ON IMPORTANT INFORMATION DYNAMICS

A.K. ALEXEEV[1], A.E. BONDAREV[2]

[1] Moscow Institute of Physics and Technology
Moscow, Russia
e-mail: aleksey.k.alekseev@gmail.com

[2] Keldysh Institute of Applied Mathematics Russian Academy of Sciences
Moscow, Russia
e-mail: bond@keldysh.ru

Key words: Shannon information, Fisher information, goal functional gradient, adjoint equations, Inverse Problems

Summary. The “important information” is considered as the formalization of the intuitive notions of the information related with “goal”, “instruction”, “control”, “intention”. The equations of dynamics for both vector and tensor components of important information are considered. The applications of this important information in Inverse Problems, its relations with adjoint equations and existing information theories and notions are discussed. The classic and quantum cases are compared. The illustrations are provided that present the generation and transfer of the vector component of important information in inverse computation fluid dynamics problems.

1 INTRODUCTION

The information is one of most popular modern notions used, practically, everywhere. A significant number of concepts related with the information are used on semiformal or totally intuitive levels.

The modern standard, Shannon Information theory [1], is aimed only on the communication problems (coding, compression, noise resistance), without any relations with a meaning of the information. Shannon Information content of text string containing m signs, having probability P_i, has a form: $I_{Sh} = -\sum_i^m P_i Log_2 P_i$. It may be seen, that the Shannon information content does not changes under the permutation of signs. It is difficult to accept that the permutation of characters in Tolstoy’s “War and society” does not affect the information contained in this novel. The reason for this paradox is the lack of coincidence of the intuitive “every day” conceptions of information with the Shannon information.

There are a lot of different intuitive notions of information, commonly used at a “folk” level of rigorousity. For example, data base is considered as information, the receipt of pie and the road arrow are another option. The information considered here is close to latter one. Such information should have “meaning”, be useful, be “valuable”. It is clear, that in certain situation a bit of information may be more important if compare with the gigabyte (an

2010 Mathematics Subject Classification: 35Q93, 35R30, 94A15.
Key words and Phrases: Shannon information, Fisher information, goal functional gradient, adjoint equations, Inverse Problems

information in the arrow that show the way from the library at fire may be more valuable than the information in any bookshelf). Obviously, the same formula may have absolutely different meaning for different persons, i.e. the information should depend on the state of system. The information should also depend on the goal and the available means. Intuitive information implies the existence of the disinformation (has sign). The information notion related to "instructions", "recipes" is widely used at everyday life and in science. It has such properties as "importance of information", "value", "sense", etc. The variants of such information are presented in the books [2,3]. The lack of quantitative definitions is the general drawback of these approaches. However, in a number of problems the corresponding formalism exists for a long time and it seems that troubles are caused by "the effect of the tower of Babel" -the lack of common language and proper interpretation. For example, in statistic estimation[4,5], Inverse statistic problems[6], variational data assimilation[7,8], extreme statements of Inverse problems[9] this concept of information may be formalized in a unique way. For this formalism dubbing we use herein the term "important information". We mean under the important information the knowledge of action that may affect the control parameters in order to reach certain goal with the known uncertainty under current state of system. Below we consider the variational statement of Inverse Problems from this viewpoint . We also discuss the relation of the important information with such notions as intentional information [10], "active information" [11] that treats the information by the quantum mechanics means.

2 VARIATIONAL AND STATISTIC STATEMENT OF INVERSE PROBLEMS

The variational statement of Inverse Problems is rather universal approach[7,8,9]. We consider it, herein, at the example of variational data assimilation problems[7,8] which concern the estimation of control variables u vector of some distributed dynamic system (atmosphere, for example) from observations f^{obs} (pressure, temperature, velocity) under the presence of covariation matrix of the observation error W^{-1}. These problems may be recast into the optimization of the goal functional

$$\varepsilon(u) = ((Au - f^{obs}), W(Au - f^{obs}))/2, \tag{1}$$

where A is the evolution operator (propagator)

$$f = Au \tag{2}$$

and f is the prognosis of the system state vector.

The functional (1) is the discrepancy of the numerical computation (prognosis) and measurements, normalized by the data error. Such functionals are usually constructed via a scalar product in L_2, however, sometimes another spaces are used, H_1, for example.

Sometimes[9] the uncertainty of measurements is not accounted and the goal functional has a simpler form $\varepsilon(u) = ((Au - f^{obs}), (Au - f^{obs}))/2$.

The exact form of the operator A is unknown if the dynamics is governed by a system of partial differential equations (PDE). The numerical solving of PDE system enables to determine the action of this operator $f = Au$. If we can compute the action of the operator (and adjoint operator), we may avoid the usage of operator A explicit form in many practical events. With these reservations, we shall further use the operator A as a symbol, that enables significant reduction of the treatments.

The gradient of the functional (1) with respect to control parameters may be presented as

$$\nabla\varepsilon = A^{*}W(Au - f^{obs}) . \tag{3}$$

The optimal solution may be obtained by the Newton iteration step, based on the assumption of $\nabla\varepsilon = 0$ at the extremum point

$$u = (A^{*}WA)^{-1}b \tag{4}$$

where

$$F = A^{*}WA \tag{5}$$

and

$$b = A^{*}Wf^{obs} . \tag{6}$$

Usually the operator A is defined implicitly via the system of PDE and the term $A^{*}Wf^{obs}$ is also not provided in the explicit form. Fortunately, there exists an efficient technique for determination of $\nabla\varepsilon = A^{*}W(Au_0 - f^{obs})$ by the solution of adjoint equations[9,12]. Using the gradient the Newton iterations may be stated in the following form

$$\Delta u = u - u_0 = -(A^{*}WA)^{-1}\nabla\varepsilon . \tag{7}$$

Much more often, the gradient based methods are applied, which implicitly use some primitive approximation of Fisher matrix:

$$u_i^{n+1} = u_i^n - \tau^n \nabla\varepsilon_i^n . \tag{8}$$

Iterations (7,8) in the contrast to (4), operate in the nonlinear case also.

The variational form of Inverse Problem (usually, Ill-posed) may be easily regularized by adding a penalty term, for example:

$$\varepsilon(u) = (A_{ij}u_j - f_i^{obs})W_{ik}(A_{km}u_m - f_k^{obs})/2 + u_i D_{ij} u_j /2 . \tag{9}$$

The matrix $D_{ij} = \alpha E_{ij}$ (α - regularization coefficient) corresponds the Tikhonov regularization of zero order[13]. Since $\Delta\varepsilon(u) = (\nabla\varepsilon_j \Delta u_j) = A_{ij}\Delta u_j W_{ik}(A_{km}u_m - f_k^{obs}) + u_i D_{ij}\Delta u_j$ and $\nabla\varepsilon_j = A_{ij}W_{ik}(A_{km}u_m - f_k^{obs}) + u_i D_{ij} = 0$ at minimum, the Newton iteration step has the form:

$$u = (A^{*}WA + D)^{-1} A^{*}Wf^{obs} . \tag{10}$$

Generally, the expressions (4-8) are well known in the statistic theory of estimation[4,5]. $F = A^{*}WA$ is the Fisher information matrix, the expression $\nabla\varepsilon = A^{*}W(Au_0 - f^{obs})$ is the informant, score. So, no wonder that the statistic theory of estimation may be easily applied to Inverse Problems. The statistic Inverse Problems approach[6] may be used in linear event if the measurement error is available. In a contrast to deterministic statements of Inverse Problems, this approach provides both the value under the search and its uncertainty. From this viewpoint, the expression (10) defines the solution corresponding the probability density maximum. The matrix D contains a priori information $P(u) = C\exp(-((u, Du)/2)$. The error of Inverse Problem solution is defined by the covariation matrix (inverse Fisher matrix)

$$F^{-1} = (A^* WA + D)^{-1} \tag{11}$$

So, the statistic approach enables to solve an Inverse Problem in the sense of calculating most probable estimation and the covariation matrix of solution error.

3 IMPORTANT INFORMATION

In the domain of Inverse Problems one may define the important information (knowledge that is necessary to reach a goal and to obtain uncertainty of results) that consists of the vector Δu and tensor the $(A^* WA)^{-1} = F^{-1}$ (inverse Fisher information matrix).

The tensor component of the important information may be considered as certain "ellipsoid of concentration", "ellipsoid of errors", which describes the uncertainty of goal estimation. The tensor component of important information is defined by the evolution operator and the error of measurements.

The vector component of important information may be considered as coordinates of aim in the control variables space Δu. Due to linear transformation (7) $\Delta u = (A^* WA)^{-1} \nabla \varepsilon$, the vector component Δu may be considered to be equivalent the goal functional gradient $\nabla \varepsilon$. So, we use $\nabla \varepsilon$ as the vector component of the important information due to well-developed adjoint technique and absence of corresponding equations stated in terms of Δu. Certainly, $\nabla \varepsilon$ provides exact direction to the goal only if $F = E$. The vector component of important information depends on the discrepancy of the prognosis and observation and is propagated by adjoint equations.

4 THE VECTOR COMPONENT OF IMPORTANT INFORMATION IN ADJOINT EQUATIONS

Thus, the goal functional gradient is the key element at solving the Inverse Problem. For the problems governed by PDE, the efficient method for the gradient calculation is based on adjoint problems [12]. Usually, the gradient $\nabla \varepsilon$ is expressed via a combination of the physical and adjoint parameters, or, sometimes, via adjoint parameters only. We consider a model problem of estimation of source term $u(t, x)$ from observations $f^{obs}(t, x)$ in the system governed by the transfer equation

$$\frac{\partial f}{\partial t} + a\frac{\partial f}{\partial x} + u(t, x) = 0 . \tag{12}$$

The goal functional (for lucidity posed without an account of error) is engendered by the scalar product in L_2

$$\varepsilon(u) = \int (f(t, x) - f^{obs}(t, x))^2 / 2dxdt . \tag{13}$$

All necessary expressions may be obtained from the stationarity of the Lagrangian

$$L(u) = \varepsilon(u) + \int \left(\frac{\partial f}{\partial t} + a\frac{\partial f}{\partial x} + u(t,x)\right)\psi(x,t)\,dxdt \ .$$

On the solution of main (12) and adjoint problems

$$\frac{\partial \psi}{\partial t} + a\frac{\partial \psi}{\partial x} + (f - f^{obs}) = 0 \ , \qquad (14)$$

$$\Delta L(u) = \Delta\varepsilon(u) = \int \Delta u(t,x)\psi(x,t)\,dxdt \ ,$$

and the goal functional gradient is equal to the adjoint function

$$\nabla\varepsilon = \psi(t,x). \qquad (15)$$

The source term $f - f^{obs}$ corresponds the discrepancy of calculation and observation and causes the origin of vector component of the important information. The adjoint equation describes a transfer of this component. It should be stressed that in adjoint equations the evolution occurs in the reverse time direction.

The correlation of the of vector component of the important information with the goal functional gradient enables to define some intuitive notions.

The goal may be considered as a goal functional minimum point in the control variables space.

The goal functional gradient may be considered as an instruction, receipt (it shows the direction to the goal).

The dependence of the gradient on the initial state u_0 correlates with the dependence of information on an initial, “a priori” knowledge.

The importance of information (from present goal viewpoint) may be estimated as maximum variation of the goal functional at the movement along gradient.

If we know the exact information $\nabla\varepsilon_{true}$, we may analyze an additional information vector M (for example, message) using the scalar product $(M, \nabla\varepsilon_{true})$ or the distance $\|M - \nabla\varepsilon_{true}\|$.

It enables to define information inaccuracy $\|M - \nabla\varepsilon_{true}\|$, zero information $(M, \nabla\varepsilon_{true}) = 0$ and disinformation $(M, \nabla\varepsilon_{true}) < 0$.

The extent of information novelty may be estimated as the orthogonal part of $M_{\perp}$ with respect to $\nabla\varepsilon_{true}$.

The “price/quality” ratio may be explicitly expressed in the event of Tikhonov regularization of zero order[14]. The relation $\alpha\Delta u + \nabla\varepsilon = 0$ may be recast as $((\alpha\Delta u + \nabla\varepsilon), \Delta u) = 0$. Correspondingly $\alpha(\Delta u, \Delta u) + (\nabla\varepsilon, \Delta u) = \alpha\|\Delta u\|^2 + \Delta\varepsilon$ and “price/quality” ratio may be expressed as the inverse of the regularization coefficient $1/\alpha = \frac{\|\Delta u\|^2}{\Delta\varepsilon} = \frac{(\Delta u, \Delta u)}{(\nabla\varepsilon, \Delta u)}$.

The correlation of the important information vector component and the adjoint parameters sheds a light for the old mystery. The adjoint equations are used in the wide range of practice problems and results are published in thousands of references. However, from the mathematical viewpoint, adjoint equations may be obtained from the Lagrange identity for a scalar product, from the stationary conditions for Lagrangian or from the Green function. It does not indicate the presence of a general nature at all. Nevertheless, in all cases the adjoint equations have the same form and the manner of application. At present, the pragmatic approach dominates and all attention is concentrated on the methods for calculation of adjoint variables without any discussion of their nature. However, it is very strange that the object, used in a standard way in the wide range of applications, has no any physical interpretation. By this reason, the works concerning different treating of adjoint equations' "physical meaning" appear regularly, for example [15].

Lewins (1965)[16] used a notion "importance" for the adjoint function, Marchuk[12] designated the adjoint function as the "function of information importance". These terms have rather vague sense and are of no use. However, the idea to provide some information sense to the adjoint variables seems to be highly perspective one. As we have seen, the goal functional gradient is correlated with the adjoint function (or, even equal to it). So, a meaning of the vector component of important information may be attributed to adjoint parameters, while adjoint equations describe its transfer. Thus, the troubles met at the search of the physical meaning of adjoint equations are caused not by the treating of adjoints, but by the absence of generally accepted qualitative statement of commonly used intuitive information.

For example, let's consider the supersonic flow around circular cylinder under the impinging shock that is governed by the unsteady two-dimensional Euler equations

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho U_k)}{\partial x^k} = 0, \tag{16}$$

$$\frac{\partial (\rho U_i)}{\partial t} + \frac{\partial (\rho U_k U_i + P\delta_{ik})}{\partial x^k} = 0, \tag{17}$$

$$\frac{\partial (\rho E)}{\partial t} + \frac{\partial (\rho U_k h_0)}{\partial x^k} + q(x, y) = 0\,. \tag{18}$$

Here $U_1 = U, U_2 = V$, $\gamma = \dfrac{C_p}{C_v}$, $h = \dfrac{\gamma}{\gamma - 1}\dfrac{P}{\rho} = \gamma e$, $e = \dfrac{RT}{\gamma - 1}$, $E = \left(e + \dfrac{1}{2}(U^2 + V^2)\right)$,

$$h_0 = (U^2 + V^2)/2 + h,\ P = (\gamma - 1)\rho e\,.$$

The Edney-IV scheme of shock interferention occurs at this flow pattern[17] (Fig. 1) with the high rise of the pressure at the body surface. It is interesting to estimate a feasibility to reduce surface pressure by the volume heat sources $q(x, y)$ (laser heating, for example). With this purpose, we consider the minimization of the functional

$$\varepsilon(q) = 1/2 \int_\Gamma p^n d\Gamma\,. \tag{19}$$

The adjoint equations are

$$\frac{\partial \Psi_\rho}{\partial t} + \frac{\partial \Psi_i}{\partial t} U^i + \frac{\partial \Psi_e}{\partial t} E + U^k \frac{\partial \Psi_\rho}{\partial X^k} + U^k U^i \frac{\partial \Psi_i}{\partial X^k} + (\gamma - 1)\frac{\partial \Psi_k}{\partial X^k} e + U^k h_0 \frac{\partial \Psi_e}{\partial X^k} = 0, \tag{20}$$

$$\frac{\partial \Psi_i}{\partial t}\rho+\frac{\partial \Psi_e}{\partial t}\rho U^i+U^i\frac{\partial \Psi_k}{\partial X^i}+U^i\frac{\partial \Psi_i}{\partial X^k}+\frac{\partial \Psi_\rho}{\partial X^k}+h_0\frac{\partial \Psi_e}{\partial X^k}+U^iU^k\frac{\partial \Psi_e}{\partial X^i}=0\,, \tag{21}$$

$$\frac{\partial \Psi_e}{\partial t}+\gamma U^k\frac{\partial \Psi_e}{\partial X^k}+(\gamma-1)\frac{\partial \Psi_k}{\partial X^k}=0\,. \tag{22}$$

Initial conditions: $\Psi_{\rho,U,V,e}\big|_{\tau=0}=0$.

Boundary conditions:

$\Psi_{\rho,U,V,e}\big|_{\partial\Omega}=0$. $\Psi_\rho\big|_\Gamma=nP^{n-1}(\gamma-1)e$, $\Psi_e\big|_\Gamma=nP^{n-1}(\gamma-1)\rho$.

The goal functional gradient coincides with the adjoint inner energy

$$\nabla\varepsilon_q=\Psi_e(t,x,y)\,. \tag{23}$$

Thus, the adjoint parameters may be also considered as the vector component of important information, also. Fig. 2 demonstrates isolines of the adjoint inner energy $\nabla\varepsilon_q=\Psi_e(t,x,y)$, which may be considered as the flowfield of the valuable information.

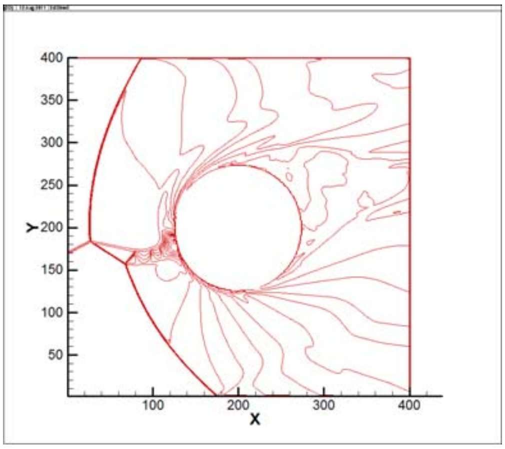

Fig. 1: Density isolines

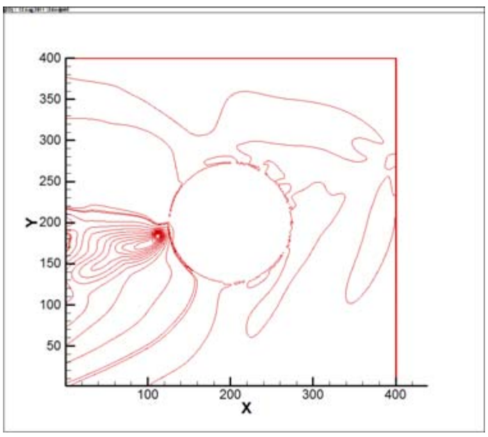

Fig. 2: Gradient isolines $\nabla\varepsilon_q=\Psi_e(t,x,y)$

5 TENSOR INFORMATION

The tensor component of the important information is the most difficult object from the computation viewpoint. Formally, it may be calculated via information dynamics equations of Einstein type[18] for the Fisher information matrix as the metric tensor. Unfortunately, this approach engender extraordinary obstacles from the computation viewpoint. However, the Hessian $H_{jm} = \dfrac{\partial^2 \varepsilon}{\partial u_j \partial u_m}$ of the functional $\varepsilon(u) = (A_{ij}u_j - f_i^{obs})W_{ik}(A_{km}u_m - f_k^{obs})$ may be used instead the Fisher information matrix $F = A^* W A$ ($A_{ij} = \partial f_i / \partial u_j$) in the vicinity of optimal solution. Really, by differentiation of the gradient $\nabla \varepsilon_j = \dfrac{\partial \varepsilon}{\partial u_j} = \dfrac{\partial f_i}{\partial u_j} W_{ik}(f_k - f_k^{obs})$, one may obtain

$$H_{jm} = \frac{\partial^2 \varepsilon}{\partial u_j \partial u_m} = \frac{\partial f_i}{\partial u_j} W_{ik} \frac{\partial f_k}{\partial u_m} + \frac{\partial^2 f_i}{\partial u_j \partial u_m} W_{ik}(f_k - f_k^{obs}) \tag{24}$$

and $H_{jm} \approx F_{jm}$, since in the vicinity of the optimal point $(f_k - f_k^{obs}) \approx 0$. The computation of Fisher matrix via Hessian is much simpler if compare with approach[18], and may be conducted either by second order adjoint equations[19] or by differentiation of the gradient obtained from adjoint equations. However, these approaches need the number of runs of main and adjoint problems proportional to the length of control parameter vector p that may be computationally expensive. This difficulty may be resolved by the iteration approach, where the action of the Hessian by the vector δu is obtained by the gradient differentiation

$$H\delta u = (\nabla \varepsilon(u + a\delta u) - \nabla \varepsilon(u)) / a \tag{25}$$

Since Hessian contains relatively small number of nonzero eigenvalues, corresponding eigenvectors (senior) may be obtained by Lanczos/Arnoldi iterations. The approximation of Hessian based on these vectors may be constructed according to paper[20] as $H_{reduced} = V_r \Lambda_r V_r^*$. Here the matrix $V_r \in R^{p \times r}$ is constructed using r senior eigenvectors of H, corresponding r senior eigenvalues are collected in a diagonal matrix $\Lambda_r = diag(\lambda_i) \in R^{r \times r}$. This trick enables the significant reduction of computations for the case $r << p$.

6 FISHER MATRIX SCALAR INVARIANTS

Fisher's information matrix has the scalar invariants: determinant $\det(F)$ and trace $tr(F)$.

For the normal probability distribution, the Fisher matrix determinant is related with Shannon information content[21]:

$$I_{Sh} = 1/2 \log_2 \det(F) \tag{26}$$

The trace is equal to Fisher information content

$$I_F = tr(F) \tag{27}$$

These invariants have a lot of practical applications, for example, in the paper[21] they are used for the estimation of information content of experimental data.

Shannon information content forms the basis of modern information theory and is extremely widely used.

Fisher information (content) is less popular object, however, of great generality also. For example, Frieden and Soffer [22] applied it for the statement of fundamental physics equations.

It should be noted that the Shannon and Fisher information contents are scalar values, so they provide relatively small knowledge regarding tensor information $F \in R^{p\times p}$ and are not related with the vector component of important information. Shannon and Fisher information contents are suitable for the description of the systems with zero gradient and diagonal Fisher matrices. The state of such system is close to the goal functional extremum and the functional is isotropic in the space of parameters u. The Boltzmann/Shannon information/entropy is widely applied for the thermodynamic systems, the Fisher information also may be considered as another kind of entropy. No wonder that their application is restricted by the vicinity of thermodynamical equilibrium.

7 DISCUSSION

At present, the notion "information" is applied in extremely wide set of contents of different formalization extent. Respectively, the correlation of considered important information with another "informations" varies in broad margins. Several congeneric concepts we should mention.

Shaw et al.[10] for systems with intentional dynamics stressed the role of adjoint equations describing a "field of control-specific information in which the actor and the intended goal both participate". The successful intentional behaviour was assumed to be feasible for self-adjoint equations and was related with the quantum mechanics. Hiley and Pylkkanen[11] defined the information as the quantum potential in the hydrodynamic form of Schrodinger equation (Madelung equation [23]). This potential is related to the Fisher information content. It is main reason for the correlations of quantum mechanics and the information. Reginatto [24] demonstrated that the extreme of Fisher information content

$$I_F = \int \rho\left(\frac{\partial \ln\rho}{\partial x}\right)^2 dxdt \tag{28}$$

on the Hamilton-Jacoby equation may be used for derivation of Schrodinger equation in Madelung form. The conditions of stationarity of the Lagrangian

$$L = \int \left(\rho\frac{\partial S}{\partial t} + \frac{\rho}{2m}(\frac{\partial S}{\partial x})^2 + U(\rho)\right) + \hbar^2 I_F \tag{29}$$

engender the Madelung equations

$$\frac{\partial\rho}{\partial t} + \frac{1}{m}\frac{\partial}{\partial x}\left(\rho\frac{\partial S}{\partial x}\right) = 0,\ \frac{\partial S}{\partial t} + \frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2 + \left(U - \frac{\hbar^2}{2m\rho^{1/2}}\frac{\partial^2\rho^{1/2}}{\partial x^2}\right) = 0,$$

where S is the action and ρ is the probability density ($\rho = \psi\psi^*$), $\psi(t,x) = \rho(t,x)^{1/2} e^{iS(t,x)/h}$ is the wave function, $u = \nabla S$.

This form practically coincides with the Euler equations (16-17)

$$\frac{\partial \rho}{\partial t}+\frac{1}{m}\frac{\partial}{\partial x}\left(\rho \nabla u\right)=0 \tag{30}$$

$$\frac{\partial u}{\partial t}+\frac{1}{2m}\nabla\left(u\right)^2+\nabla\left(U-\frac{h^2}{2m\rho^{1/2}}\frac{\partial^2\rho^{1/2}}{\partial x^2}\right)=0 \tag{31}$$

with the main difference in the pressure term related to the Fisher information. It should be mentioned that, in the contrast to equations (12-14) and (16-17) designed for the search of preassigned goal, this system provides the search for the solution that is maximally uncertain. Frieden[22] considered this situation as a contest between "Mind" (observer, which search for the maximum of information) and Nature, which maximizes disorder (minimizes the Fisher information).

From other viewpoint, it is easy to add the goal term to the Lagrangian (29) and obtain corresponding adjoint equations, as in the paper[25], providing the transfer of vector information. Hawkins and Frieden[26] enhanced this approach by including experimental data in the inverse problem stile (Eq. 14) for the financial economics.

So, in quantum mechanics, the invariant of the tensor component of the important information directly affect the physical dynamics. This provides the basics to relate quantum mechanics with some information notions, such as "intentional information" by Shaw[10] and "active information" by Hiley[11].

However, the only domain where the dynamics of the tensor component of important information is explicitly used is the famous Kalman filter[27]. It estimates a point at trajectory with the ellipsoid of concentration and is also based on the search of Fisher information content extreme. The covariation matrix is propagated by the evolution operator. However, this approach need for the prohibitively high computer resources for PDE based systems.

In general, the logical structure of important information is the same as in statistic estimation problems [4,5]. The distinctions concern interpretations and technique, especially in domains where the dynamics is governed by PDE systems.

As we have already seen, the important information is intimately related with the evolution operator that, herein, is determined by a priori known dynamics of the system, governed by PDE. In this approach, the important information is totally objective one since it is not connected with an activity of any natural or artificial intellect.

This standpoint seems to be too narrow since it is naturally to expect certain isomorphism of information origin and transfer processes both in the object and in the subject. Such isomorphism may be detected for neural networks, since the evolution operator A_{ij} corresponds perceptron weights matrix[28]. So, the above described formalism may be applied to the origin and transfer of the important information in a subject (naturally, if it has the structure of some neural network). If the evolution operator is unknown, it may be determined in the learning process via the set of input data X and the set of observations Y (formally as $A=YX^+$, where X^+ is a pseudoinverse matrix, really using such algorithms as the backpropagation). Thus, the propagator at the learning is constructed using an extra sort of information ("raw information", similar to some data base), which contains input data X and corresponding responses Y.

8 CONCLUSIONS

A formal definition for the important information that possesses many properties of the intuitive concepts of information such as described by Bongard [2] or Chernavskii [3] , may be stated in Inverse Problems.

The important information corresponds the knowledge of the movement in the control variables space that is necessary to achieve the goal with the known uncertainty under the present system state.

The important information has the vector and the tensor components.

The vector component of important information may be related with the goal functional gradient showing a direction to the point of goal functional minimum in control variables space. If the dynamics is governed by PDE, the gradient may be computed using adjoint equations.

The tensor component of important information defines the uncertainty at goal point (ellipsoid of concentration) and corresponds the Fisher's information matrix, which (in PDE event) also may be computed by adjoint equations.

The important information formalism may be applied to modeling of the origin and propagation of information both in the natural systems and in the neural networks.

Acknowledgements

Authors acknowledge the partial support by grants of RFBR № 14-01-00769A and № 16-01-00553A.

REFERENCES

[1] C.E. Shannon, "A Mathematical Theory of Communication", *The Bell System Technical Journal*, **27**, 379–423, 623–656, (1948).

[2] M.M. Bongard, *Pattern Recognition*, New York: Spartan Books, (1970)

[3] D.S. Chernavskii, "The origin of life and thinking from the viewpoint of modern physics", *PHYS-USP*, **43** (2), 151–176, (2000).

[4] R.A. Fisher, *Theory of statistic estimation*, Proc. Cambridge Phil. Soc, Vol 25, (1925).

[5] S. Zacks, *The theory of statistical inference*, NY: Wiley,. (1971).

[6] V.Yu. Terebigh, "Image restoration with minimum a priori information", *Physics – Uspekhi*, **38** (2), 137 – 168, (1995).

[7] F.X. Le-Dimet, O. Talagrand, "Variational algorithms for analysis and assimilation of meteorological observations", *Tellus*, **38A**, 97-110, (1986).

[8] A.C. Lorenc, "Analysis methods for numerical weather prediction", *QJR Meteorol Soc*, **112** (474), 1177-1194, (1986).

[9] O.M. Alifanov, E.A. Artyukhin and S.V. Rumyantsev, *Extreme Methods for Solving Ill-Posed Problems with Applications to Inverse Heat Transfer Problems*, Begell house inc. Publishers, (1996).

[10] R. Shaw, E. Kadar, & Kinsella-Shaw, "Modelling systems with intentional dynamics: A lesson from quantum mechanics", *In K. Pribram (Ed.) Appalacia II: Origins of self-organization. (The Report of the Second Annual Appalachian Conference on Neurodynamics), Hillsdale NJ: Lawrence Erlbaum & Associates*. (1995).

[11] B.J. Hiley, P. Pylkkanen, "Can Mind Affect Matter Via Active Information?", *Mind & Matter,* **3** (2), 7, (2005).

[12] G.I. Marchuk, *Adjoint Equations and Analysis of Complex Systems*, Kluwer Academic Publishers, Dordrecht, Hardbound, (1995).

[13] A.N. Tikhonov and V.Ya. Arsenin, *Solution of Ill-Posed Problems*, N.Y.: Halsted Press, (1979).
[14] J.L. Lions, *Optimal control of systems governed by partial differential equations*, Berlin: Springer-Verlag, (1971).
[15] Q. Wang and J-H. Gao, “The drag-adjoint field of a circular cylinder wake at Reynolds numbers 20, 100 and 500”, *Journal of Fluid Mechanics*, *Cambridge University Press*, **730**, 145- 161 (2013).
[16] J. Lewins, *Importance - The Adjoint Function*, Pergamon Press, (1965).
[17] B. Edney, *Anomalous heat transfer and pressure distribution on blunt bodies at hypersonic speeds in the presence of an impinging shock*, Report N 115, Aeron. Res. Institut of Sweden, (1968).
[18] X. Calmet, J. Calmet, “Dynamics of the Fisher information metric”, *Physical Review E*, **71** (5), 056109, (2005).
[19] A.K. Alekseev, and I.M. Navon, “On Estimation of Temperature Uncertainty Using the Second Order Adjoint Problem”, *Int. J. of Comput. Fluid Dynamics*, **16** (2), 113–117, (2002).
[20] H.P. Flath et al, “Fast algorithms for Bayesian uncertainty quantification in large-scale linear inverse problems based on low-rank partial Hessian approximations”, *SIAM J. Sci. Comp.*, **33** (1), 407-432, (2011).
[21] K. Singh, A. Sandu, M. Jardak, K.W. Bowman, and M. Lee, “A practical method to estimate information content in the context of 4D-Var data assimilation”, *SIAM/ASA Journal of Uncertainty Quantification*, **1**, 106-138, (2013).
[22] B.R. Frieden, B.H. Soffer, “Lagrangians of physics and the game of Fisher-information transfer”, *Phys. Rev. E,* **52**, 2274–2286, (1995).
[23] E. Madelung, “Quantum Theory in Hydrodynamical Form”, *Zeit. f. Phys.*, **40**, 322-325, (1927).
[24] M. Reginatto, “Derivation of the equations of nonrelativistic quantum mechanics using the principle of minimum Fisher information”, *Phys. Rev. A*, **58** (3), 1775–1778, (1998).
[25] B. Dey, H.A. Rabitz and A. Askar, “Optimal control of molecular motion expressed through quantum fluid dynamics”, *Phys. Rev. A.* **61**, 043412-1, (2000).
[26] R.J. Hawkins and B.R. Frieden, “Asymmetric Information and Quantization in Financial Economics”, *International Journal of Mathematics and Mathematical Sciences*, **2012**, Article ID 470293, 11 pages, (2012), doi:10.1155/2012/470293.
[27] R.E. Kalman, “A new approach to linear filtering and prediction problems”, *J. Basic Eng., Ser. D*, **82**, 34–45, (1960).
[28] C. Gershenson, “Artificial Neural Networks for Beginners”, *arXiv: preprint cs/0308031*, (2003).

Received January, 25 2016