}\, N + \sum_{j=1}^M \mu_j),$$

where $\mu_j > 0$ satisfy the equations

$$\lambda_j \langle 0|\psi_j^+ a|K^{-1}|a^+ \psi_j|0\rangle = 1,$$

$$K = \Big(\sum_{k=1}^{M} \mu_k \lambda_k a^+ \psi_k|0\rangle\langle 0|\psi_k a\Big)^{1/2}.$$

If the matrix $S = \|\lambda_k^{1/2}\psi_k^+\psi_l\lambda_l^{1/2}\|$ satisfies the condition $(S^{1/2})_{kk} = (S^{1/2})_{ll}$, the minimum probability of error can be represented in the following invariant form

$$P_e^0 = 1 - M^{-1}(1 + ((\mathrm{Tr}\, S^{1/2})^2 - \mathrm{Tr}\, S)/\mathrm{Tr}(I+N+S)).$$

In conclusion the author expresses gratitude to Professor Stratonovich for fruitful discussion of the results and for useful remarks.

References

[1] C. W. Helstrom, Detection theory and quantum mechanics, *Inform. Control* **10** (1968), No. 3, 264–291.
[2] H. P. Yuen, K. S. Kennedy and M. Lax, On optimal quantum receiver for digital signal detection, *Proc. I.E.E.E.* **58** (1970), 1770–1773.
[3] R. L. Stratonovich, The quantum generalization of optimal estimation and hypothesis testing, *Stochastics* **1** (1973), 87–126.
[4] B. A. Grishanin, Some methods of solution of quantum detection and measurement problems, *Proc. A.S., U.S.S.R., Tecnich. Cybernetica* **5** (1973), 127–137 (in Russian).
[5] V. P. Belavkin and R. L. Stratonovich, On optimization of processing of quantum signals by information criterion, *Radio Eng. Electron. Phys.* **18** (1973), 1839–1844.
[6] C. W. Helstrom, Quantum detection and estimation theory—A review, I.E.E.E. International Symposium on Information Theory, Ashkelon, Israel, June 25, 1973.
[7] V. P. Belavkin, Linear estimation of non-commuting observables by their indirect measurement, *Radio Eng. Electron. Phys.* **17** (1972), 2534–2540.
[8] P. A. Bakut and S. S. Shchurov, Optimal detection of quantum signal, *Probl. Peredachi Informatsii* **4** (1968), No. 1, 77–82 (in Russian).
[9] J. V. S. Liu, Reliability of Quantum-mechanical communication systems, *I.E.E.E. Trans. Inform. Theory*, **IT-16** (1970), 319–329.
[10] A. S. Holevo, Statistical problems in quantum physics, *Proc. 2nd Japan-U.S.S.R. Symposium on Probability Theory* **1** (1972), 22–40.
[11] V. P. Belavkin and A. G. Vancjan, On sufficient conditions of optimality of quantum signal processing, *Radio Eng. Electron. Phys.* **19** (1974), 1397–1401.
[12] M. A. Neumark, On a representation of additive operator set functions, *Comptes Rendus de l'Acad. des Sciences de l'U.R.S.S.* **41** (1943), 359–361.
[13] R. Bellman, *Introduction to the Matrix Analysis*, McGraw-Hill Book Company, Inc., New York, Toronto, London, 1960.
[14] V. P. Belavkin, Optimal linear random filtration of quantum boson signals, *Problems of Control and Information Theory* **3** (1974), 47–62.

[15] J. R. Klauder and E. C. G. Sudarshan, *Fundamentals of Quantum Optics*, W. A. Benjamin, Inc., New York, Amsterdam, 1968.
[16] R. J. Glauber, Optical coherence and foton statistics, in *Quantum Optics and Electronics*, Gordon and Breach, New York, 1965, 65–185.