ОБЪЕДИНЕННЫЙ
ИНСТИТУТ
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ
ДУБНА

E5-83-693

G.A.Emelyanenko, Š.Šujan

A NEW LOOK UPON INFORMATION PROCESSING IN EXPERIMENTAL PHYSICS. GENERAL STRATEGY

Submitted to "Kybernetika"

1983

1. INTRODUCTION

In contemporary high-energy experimental physics a rather dominant role is played by technique based on film and filmless principles of gaining information, respectively. There exist considerable differences between the two approaches in many respects, including, e.g., the control of measuring systems, the manner of registration, pre-processing, and storage of primary physical information, as well as its employment for subsequent processing. On the other hand, the actual information processing is almost the same within both approaches. The only differences consist of different degrees of precision with which the given structure of physical processes is involved, and of the level of completeness of incorporating prior information. In this paper we shall concern merely the above-mentioned common part of information processing problems.

Theoretical and applied aspects of information processing in high-energy experimental physics are widely investigated (see, e.g., publications of more or less review nature [1-6]). However, a more thorough mathematical analysis of such problems is still merely in initial stage. Consequently, most of contemporary information processing techniques represent a heterogeneous mixture of rigorous mathematical reasoning and of intuitive and heuristic inference.

When compared with previous publications, in this and the subsequent papers, much more attention is paid to the geometrical problems. There exist several reasons for doing so:

(1) Most physical quantities are not directly observable, and their "measurements" are performed in a corresponding geometrical space rather than in the natural physics one;

(2) The optimum choice of a geometric space makes it possible to incorporate more precisely the relations among separate physical processes which, in the whole, influence the geometrical structure and statistical properties of experimental data;

(3) a convenient geometric representation provides us with an exact and clear separation of geometrical and physical parts of processes involved.

In particular, by a proper choice of a geometric representation we may find new connections among physical processes and new factors which influence the properties of experimental data. All these connections have been considered more or less explicitly in previous publications. However, because of lack of sufficiently clear understanding of such connections, they have been usually eliminated from investigations. This amounts to adopting various simplifying assumptions. Contrary to that, we aim at not saying merely that the influence of a certain factor is "small" but also saying which is the "degree of smallness". Finally, our approach contains the previous ones as special cases ,that is, the main difference consists in the fact that we shall not impose simplifications until they will appear really unavoidable.

2. THE PROBLEM OF MODELLING

The contemporary practice of information processing, especially within high-energy experimental physics, typically focuses on problems related to statistical elaboration of data and connected computational questions. The results of theoretical physics (e.g., the theory of scattering processes, etc.) usually are considered as final results, the application of which cannot lead to misunderstandings. Fortunately, this is the actual situation, however, the prize paid for this is rather high because the (sometimes) low adequacy of models is compensated by an extremely large (and extremely expensive) statistics.

In this section we intend to point out that for any concrete experiment it is unaviodable to start with modelling problems, which should involve both theoretical results and the concrete setup of a given experiment. A formal application of theoretical results may lead to unprecise or even false conclusions.

A basic information-processing problem (for a large class of physical experiments), from the mathematical point of view, is the evaluation of an unknown physical parameter $\vec{\theta}$; $\vec{\theta}$ is considered as a vector which is described by a pair $(\vec{n}, \theta)$, where $\theta = |\vec{\theta}|$ is the modulus of $\vec{\theta}$ and $\vec{n} = \vec{\theta}/\theta$.

We are interested mainly in experiments based on various kinds of registration of the particle tracks. Hence, it is necessary to add the dynamical information about $\vec{\theta}$. Let a time interval $[0, T]$ be given. We assume that the dynamics is described (at present, purely formally) by an equation of motion

$$\frac{d\vec{\theta}}{dt} = \vec{A}(t); \quad t \in [0, T]. \tag{2.1}$$

A typical problem we meet is to find an estimate of $\vec{\theta}(r)$ at some $r \in [0, T]$, usually at $r = 0$. Suppose for a moment that the space $\mathcal{P}$ of parameters $\vec{\theta}$ is directly observable. In this case the problem reduces to a classical filtration problem. Indeed, let $(\vec{\theta}(t_1), \vec{\theta}(t_2), \ldots, \vec{\theta}(t_N))$, $t_i \in [0, T]$ denote the vector of observations along the trajectory of the process

$$\frac{d\vec{\theta}}{dt} = \vec{A}(t) + \vec{E}(t), \tag{2.2}$$

where $\vec{E}$ designates the measurement noise process. Depending on its statistical properties we define a norm $\|\cdot\|$ on the sample space and find then the value $\vec{\theta}(0)$ for which the trajectory of (2.2) exhibits the best fit to the observed data.

Unfortunately, the problem becomes much more involved for:

(1) We carry over our measurements within a different space $\mathcal{K}$ (usually, the coordinate-space determined by the registration device).

(2) The formal operator $\vec{A}(t)$ in (2.1) depends on chance, too, so that there appear problems even when attempting to define what is a trajectory of that equation. As a consequence of (1) and (2), a third difference appears which prevents us from using the filtration approach. Namely, the formal sum $\vec{A}(t) + \vec{E}(t)$ in (2.2) in general does not make sense, for the operators $\vec{A}(t)$ and $\vec{E}(t)$ act in different spaces.

(3) The operator $\vec{A}(t)$ should describe all physical processes which influence upon the change $d\vec{\theta}/dt$ of the physical parameter we are interested in.

This latter assertion deserves a more detailes explanation. Suppose the time change $d\vec{\theta}/dt$ is a consequence of a finite family $P_1, \ldots, P_K$ of physical processes. For example, let $\vec{\theta}(t) = \vec{p}(t)$ is the impulse-vector of a particle, P_1 - braking process, P_2 - process of multiple scattering, P_3 - process of large-angle scattering, P_4 - process of conservation laws, P_5 - process of influence of magnetic field, P_6 - measurement noise process.

At present, we imagine all those processes as deterministic. From a geometrical point of view, the processes P_1 and P_2 act essentially in mutually orthogonal subspaces of $\mathcal{P}$, because the process P_1 works essentially in the direction of $-\vec{n}(t) = -\vec{p}(t)/p(t)$ at each $t \in [0, T]$, while P_2 acts in the hyperplane orthogonal to that vector. On the other hand, we shall see that the processes P_1 and P_2 will be related through P_5 which, geometrically, acts in another subspace.

Hence, the operator $\vec{A}(t)$ becomes a function of not only the "pure" separate processes $P_1, \ldots, P_K$, but also of processes $Q_{i_1 i_2}$, $Q_{i_1 i_2 i_3}, \ldots, (i_1, i_2 \in \{1, \ldots, K\})$, which describe interactions among pure processes. In most situations, pairwise interactions result in a sufficiently complete picture, i.e., we may assume that

$$\vec{A}(t) = \vec{F}(t, P_i, Q_{ij}; 1 \leq i, j \leq K). \tag{2.3}$$

Of course, the criteria for evaluating which processes P_i and pair processes Q_{ij} are negligible, must arise from the information processing problem itself. Indeed, there may exist processes extremely important from the physical point of view, but contributing almost nothing to the precision of the result we seek for. The point here is that it is not at all justifiable to exclude some of those processes in case we are not able to evaluate the loss of information caused by this exclusion.

All this shows that the role of modelling has been underestimated. As a result, we often have to deal with a high amount of data without being able to specify in more detail their physical origin, i.e., we are forced to use a kind of a "black-box" approach. This does not seem reasonable for we usually have a good deal of additional information at our disposal.

3. GEOMETRICAL REPRESENTATION AS A MODELLING DEVICE

As we shall see in subsequent papers, the complexity of the construction of $\vec{A}(t)$ (cf. (2.3)) depends sensitively upon the choice of a geometric representation of the dynamics (2.1). Furthermore, not all interactions are recognized in all representations (in particular, this concerns the processes P_2 and P_5 mentioned above). In fact, it was this latter fact that motivated the need for a more thorough investigation on the geometrical picture.

The basic idea is very simple. Since

$$\vec{\theta}(t) = \theta(t)\,\vec{n}(t), \tag{3.1}$$

formal differentiation yields

$$\frac{d\vec{\theta}}{dt} = \frac{d\theta}{dt}\vec{n}(t) + \frac{d\vec{n}}{dt}\theta(t). \tag{3.2}$$

Hence (2.1) assumes on the form $\frac{d\theta}{dt}\vec{n}(t) + \frac{d\vec{n}}{dt}\theta(t) = \vec{A}(t)$, i.e.,

$$\frac{d\vec{n}}{dt} = \frac{1}{\theta(t)}\vec{A}(t) - \frac{1}{\theta(t)}\frac{d\theta}{dt}\vec{n}(t). \tag{3.3}$$

As the unique natural geometric representation of the dynamics is the time change of the direction of $\vec{\theta}$, using an appropriate normalization we may suppose that $d\vec{n}/dt$ is a unit vector, too. This results in a nonlinear algebraic equation

$$\vec{\eta}(t) = N(t)\left[\frac{1}{\theta(t)}\vec{A}(t) - \frac{1}{\theta(t)}\frac{d\theta}{dt}\vec{n}(t)\right], \tag{3.4}$$

where $N(t) = |d\vec{n}/dt|^{-1}$.

Consequently, the geometrical representation of (2.1) is a walk on the surface of the three-dimensional sphere. Of course, we did not touch technical questions related with such a representation, and we will devote one of the subsequent papers to this.

As we observed (see (2) in Section 2), the measurements are carried over in a different, "geometrical", space K. Thus, we have to obtain an analogous representation for the geometry of the measurement process as well. Let us explain this in more detail.

Fix an orthogonal coordinate system (0xyz) so that the z-axis coincides with the time-axis (e.g., this is the case in the experiment described in Ref. /2/). Let $(\vec{\theta}(t); 0 \leq t \leq T)$ be an integral curve of (2.1). Suppose it is in a "general position" relative to (0xyz). For any $t \in [0,T]$ we may imagine the process P_2 as acting in the hyperplane orthogonal to $\vec{n}(t)$, and the measurement process P_6 in the hyperplane $(t = \text{const}) \equiv (z = \text{const})$. Our assumption entails that these two hyperplanes do not coincide. Consequently, we cannot simply add the two processes but we must first perform a transformation (depending on t) of one process before we may add them. Problems of this kind are usually treated neither in statistical literature nor in the literature devoted to information processing.

4. RANDOM FACTORS

Suppose now the dynamics is described by (2.1), the operator $\vec{A}(t)$ by (2.3), and the processes P_i and Q_{ij} are random ones. In the most general setup this says that there exist functional spaces M_i, $M_{ij} \subset \mathcal{P}^{[0,T]}$, σ-fields $\mathfrak{M}_i$, $\mathfrak{M}_{ij}$, and corresponding probability measures μ_i, μ_{ij} $(1 \leq i, j \leq K)$. Having a finite family of probability spaces as above we may construct a common probability space $(M, \mathfrak{M}, \mu)$ so that each process is defined on $(M, \mathfrak{M}, \mu)$ in a canonical manner /7/. However, we do not go into details of that construction, for we have not enough information at our disposal in order we can construct the above-mentioned probability spaces. Furthermore, since there exist, in general, infinitely many probability spaces $(M, \mathfrak{M}, \mu)$ which can serve the purpose of canonical representation, this idea would include in our consideration an element lacking both a geometrical and a physical interpretation.

Therefore we propose another way, also related to the geometric representation. Indeed, that representation enables us to separate the physical and the geometrical constituents of the actual states of the system (2.1). Based on microscopic theory of processes involved we shall formulate a phenomenological model of random factors. In analogy to the deterministic case[8], we shall construct the "stochastic" flow $(\tilde{T}_t ; t \in [0, T])$, where

$$\tilde{T}_t(p, g) = (p(t), g(t)) \tag{4.1}$$

if $p(0) = p$, $g(0) = g$, and $t \to (p(t), g(t))$ is the trajectory of the stochastic equation of motion (2.1); with p indicating the physical, and g indicating the geometrical parts of the state vectors. The construction will proceed via an operator representation of all stochastic forces.

On the other hand, the following natural question appears. Suppose we have been able to determine the flow $\tilde{T} = (\tilde{T}_t ; t \in [0, T])$ (i.e., we have been able to determine at least its second order properties). We can include also the measurement noise process into $\tilde{T}$ (e.g., modelling it in terms of a white noise process as the majority of experiments we are interested in are made as independent position measurements). As a result, we shall get a continuous-time description of the entire family of random factors. However, all measuring devices are constructed so that only measurements at a certain discrete finite set of instants is possible. Therefore, one may ask whether our approach enjoys advantages when compared with the usual approaches[3,4,8], which directly take into account the discrete nature of experimental data. This question is particularly vital when we realize known theoretical results on estimation problems for random functions[9] which, roughly speaking, assert that the optimum estimates obtained from the entire continuous-time sample paths are dominated by those ones obtained from certain discrete set of data lying on these paths.

Here our arguments go as follows. For "old" experiments, we can integrate in between the discrete observations. Since we did not make any simplifying assumption, the resulting probabilistic picture should be at least as adequate as that one obtained from the usual approaches. For "new" experiments our approach will result in an optimum design (in particular, the optimum number and location of discrete points of observation - similar considerations have been made also previously but only at a heuristic level without a sufficiently rigorous mathematical background).

5. CONCLUSIONS

The basic aim of the present paper was to explain several ideas on which a new approach to information processing in experimental physics is based. This approach will be developed in subsequent papers.

It is clear that even within the proposed new approach certain simplifying assumptions will become unavoidable. Nevertheless, unlike the usual approaches, our approach will allow for estimation of errors resulting from such assumptions.

The latter fact seems to us to be of primary importance. In fact, in majority of contemporary experiments the result is an estimate of an unknown parameter and determination of some of its statistical properties, e.g., its dispersion. But it does not follow automatically that the dispersion calculated is close to the theoretically optimum one (the problem of efficiency). As a consequence of more or less uncontrolable simplifications, the problem of efficiency has not yet been solved in a mathematically rigorous manner. These facts, according to our opinion, sufficiently well justify the forthcoming series of papers.

ACKNOWLEDGEMENT

The authors would like to express their sincere thanks to Prof. N.N.Govorun for interest in the problems considered and for valuable discussions.

REFERENCES

1. Moroz V.I. JINR, 10-4112, Dubna, 1968.
2. Zupančič C. CERN/EP/NA-4, Note 81-8.
3. Soop K. Bubble Chamber. Measurements and Data Processing. (Russian). "Nauka", M., 1970.
4. Emelyanenko G.A. JINR, 11-6076, Dubna, 1971.
5. Pascaud C., Morellet P. L.A.L. Rapport No.1227, Orsay, 1970.
6. Zhigunov V.P., Sokolov S.N. Sov.J.Part.Nucl., 1982, vyp.5, p.1024-1069.
7. Gichman I.I., Skorochod A.V. Theory of Random Processes I. (Russian). "Nauka", M., 1971.
8. Nitecki Z. Differentiable Dynamics. MIT Press. Cambridge and London, 1971.
9. Vilenkin S.Ya. Statistical Evaluation of Results of Observation of Random Functions. (Russian). "Energia", M., 1979.

Received by Publishing Department
on October 5, 1983.

Емельяненко Г.А., Шуян Ш. E5-83-693

Новый взгляд на обработку информации в экспериментальной физике. Общая стратегия

Излагаются основы нового подхода к решению теоретических и прикладных проблем обработки информации для широкого класса физических экспериментов. Особое внимание уделяется выбору геометрического пространства, а также вопросам согласованного описания взаимодействия между физическими процессами, которые обусловливают возникновение измеряемых пространственных структур. Отдельные, более технические проблемы будут изложены в последующих публикациях, к циклу которых настоящая работа является идейным введением.

Работа выполнена в Лаборатории вычислительной техники и автоматизации ОИЯИ.

Препринт Объединенного института ядерных исследований. Дубна 1983

Emelyanenko G.A., Sujan S. E5-83-693

A New Look upon Information Processing in Experimental Physics. General Strategy

The fundamental ideas of a new approach towards the solution of theoretical and applied problems of information processing for a large class of physical experiments are explained. A particular attention is paid to the choice of the geometric space as well as to the problem of a consistent description of interrelations among physical processes which influence upon measured spatial structures. The paper is considered as an introduction to a series of publications devoted to separate, more technical questions.

The investigation has been performed at the Laboratory of Computing Techniques and Automation, JINR.

Preprint of the Joint Institute for Nuclear Research. Dubna 1983