Automatica, Vol. 22, No. 6, pp. 657–674, 1986
Printed in Great Britain.

0005-1098/86 $3.00 + 0.00
Pergamon Journals Ltd.

A New Theoretical and Algorithmical Basis for Estimation, Identification and Control*

P. KOVANIC†

A non-statistical (gnostical) theory may be used to develop algorithms for robust estimation, robust identification and robust control suitable for treating small samples of inexact data, the statistical model of which is unknown.

Key Words—Cognitive systems; computer software; data processing; estimation; filtering; non-linear filtering; robust control; robust identification; (gnostics).

Abstract—A new theory applicable to data treatment is briefly exposed. This (gnostical) theory derives a mathematical model of data disturbed by uncertainty, the statistical model of which may be unknown or even unjustifiable. Gnostical theory is based on two simple axioms. It results in laws governing the uncertainty of each individual datum such as variational principles of virtual kinematics of real data and of their dynamics, closely related to entropy and information of data. Algorithms resulting from gnostical theory maximize the information obtained from data and yield data characteristics robust with respect to outlying or inlying data. Fields of application include the estimation of both location and scale parameters of small data samples and of their generalized correlations, robust estimation of probability and of probability distribution, non-linear discrete filtering, prediction and smoothing, identification of systems under strong disturbances and adaptive setting of alarm systems, robust identification of regression models, robust control systems etc. The main advantage of the new approach is that it leads to algorithms efficient even in applications to small samples of bad data.

1. INTRODUCTION

CONTROL necessarily includes estimation of real quantities and identification of a system. The better the cognition, the better is the control. But real cognitive processes are unavoidably disturbed by uncertainties of various origins and natures. Practical situations are ordinarily neither stationary nor ergodic and a full statistical description of the uncertainties is rarely possible. Available data samples are often "small", i.e. their size and/or quality are not sufficient to support a hypothesis on a statistical model reliably. The idea of the gnostical approach exposed here is to base the treatment of small data samples on a theory of individual uncertainty of a datum instead of using a statistical approach assuming some large collections of random events. A foundation of a theory of individual uncertainty can, of course, be expected to differ from that of statistics in substance.

Each theory of uncertain data should answer two fundamental questions.

(1) How should the amount of uncertainty of an individual datum be evaluated?

(2) How should the data of a sample be composed to suppress influence of their individual uncertainties on a characteristic of the whole sample?

The former problem is closely related to the problem of defining metrics on spaces. If a datum has a value z instead of a true, ideal value z_0 then the effect of uncertainty is usually evaluated by the difference $z - z_0$. The quantity $|z - z_0|$ characterizing the error should represent the distance of two points z and z_1 of one-dimensional real variety S_1. But such a formula is true only under the assumption that the metric on the variety S_1 is of the Euclidean type. *Should* the variety of uncertain data be a Euclidean space? If not, then a more general formula for the distance should be used which would be valid for a more general type of a metric space of the Riemannian type. A one-dimensional Riemannian metric space is defined by the variety S_1 on which a positive weight $g(z)$ is given such that an element dl of the distance equals $dl = g(z)dz$. The weight $g(z)$ equals identically 1 only in the Euclidean case. In other geometries it depends on the point z of the real variety S_1. Thus the evaluations of individual uncertainty can be interpreted as a geometrical problem. In gnostical theory the metric enabling one to measure individual uncertainty is fully determined as a consequence of

* Received 26 November 1984; revised 11 December 1985; revised 3 April 1986; revised 18 June 1986. The original version of this paper was presented at the 9th IFAC Congress on A Bridge Between Control Science and Technology which was held in Budapest, Hungary during July 1984.

The published proceedings of this IFAC meeting may be ordered from Pergamon Books Limited, Headington Hill Hall, Oxford OX3 0WB, England.

This paper was recommended for publication in revised form by Associate Editor T. Basar under the direction of Editor H. Kwakernaak.

† Institute of Information Theory and Automation, Czechoslovak Academy of Sciences, Pod vodárenskou věží 4,18208 Prague, Czechoslovakia.

the first axiom of the theory although this axiom seems to have nothing common with the geometry on first sight. It expresses some evident metrological facts.

The problem of data composition is also to be interpreted as a more complicated one than in the classical approach; for example, in evaluating the arithmetic mean the linear composition law is used, whereby bad data obtain the same weight as good ones. Ordinary estimates of variances and covariances assume a quadratic composition law as the proper one, giving individual outliers a strong influence on the result. But why should the same composition laws as those of classical mechanics related to the centre of mass and the inertial momentum of a system of a mass points be acceptable? The coherence of statistics with Newtonian mechanics based on Euclidean geometry was certainly important two centuries ago but present scientific paradigm is neither Euclidean nor Newtonian. In gnostical theory the data composition law is given by the second axiom based, among others, on the requirement of coherence of the gnostical theory of data with the relativistic mechanics.

As shown in this paper, the two gnostical axioms mentioned and their consequences are sufficient to develop algorithms for the treatment of small samples of real data under a strong influence of uncertainty which may admit no statistical model. Only data should be used, the idea is: "Let data speak for themselves!"

A first exposure of the gnostical theory has been presented in Kovanic (1984a–c). The aim of the present paper is to attract the attention of readers feeling the need for an efficient complement to existing models of uncertainty which would be theoretically founded and practically applicable. The reader interested only in practical aspects, i.e. algorithms and applications may find it desirable to move on immediately to Section 5.

Some references are in order here. As already mentioned, there exists a deep connection between the gnostical theory of uncertain data and relativistic physics. The uncertain events considered by the gnostical theory have a general nature. However, a general theory of uncertainy should also hold in a simple particular case of the uncertainty related to the physical motion. The correspondence of such seemingly different theories is thus not surprising; it is a manifestation of Bohr's principle of correspondence. The idea on relations between relativistic physics and cybernetics appeared probably for the first time in a paper by Jumarie (1975) and was treated further in a series of his papers and books on relativistic cybernetics (e.g. Jumarie, 1985). Jumarie's idea can perhaps be formulated as a necessity to take into account the subjectivity of the observer in all cybernetical considerations. Jumarie uses theoretical arguments of a very general nature to support his point of view. They differ from those of the gnostical theory substantially. Nevertheless, results of the gnostical theory also support the idea that cybernetics should be relativistic. In gnostical theory, a non-Euclidean character of metric results in different weights of data having different errors. Each weight depends on the uncertainty of the particular datum in a way which is reminiscent of the dependence of the mass of a relativistic particle on its relative velocity.

Individual weights of individual data derived from statistical assumptions also appear in the framework of robust statistical theory. The gnostical approach is quite different but its practical goal is the same: to treat data. Both approaches may thus be compared by practical results. An example of such a comparison is given in Section 6.

2. GNOSTICAL KINEMATICS OF AN INDIVIDUAL DATUM

Kinematics deals with both actual and virtual motion of objects. The object is a datum. Its value differs from the result of an ideal qualification because of influence of uncertainty. Changes in its value caused by the uncertainty may be thought of as being virtual motion along a path. Behaviour of errors of a datum can thus be studied by methods similar to those of kinematics, but it is necessary to have a realistic and sufficiently general starting point. Such a point is the model of data.

2.1. *First gnostical axiom—the model of data*

One usually considers an additive model

$$v_i = v_0 + s\Omega_i \quad v_0, s, \Omega_i \in R_1, i = 1, \ldots, n, \quad (1)$$

where v_0 is an undisturbed value of the variable v, s is a *scale parameter* and Ω_i is a numerical characteristic of the influence of the ith particular uncertainty. The scale parameter is constant for all observations of the series. It characterizes their variability.

However, of great importance in this exposition is a multiplicative version of model (1), easily obtained by exponentiating

$$z_i = z_0 e^{s\Omega_i}, \quad z_i, z_0 \in R_+, \quad (2)$$

where $z_i = e^{v_i}$, $z_0 = e^{v_0}$ and R_+ denotes the interval of positive finite real numbers. If not specified otherwise, the notion "data" will be used for the multiplicative model (2), which will now be considered in more detail.

A quantitative cognition process includes two phases: the quantification and the estimation. The *quantification* is a mapping of a structure of empirical quantities into an interval of real numbers. Results of quantification are *data*. Only two quantification procedures will be considered here, *measurements* of real quantities and *countings* of real objects having a defined quality. An ideal quantification would result in a precise datum z_0 (called the ideal value). As a rule, a real quantification process involves uncertainty. A datum really produced in an ith quantification procedure will be denoted by z_i. The *estimation* is a mapping of a collection of data into an interval of real numbers. It should produce an estimate of the (unknown) ideal quantity z_0. To develop a mathematical model of quantification the data model (2) will be used, but to stress its fundamental features it is generalized as Axiom 1 given below.

Axiom 1 of the gnostical theory. Let z be either an actual datum already produced by quantification or an arbitrary possible result of quantification. Let z_0 be the ideal value of this datum. Let ζ be a parameter characterizing the influence of the uncertainty on this datum. Then

$$z = z_0\zeta \quad z_0, \zeta \in R_+, \tag{3}$$

where possible values of ζ cover the whole continuum R_+.

This data model has its own motivation independent of (1). To obtain a realistic theory one must start with realistic axioms. The realism of Axiom 1 is in its correspondence to the nature of data. They are not just any real numbers but results of a practical activity of men and of their manipulations with real objects, measuring tools, devices and apparatuses. They are products of quantification, which is a special technology having its own theory—metrology (Krantz *et al.*, 1971). The theory of measurement has its axiomatics reflecting the fundamental assumptions of real quantifying procedures. A result of a primary measurement says how many times a measured quantity exceeds a measuring unit or vice versa. Hence, the result of such a measurement is positive and finite. The same holds for counting.

Metrological axiomatics also includes a requirement of linearity of the mapping of empirical structures into numerical structures, but metrology deals only with precise quantification, leaving the problem of uncertainty of quantification to post-processing. The author wishes to include the uncertainty in the quantification stage, extending the one-dimensional linear metrological model by another variable ζ characterizing the effect of uncertainty. However, this quantity is often an object of quantification too. The roles of variables z_0 and ζ should thus be interchangable, symmetrical. This is how one arrives at the bilinear function (3) of a pair of finite positive factors.

The metrological requirement of positiveness of data is related only to results of primary quantifying operations. There exist secondary manipulations with data such as shifting of the origin of the measurements and so on. Some processes involve additive disturbances and additive transformations of data. Sometimes the result of data treatment is required to be translation equivariant (the case of a location parameter) or translation invariant (the case of the scale parameter). The additive data model (1) is suitable in such situations. All formulae of the gnostical theory as well as gnostical algorithms are applicable to multiplicative data (3) directly and to additive data (1) after their exponentiation.

2.2. *Virtual kinematics of quantification*

A datum z_i may be interpreted as a result of virtual motion caused by a continuous change of the factor ζ from 1 to its final value ζ_i. Features of the path passed by the datum ("kinematics" of the datum) may be derived from Axiom 1. Using the form (2) of this axiom and the identity

$$e^{\Omega} = \cosh\Omega + \sinh\Omega \tag{4}$$

a Cartesian coordinate system is introduced,

$$x := z_0^{1/s}\cosh(\Omega) \quad y := z_0^{1/s}\sinh(\Omega). \tag{5}$$

Using the matrix notation

$$\mathbf{u} := \begin{bmatrix} x \\ y \end{bmatrix} \qquad {}_c\mathbf{u} := \begin{bmatrix} y \\ x \end{bmatrix}$$
$$\mathbf{u}_0 := \begin{bmatrix} z_0^{1/s} \\ 0 \end{bmatrix} \qquad {}_c\mathbf{u}_0 := \begin{bmatrix} 0 \\ z_0^{1/s} \end{bmatrix} \tag{6}$$

$$\mathbf{K}_q(\Omega) \equiv \mathbf{K}_q := \begin{bmatrix} \cosh\Omega & \sinh\Omega \\ \sinh\Omega & \cosh\Omega \end{bmatrix} \tag{7}$$

gives

$$\mathbf{u} = \mathbf{K}_q(\Omega)\mathbf{u}_0 \quad {}_c\mathbf{u} = \mathbf{K}_q(\Omega)_c\mathbf{u}_0. \tag{8}$$

The matrix $\mathbf{K}_q$ will be called the *quantifying channel*, the vectors $\mathbf{u}$ and ${}_c\mathbf{u}$ will be gnostical *events* of the first and second kind, respectively. Vectors $\mathbf{u}_0$ and ${}_c\mathbf{u}_0$ are *ideal* events. If $\mathbf{u} = \mathbf{K}_q(\Omega')\mathbf{u}_0$ and $\mathbf{u}'' = \mathbf{K}_q(\Omega'')\mathbf{u}_0$ are two gnostical events then the transformations

$$\mathbf{u}'' = \mathbf{K}_q(\Omega)\mathbf{u}' \quad {}_c\mathbf{u}'' = \mathbf{K}_q(\Omega)_c\mathbf{u}' \tag{9}$$

take place, where

$$\Omega = \Omega'' - \Omega' \tag{10}$$

for parameters of the channels. Thus quantifying channels transform, in a general case, arbitrary gnostical events of both kinds into other ones of the same kind in the same way (they are "symmetrical"). Parameter Ω of a channel characterizes the relation between the two events due to the effect of uncertainty. An ordinary matrix product of two channels $\mathbf{K}_q(\Omega')$ and $\mathbf{K}_q(\Omega'')$ has the parameter equal to the sum $\Omega' + \Omega''$ of parameters of both factors. A set G_q of all possible quantifying channels is easily shown to be a commutative group with respect to multiplication of channels which corresponds to a "serial connection" of these channels. The group G_q is an isomorphic map of the additive group of parameters Ω. This parameter is an independent variable of the quantifying process. It characterizes a relation between two gnostical events. The group G_q is a mathematical model of gnostical "motion" caused by uncertainty, which may increase as well as decrease the data. This is a good feature leading to the possibility of mutual compensation of different gnostical motions. It will be important in the sequel that the group G_q is an isomorphic map of the group of Lorentz transformations. Yet another interesting feature of all quantifying channels may be seen in (7):

$$\mathrm{Det}\{\mathbf{K}_q\} = 1. \tag{11}$$

Channels are thus "uniformly regular".

Let the events of both kinds be combined into a matrix $\mathrm{U} = [\mathbf{u}, {}_c\mathbf{u}]$. Then (9) becomes

$$\mathrm{U}'' = \mathbf{K}_q(\Omega)\mathrm{U}' \tag{12}$$

and

$$\mathrm{Det}\{\mathrm{U}''\} = \mathrm{Det}\{\mathrm{U}'\} = z_0^{2/s} \tag{13}$$

because of (11). This says that the ideal, true value z_0 is an invariant of the group of quantifying motions. In spite of the "falsifying" influence of uncertainty the "truth" is thus present in each gnostical event.

2.3. *Metric of quantification and estimation*

Nothing has yet been said on the choice of a metric of the space occupied by gnostical events. What is known is only that it is a Cartesian product $R_+ \times R_1$ with coordinates x, y transformed by the group G_q of channels $\mathbf{K}_q$. Could a metric, a scalar product to measure distances between two gnostical events, be chosen freely? The answer is negative.

Many mathematicians believe that they have freedom in the choice of a metric for their mathematical model. This is true for pure mathematics but it may not be true for mathematical modelling of some real processes. A long time ago, Albert Einstein had already acknowledged Riemann's idea on the existence of ties between metric tensors of spaces for mathematical descriptions of reality on the one hand and such facts of real life as real forces on the other. Both of Einstein's theories of relativity corroborated Riemann's daring hypothesis for mechanics, electrodynamics and optics decades after his death. The physical facts implying the metric in special relativity are the invariance of the speed of light and momentum conservation. In general relativity theory, such facts are the equivalency of inertial and gravitational effects of mass and the distribution of mass in space.

In this case the governing facts have been expressed by Axiom 1. As proved in Kovanic (1984a) there exists exactly one metric on the space of gnostical events invariant under the group G_q. It is the Minkowskian metric for which a scalar product is

$$s_{12} = \mathbf{u}_1^{\mathrm{T}}\mathbf{g}_{\mathrm{M}}\mathbf{u}_2, \tag{14}$$

where

$$\mathbf{g}_M = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{15}$$

is the matrix representation of the Minkowskian metric tensor. The scalar product of two events of the second kind obviously differs from (14) only by sign.

The proof of this important statement is simple. A metric is invariant with respect to transformations performed by $\mathbf{K}_q \in G_q$ iff the equation $\mathbf{u}_1^{\mathrm{T}}\mathbf{g}\mathbf{u}_2 = \mathbf{u}_1^{\mathrm{T}}\mathbf{K}_q\mathbf{g}\mathbf{K}_q\mathbf{u}_2$ holds identically for a symmetric matrix $\mathbf{g}$ and for all events $\mathbf{u}$ and all channels $\mathbf{K}_q$. It should thus hold: $\mathbf{g} \equiv \mathbf{K}_q\mathbf{g}\mathbf{K}_q$ for all $\mathbf{K}_q \in G_q$. This identity is satisfied only by matrices having the form $c\mathbf{g}_{\mathrm{M}}$ where $\mathbf{g}_{\mathrm{M}}$ is (15) and c is a non-zero constant. But a constant multiplier of a metric tensor does not produce a new metric. Therefore $c = 1$ may be chosen.

The quantity $\mathrm{Det}\{\mathrm{U}\} = z_0^{2/s}$ is thus a square of the length of the event $\mathbf{u}$. It remains invariant under transformation. The channels $\mathbf{K}_q$ represent (Minkowskian) orthogonal rotation of events $\mathbf{u}$ (or ${}_c\mathbf{u}$) without a change of their (Minkowskian) length equalling

$$z_0^{1/s} = \sqrt{x^2 - y^2} \tag{16}$$

in the case of events of the first kind and $\sqrt{-1}\, z_0^{1/s}$ in the case of events of the second kind.

When a metric tensor of the space is known, it is possible to study geodesic lines. Two families of geodesics can be proved to exist for the particular case under consideration: straight lines and Minkowskian circles. As known, a distance between two points on a geodesic is extremal when measured along the geodesic. The gnostical quantifying "motion" along a path corresponding to a quantifying channel (along a Minkowskian circle) should therefore satisfy a variational principle. It can indeed be shown that the modulus $|\Omega|$ of the quantifying "time" Ω is a relative length of such a path and that it is the maximal possible relative length of this path. The variational principle is thus the maximization of the relative distance of data from the ideal value. Thus Nature "aims" to falsify the data by uncertainty in the most effective way. It is done by means of symmetrical (9) and uniformly regular (11) channels.

To "move" in the "gnostical game" against Nature one chooses an estimating path which would minimize the damage caused to data by quantifying uncertainty. As shown in Kovanic (1984a) there exists exactly one *estimating channel* $\mathbf{K}_e(\omega)$ corresponding to each quantifying channel $\mathbf{K}_q(\Omega)$ (where $\tan\omega = -\tanh\Omega$), which is symmetric and uniformly regular. Its explicit form is

$$\mathbf{K}_e(\omega) = \begin{bmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{bmatrix} \quad (\text{where } \tan\omega = -\tanh\Omega). \tag{17}$$

The estimating transformation is thus

$$\mathbf{u}^* = \mathbf{K}_e(\omega)\mathbf{u}. \tag{18}$$

A set of all possible matrices $\mathbf{K}_e$ is also a commutative group G_e with respect to matrix multiplication. It represents a group of estimating "motions", the estimating independent variable being the parameter ω. The unique metric invariant under the group G_e is the Euclidean one. The invariant of the group is the quantity

$$r := \sqrt{x^2 + y^2}, \tag{19}$$

the (Euclidean) radius of a (Euclidean) circle. Estimating motion is thus a (Euclidean) orthogonal rotation of gnostical events. The relative length of the path along the circle is $|\omega|$. It is the minimum of lengths of varying paths. Such a choice of estimating channel makes estimation dual to quantification.

2.4. *The ideal gnostical cycle*

A third type of gnostical channel can be introduced to close a cycle of these transformations: the attenuating/amplifying channel, which has the form of a matrix

$$\mathbf{K}_a := \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}, \quad k \in R_+. \tag{20}$$

The set G_a of all possible channels $\mathbf{K}_a$ is also a commutative group. This group represents a generalized real motion because it is connected with variations of the ideal quantity z_0. Both gnostical motions are orthogonal to the generalized real motion with respect to their "own" metric. The three types of motions are summarized in Table 1.

There are three mutually complementary motions. They may be used to create a closed gnostical cycle—*the ideal gnostical cycle* $(\mathrm{IGC})_i$ defined by a real datum z_i ($z_i \neq z_0$) which is a triple of continuous segments of lines successively interconnecting three points of the variety of gnostical events, namely

—a segment of Minkowskian circle from $(z_0^{1/2}, 0)$ to $(z_0^{1/s}\mathrm{ch}\,\Omega_i,\ z_0^{1/s}\mathrm{sh}\,\Omega_i)$

—a segment of Euclidean circle from $(z_0^{1/s}\mathrm{ch}\,\Omega_i,\ z_0^{1/s}\mathrm{sh}\,\Omega_i) = (r_i\cos\omega_i,\ -r_i\sin\omega_i)$ to $(r_i, 0)$

—a segment of straight line from $(r_i, 0)$ to $(z_0^{1/s}, 0)$.

A complementary ${}_c(\mathrm{IGC})_i$ exists to $(\mathrm{IGC})_i$ passing through the points corresponding to ${}_c\mathbf{u}_i$ instead of $\mathbf{u}_i$. In both cases $r_i^2 = z_0^{2/s}\mathrm{ch}2\,\Omega_i = x_i^2 + y_i^2$ and

$$z_0^{1/s}\mathrm{ch}\,\Omega_i = r_i\cos\omega_i \quad z_0^{1/s}\mathrm{sh}\,\Omega_i = -r_i\sin\omega_i. \tag{21}$$

The "idealness" of the IGC has two aspects.

(1) In practice, the exact value of z_0 is not known, therefore the cycle can never be closed.

(2) The IGC is the best of possible gnostical cycles passing through the same "edge" points. To show the sense of this statement it is necessary to proceed to dynamics.

3. GNOSTICAL DYNAMICS OF AN INDIVIDUAL DATUM

3.1. *Dissimilarity of gnostical events and its characteristics*

Squares of channels

$$\mathbf{K}_q^2(\Omega) = \mathbf{K}_q(2\Omega), \quad \mathbf{K}_e^2(\omega) = \mathbf{K}_e(2\omega) \tag{22}$$

can be shown to be transformed as tensors when gnostical events are transformed by gnostical channels. To understand the role of these tensors the concept of dissimilarity of events is introduced. Two events $\mathbf{u}'$ and $\mathbf{u}$ are said to be σ, ε—similar if

TABLE 1. THREE GNOSTICAL GROUPS

Symbol	Group Motion	Independent variable	Channel	Invariant	Time dependent variables
G_q	Quantification	Ω	$\mathbf{K}_q$	$\sqrt{x^2 y^2} \equiv z_0^{1/\beta}$	$\sqrt{x^2+y^2}, y/x$
G_e	Estimation	ω	$\mathbf{K}_e$	$\sqrt{x^2+y^2} \equiv r$	$\sqrt{x^2-y^2}, y/x$
G_a	Attenuation/ amplification	z_0, e	$\mathbf{K}_a$	$y/x = \tanh\Omega = \tan\omega$	$\sqrt{y^2+y^2}$, $\sqrt{x^2-y^2}$

TABLE 2. MEASURES OF DISSIMILARITY OF A COUPLE OF EVENTS u' AND u

Type of similarity σ	ε	Measure of dissimilarity: Minkowskian	Euclidean
+1	+1	$\sinh(\Omega' - \Omega)$	$\sin(\omega' - \omega)$
+1	−1	$\cosh(\Omega' - \Omega)$	$\cos(\omega' + \omega)$
−1	+1	$\sinh(\Omega' + \Omega)$	$\sin(\omega' + \omega)$
−1	−1	$\cosh(\Omega' + \Omega)$	$\cos(\omega' - \omega)$

$$\sigma\frac{x'}{y'} = \left|\frac{x}{y}\right|^{\varepsilon} \quad \text{where } \sigma = \pm 1, \varepsilon = \pm 1. \tag{23a}$$

This condition is obviously equivalent to

$$\sigma x'|y|^{\varepsilon} = y'|x|^{\varepsilon}. \tag{23b}$$

If a couple of events does not satisfy (23b) then the events are dissimilar to an extent which is measurable by the difference of both sides of (23b) or by a quantity proportional to this difference. Using (23b) and (21) normalized measures of dissimilarity of events are thus obtained, as summarized in Table 2.

Thus the products of channels $\mathbf{K}_q(\Omega)\mathbf{K}_q(\Omega') = \mathbf{K}_q(\Omega + \Omega')$ and $\mathbf{K}_e(\omega)\mathbf{K}_e(\omega')$, which are members of the groups G_q and G_e, as are the channels, characterize the dissimilarity of events. The special cases $\mathbf{K}_q^2$ and $\mathbf{K}_e^2$ obtained for $\Omega = \Omega'$ (or $\omega = \omega'$) thus characterize the dissimilarity of an event to itself, "via the ideal value". They can be written in the form of matrices

$$\mathbf{K}_q^2 = \begin{bmatrix} \frac{1}{f} & h_q \\ h_q & \frac{1}{f} \end{bmatrix} \qquad \mathbf{K}_e^2 = \begin{bmatrix} f & -h_e \\ h_e & f \end{bmatrix} \tag{24}$$

using the following notation.

Fidelity:

$$f: = \cos 2\omega = \frac{x^2 - y^2}{x^2 + y^2}. \tag{25}$$

Estimating irrelevance:

$$h_e: = \sin 2\omega = \frac{2xy}{x^2 + y^2} = -\tanh 2\Omega. \tag{26}$$

Quantifying irrelevance:

$$h_q: = \sinh 2\Omega = \frac{2xy}{x^2 - y^2} = -\tan 2\omega. \tag{27}$$

Infidelity:

$$\frac{1}{f} = \cosh 2\Omega. \tag{28}$$

These relations between the quantities can be derived from (7), (17) and (21) using formulae of standard functions.

Using Laplace's operator ∇^2 (which in Minkowskian metric has the form

$$\nabla_M^2(\cdot) = \frac{\partial^2(\cdot)}{\partial x^2} - \frac{\partial^2(\cdot)}{\partial y^2}$$

and in Euclidean metric

$$\nabla_E^2(\cdot) = \frac{\partial^2(\cdot)}{\partial x^2} + \frac{\partial^2(\cdot)}{\partial y^2}\Big)$$

gives the equations

$$\nabla_M^2\left(\frac{1}{f}\right) + \frac{4}{x^2 - y^2}\frac{1}{f} = 0$$

$$\nabla_E^2(f) + \frac{4}{x^2 + y^2}f = 0 \tag{29}$$

proving that both infidelity $1 - f$ and fidelity f (interpreted as scalar fields over the variety of gnostical events) diffuse from a point to another. The length of the diffusion flow vector of these quantities (of the vectors grad $(1/f)$ and grad(f)) can be shown to be proportional to irrelevance h_q and h_e, respectively. The gnostical tensors $\mathbf{K}_q^2$ and $\mathbf{K}_e^2$ can thus be interpreted alternatively as the tensors of "infidelity–infidelity flow" and of "fidelity–fidelity flow".

3.2. *Energy of a gnostical event*

Dynamics deals with energies and their changes. Can we find some energetical aspects with data? If a quantity q is measured in terms of electrical current, voltage or electrical pulses, then the electrical energy connected with this quantity is proportional to q^2. It may be measured without electrical transformations as well but the measuring process or the dependence of q on time or another quantity may be modelled using an analog computer. Then the energy of the electrical quantity modelling the q is again proportional to q^2. The coefficient of proportionality is a matter of scale, therefore q^2 can be taken to be the energy of the quantity q. For a vector quantity the energy may be attached to the square of its length. Squares $z_0^{2/s}$ and r^2 of lengths of gnostical events may thus be interpreted as their energies. Their changes within an ideal gnostical cycle can be shown (Kovanic, 1984c) to satisfy the first thermodynamical law: the overall change of the energy of a gnostical event within an IGC is zero.

The motion of energy of a datum will be used in Section 3.3 to show the relations between thermodynamical entropy of data and their geometrical features such as their mutual dissimilarity.

3.3. *Thermodynamical entropy of a gnostical event*

Fidelity has been introduced as one of the measures of uncertainty. As shown in Kovanic (1984c) it may also be interpreted as a linear function of thermodynamical entropy S defined by $\int T^{-1}\mathrm{d}Q$, where Q is an amount of heat and T the temperature at which the heat $\mathrm{d}Q$ is transferred into or from the system. The invariant energy z_0^2 of a quantifying channel may be thought to be converted into heat, the amount of which can be measured by temperature using a calorimeter. Thus a constant temperature may be attached to the quantifying channel proportional to z_0^2. The amount of heat transferred into the channel during quantification is proportional to the increase of energy of the event from z_0^2 to r^2. Analogical Gedanken experiments may be made with the estimating channel. The following relations thus hold between the changes S_q and S_e of thermodynamical entropy during quantification and estimation and the infidelity $1/f$ and fidelity f, respectively:

$$S_q = c\cdot\left(\frac{1}{f}-1\right) \quad S_e = c\cdot(f-1), \tag{30}$$

where c is a constant.

Neither the uncertainty of an event nor the entropy changes during an attenuating phase of the IGC. The overall change of entropy within an IGC therefore equals $S_q + S_e$. For a datum $(z_i \neq z_0)$

$$\oint_{(\mathrm{IGC})_i} \mathrm{d}S > 0 \tag{31a}$$

for the integrating path coinciding with $(\mathrm{IGC})_i$. This is an analogy of the second law of thermodynamics for non-invertible closed cycles: even in the case of an ideal estimation process an increase of entropy is unavoidable. Now consider a varying gnostical cycle $(\mathrm{VGC})_i$ passing through the same "edge" points as $(\mathrm{IGC})_i$, defined by a couple (z_0, z_i). Both the quantifying and estimating variational principles related to Ω and ω are already known. Using them together with (25) and (28) gives

$$\oint_{(\mathrm{VGC})_i} \mathrm{d}S \geq \oint_{(\mathrm{IGC})_i} \mathrm{d}S \tag{31b}$$

proving the *thermodynamical optimality* of the ideal gnostical cycle.

Equalities (30) and (29) offer a further interpretation of gnostical tensors $\mathbf{K}_q^2$ and $\mathbf{K}_e^2$ as *tensors of entropy–entropy flow.*

3.4. *"Probability" of an individual datum*

Both equations (29) hold for all points of the variety, however only the circles corresponding to the paths of events within channels for a fixed z_0 or fixed r will be considered. Points of these paths will be denoted x' and y'. The first terms $\nabla_M^2(\cdot)$ and $\nabla_E^2(\cdot)$ of (29) then represent the distribution of sources of both fields $1/f$ and f along the paths corresponding to channels. The same holds for entropy changes S_q and S_e due to (30). The second terms $4f^{-1}/(x^2-y^2)$ and $4f/(x^2+y^2)$ can be rewritten to obtain the form of sources of scalar fields I_q and I_e over the intervals of irrelevancies $-\infty < h_q < \infty$ and $-1 < h_e < 1$, respectively. The equivalences of sources of both kinds of scalar fields then take the form

$$[\nabla_M^2(S_q)]_{x',y'} = c_1\cdot\left[\frac{\mathrm{d}^2 I_e}{\mathrm{d}h_e^2}\right]_{x',y'}$$

$(x'^2 + y'^2$ and c_1 are constants) (32a)

$$[\nabla_E^2(S_e)]_{x',y'} = c_2\cdot\left[\frac{\mathrm{d}^2 I_q}{\mathrm{d}h_q^2}\right]_{x',y'}$$

$(x'^2 - y'^2$ and c_2 are constants). (32b)

The quantities I_q and I_e (interpretable as changes of information) can thus be defined by right sides of (29) and (32) by integrals

$$I_q := -\int_0^{h_q}\int_0^{\eta}\frac{(d\eta)^2}{1+\eta^2} \quad I_e := \int_0^{h_e}\int_0^{\eta}\frac{(d\eta)^2}{1-\eta^2} \quad (33)$$

or after integration—

$$I_q = H(1/2) - H(p_q) \quad I_e = H(1/2) - H(p_e), \quad (34)$$

where

$$H(p) = -p\ln p - (1-p)\ln(1-p) \quad (35)$$

and

$$p_q = \frac{1}{2}(1 + \sqrt{-1}\,h_q) \quad p_e = \frac{1}{2}(1 + h_e). \quad (36a)$$

Using (26) and (4) for a given datum $z = z_i$ the last quantity may be rewritten as

$$p_{ei} = 1/(1 + e^{4\Omega_i}) = 1/(1 + (z_i^4/z_0^4)^{1/s}). \quad (36b)$$

It is a monotonically increasing function of the unknown quantity z_0 taking values between 0 and 1. It can be interpreted as the distribution function of z_0 given z_i, as an analogy of the conditional probability $P(z_0 > z\,|\,z_i)$. The quantity $p_{ei}(z_{02}) - p_{ei}(z_{01})$ (where $z_{02} \geq z_{01} > 0$) induces a finite measure on Borel subsets of R_+ (given z_i). It makes possible evaluation of the confidence that the unknown, z_0, has a value from the interval $\{z_0 : z_{01} \leq z_0 < z_{02}\}$ when the value of the datum has been z_i. An estimate of the distribution function of the data sample can be obtained by application of the composition law to functions p_{ei} of all data, as shown below. Such an estimate will approach the empirical probability distribution function when the size n of the sample increases.

3.5. *Information of an individual datum*

The reasons for accepting the quantities I_q and I_e as changes of information of a datum within the quantification and estimation, respectively, are as follows.

(a) Both are monotonic real functions of uncertainty parameters Ω and ω, equalling zero for zero parameters. They may therefore be used for an evaluation of the amount of uncertainty.

(b) If p was a classical probability then (35) would be Boltzmann's statistical entropy. Consider a couple of mutually excluding messages having *a priori* and *a posteriori* probabilities (1/2, 1/2) and $(p, 1-p)$, respectively, then (34) would be the Shannon information. The parameter p_e may be interpreted as an analogy of probability $p(z_0 > z\,|\,z_i)$, also applicable without statistical model to an individual datum.

(c) Sources of the fields of I_q and I_e are proportional to sources of thermodynamical entropy according to (32). Equations (32) are thus *conversion laws* of information into entropy and vice versa for the quantification and estimation, respectively.

There are two important properties of the changes of information to be noted here.

(1) For each datum z differing from z_0, $I_q + I_e < 0$. Attenuation does not change the information. Therefore, the overall change of information within an IGC is *negative*. This is intuitively expected as a plausible counterpart of the second law of thermodynamics as well as of its gnostical version (31).

(2) Let $(\mathrm{IGC})_i$ be the closed path corresponding to the IGC defined by a datum z_i and its ideal value z_0. Let $(\mathrm{VGC})_i$ be a varied closed gnostical cycle passing through the same "edge" points like $(\mathrm{IGC})_i$. Then an analogue of the variational principle (31b) holds (Kovanic, 1984a)

$$\oint_{\mathrm{VGC}_i} dI \leq \oint_{\mathrm{IGC}_i} dI < 0 \quad (37)$$

stating the *informational optimality* of the IGC: it represents the path minimizing the (unavoidable) overall loss of information.

4. GNOSTICAL THEORY OF DATA SAMPLES

4.1. *Gnostical characteristics of a data sample*

A *data sample* is defined as an n-tuple

$$Z = Z(z_0, s, n) := \langle z_1, \ldots\ldots, z_n\rangle \; (1 < n < +\infty), \quad (38)$$

where all n data (2) have the same parameters s and z_0. Both of these parameters are to be estimated using the data sample. Formulae derived above for individual data hold for data from a sample too. To express all characteristics by means of data the ratio $(z_i/z_0)^{1/s}$ is substituted into (25)–(27) instead of e^{Ω_i} using the data model (2), giving

$$f_i = 2/(q_i^2 + q_i^{-2}) \quad (39)$$

$$h_{qi} = (q_i^2 - q_i^{-2})/2 \quad (40)$$

$$h_{ei} = (q_i^{-2} - q_i^2)/(q_i^2 + q_i^{-2}), \quad (41)$$

where

$$q_i = (z_i/z_0)^{1/s}. \qquad (42)$$

The infidelity is still f_i^{-1}, with f_i (39).

Each datum has been characterized by a gnostical event $\mathbf{u}$ or ${}_e\mathbf{u}$ to which two channels $\mathbf{K}_q$ and $\mathbf{K}_e$ correspond together with their squares, tensors of entropy–entropy flow. A data sample will be characterized in an analogous way by two *composite events* $\mathbf{u}_c$ and ${}_e\mathbf{u}_c$ having the same form (6) as the individual ones and being parameterized by Ω_c and ω_c. They should both be functions of data, the type of the functions being given by a composition law. Thus they will also be composite channels $\mathbf{K}_q(\Omega_c)$ and $\mathbf{K}_e(\omega_c)$ and composite tensors $\mathbf{K}_q^2(\Omega_c)$ and $\mathbf{K}_e^2(\omega_c)$ applicable as characteristics of a data sample. To obtain them the composition law is used.

4.2. *On a triplet of mappings*

A striking formal correspondence between the quantifying formulae of gnostical theory and the formulae of relativistic physics provides a realistic basis for a choice of data composition axiom. Consider a system S_r of relativistic events characterized by points of a two-dimensional space-time. A system S_q of gnostical events can easily be shown to exist together with three isomorphic mappings in which:

(1) a gnostical event corresponds to each relativistic event;

(2) a quantifying channel from the group G_q corresponds to each operator from the Lorentz's group of transformations of space-time;

(3) a gnostical tensor of entropy–entropy flow corresponds to the energy–momentum tensor of each relativistic event.

It is important that the mappings are *linear*. These mappings appeared as consequences of Axiom 1, having nothing common with mechanics. This axiom reflects the nature of real data obtained via quantification in an arbitrary application field, but mechanics can be a special case and a general theory should hold for a special case too. Let the points of S_r be obtained by means of an uncertain quantification and let each of them be measured using a coordinate system moving with an unknown (individual) velocity. Then the mentioned correspondence will not only be formal, but also the system S_r will turn out to be a model of S_q and vice versa. But a system of relativistic events can be represented by a single characteristic event (a "composite" event) using the known relativistic composition law which is a law of Nature: the conservation of energy and of momentum. The same mapping of composite events as for the individual ones thus warrants a correspondence of the gnostical composition law to the relativistic one. The energy–momentum tensors are composed additively and the mapping is linear. The gnostical tensors should therefore also be added to get a tensor of the proper composite event.

4.3. *Data composition law*

Axiom 2. This axiom of the gnostical theory (the data composition law) states the following. Let $\mathbf{K}_q(2\Omega_i)$ and $\mathbf{K}_e(2\omega_i)$ be gnostical tensors corresponding to data z_i from a data sample $Z(z_0, s, n)$; let Ω_c and ω_c be parameters of the composite event.

Then

$$\mathbf{K}_q(2\Omega_c) = \frac{1}{w_q}\sum_i^n \mathbf{K}_q(2\Omega_i)$$

$$\mathbf{K}_e(2\omega_c) = \frac{1}{w_e}\sum_i^n \mathbf{K}_e(2\omega_i), \qquad (43)$$

where w_q and w_e are normalizing weights of composite tensors $\mathbf{K}_q(2\Omega_c)$ and $\mathbf{K}_e(2\omega_c)$ determined by the normalizing condition

$$\mathrm{Det}\,\mathbf{K}_q(2\Omega_c) = \mathrm{Det}\,\mathbf{K}_e(2\omega_c) = 1, \qquad (44)$$

analogous to (11).

Comment. This axiom appears to be a generalization of the "relativistic" special case discussed above. The generalization extends its validity for composition of uncertain data of an arbitrary nature satisfying Axiom 1. It also relates to estimation. This extension is motivated by the duality between estimation and quantification.

4.4. *Explicit formulae of composition*

Using (43), (7), (38), (42) and (39) the infidelity and the quantifying irrelevance of the composite event of a data sample $Z(z_0, s, n)$ is obtained:

$$(f^{-1})_c = w_q^{-1}\sum_i^n f_i^{-1} \qquad (45)$$

$$h_{qc} = w_q^{-1}\sum_i^n h_{qi}, \qquad (46)$$

where

$$w_q = \sqrt{\left(\sum_i^n f_i^{-1}\right)^2 - \left(\sum_i^n h_{qi}\right)^2}. \qquad (47)$$

The estimating versions are analogously

$$f_c = w_e^{-1}\sum_i^n f_i \qquad (48)$$

$$h_{ec} = w_e^{-1} \sum_i^n h_{ei} \tag{49}$$

$$w_e = \sqrt{\left(\sum_i^n f_i\right)^2 + \left(\sum_i^n h_{ei}\right)^2}. \tag{50}$$

Having obtained estimates $\tilde{s}$ and $\tilde{z}_0$ and substituted them into (42) instead of parameters s and z_0, respectively, estimates of composite characteristics (45)–(50) are thus obtained as functions only of data $z_1, \ldots, z_n$ from the sample Z.

4.5. *Statistical interpretation of gnostical characteristics*

Denote

$$e_i = z_i/z_0 - 1, \quad i = 1, \ldots, n \tag{51}$$

$$\varepsilon = \max_i |e_i (z_i \in Z)|. \tag{52}$$

Let $\bar{e}$ and $\overline{e^2}$ be arithmetical means of all errors e_i and their squares. The case of a small ε will be referred to as the case of weak uncertainties. Using Taylor's expansion the following relations, valid for the simplified case with $s = 1$, are obtained.

$$(f^{-1})_c = 1 + 2\overline{e^2} + 0(\varepsilon^3)$$

$$f_c = 1 - 2\overline{e^2} + 0(\varepsilon^3) \tag{53}$$

$$h_{qc} = 2\bar{e} + 0(\varepsilon^2)$$

$$h_{ec} = -2\bar{e} + 0(\varepsilon^2) \tag{54}$$

$$w_q = n(1 + 2(\overline{e^2} - (\bar{e})^2) + 0(\varepsilon^3))$$

$$w_e = n(1 - 2(\overline{e^2} - (\bar{e})^2) + 0(\varepsilon^3)). \tag{55}$$

Under weak uncertainties both irrelevancies h_{qc} and h_{ec} and the infidelity $(f^{-1})_c$ as well as the fidelity f_c approach linear functions of the ordinary relative error $\bar{e}$ and of the relative value of sample variance $\overline{e^2}$, respectively. The weights w_q and w_e approach the size n. In such a situation the gnostical composition law approaches the classical statistical composition law. Otherwise the results can be quite different. It can be shown by simple analysis of formulae (39)–(42) and (45)–(50) that the estimating characteristics will be more robust with respect to outliers and the quantifying characteristics will be more robust with respect to inliers then the classical statistical characteristics. Components $(f^{-1})_c$ and f_c of gnostical tensors $\mathbf{K}_{qc}^2$ and $\mathbf{K}_{ec}^2$ are thus generalized characteristics of the variability of the data from a sample, the h_{qc} and h_{ec} being characteristics of relative errors of data measured using certain non-Euclidean geometries.

Robust generalizations of ordinary covariances appear here as well. Denoting arithmetic means by bars as before (47) and (50) are obtained in the form

$$w_q = n\sqrt{(\overline{f^{-1}})^2 - \overline{h}_q^2} \quad w_e = n\sqrt{\overline{f}^2 + \overline{h}_e^2}. \tag{56}$$

Now new characteristics of a data sample $Z(z_0, n, s)$ are introduced so that

$$\tilde{h_q^2} = \frac{1}{n}\left(\overline{h_q^2} + 2\sum_{k=1}^{n-1} \frac{n-k}{n} C_q(k)\right) \tag{57}$$

$$\tilde{h_e^2} = \frac{1}{n}\left(\overline{h_e^2} + 2\sum_{k=1}^{n-1} \frac{n-k}{n} C_e(k)\right), \tag{58}$$

where $\tilde{h_q^2}$ and $\tilde{h_e^2}$ can be interpreted as the quantifying and estimating variance, respectively, and where the quantities $C_q(k)$ and $C_e(k)$

$$C_q(k) := \frac{1}{n-k} \sum_{i=1}^{n-k} h_{qi} h_{q,i+k} \tag{59}$$

$$C_e(k) := \frac{1}{n-k} \sum_{i=1}^{n-k} h_{ei} h_{e,i+k} \tag{60}$$

are gnostical generalizations of *correlation coefficients*. It obviously holds that

$$C_q(0) = \overline{h_q^2} \quad C_e(0) = \overline{h_e^2} \tag{61}$$

as in the case of ordinary statistical variances and correlation coefficients. All these quantities can be expressed as functions of the data ratio q_i(42) using formulae (39)–(41).

4.6. *Robustness of characteristics of a data sample*

The following aspects of robustness of sample characteristics will be discussed here.

(A) Robustness with respect to assumptions on the model and to changes of the model ("model robustness").

(B) Robustness with respect to "bad" data ("data robustness"):

- (a) outlier robustness;
- (a) inlier robustness.

A datum is an *outlier* if its value is far from the values of data *(inliers)* forming a main cluster of data. A feature opposite to robustness is the *sensitivity*. High model sensitivity of sample characteristics of classical statistics motivated the development of robust statistical theory. However, most robust statistical methods are based not only on axioms of classical theory but also on some additional heuristic assumptions.

Gnostical characteristics follow directly from the gnostical axioms without using statistical assumptions at all. The robustness of gnostical characteristics appears not as a result of some additional requirements but as their inherent, natural feature. Their model robustness will depend of course on actual distribution of data but in a way determined by data robustness: inlier robust characteristics will be less sensitive to changes of central parts of distributions while outlier robustness will suppress the influence of their outer parts.

Both kinds of robustness may be required depending on the application. Noise filters, estimators of location and of scale parameters, identifiers and fitting procedures working under strong disturbances of data should be outlier robust. The inlier robustness can be practically desirable too, e.g. for detection of signals over noise background and for testing of hypotheses, in image enhancement, protection, and emergency systems.

5. APPLICATIONS OF GNOSTICAL THEORY

The theoretical results above can be used to obtain solutions of different practical problems.

5.1. *Estimation of the scale parameter of a data sample*

Problem. A data sample $Z(z_0, s, n)$ of ordered data $z_1, \ldots, z_n$ $(z_i \geq z_{i-1})$ is given, having a model (38) where z_0 and s are unknown parameters of the location and scale, respectively. It is required to find an estimate $\tilde{s}$ of the scale parameter s.

Solution.

$$\tilde{s} = \underset{s}{\operatorname{argmin}} \max_j \max(|p_{ecj} - F_{j-}|, |p_{ecj} - F_{j+}|) \tag{62}$$

where (for $j = 1, \ldots, n$)

F_{j-} and F_{j+} are values of the empirical distribution function at the points $z_j - \varepsilon$ and $z_j + \varepsilon$ (for small $\varepsilon > 0$)

$$p_{ecj} = (1 + h_{ecj})/2 \quad \text{as in (36)} \tag{63}$$

$$h_{ecj} = w_{ej}^{-1} \sum_i^n h_{eij} \quad \text{as in (49)} \tag{64}$$

and for $i = 1, \ldots, n$

$$h_{eij} = (q_{ij}^{-2} - q_{ij}^2)/(q_{ij}^2 + q_{ij}^{-2}) \quad \text{as in (41)} \tag{65}$$

$$q_{ij} = (z_i/z_j)^{1/s} \quad \text{as in (42)} \tag{66}$$

$$w_{ej}^2 = \left(\sum_i^n f_{ij}\right)^2 + \left(\sum_i^n h_{eij}\right)^2 \quad \text{as in (50)} \tag{67}$$

$$f_{ij} = 2/(q_{ij}^2 + q_{ij}^{-2}) \quad \text{as in (39).} \tag{68}$$

Comment. The quantity p_{ecj} is the value of gnostical distribution function of data at the point $z = z_j$. Equation (62) is thus a condition of the uniformly best approximation of the empirical distribution function by the gnostical distribution function.

5.2. *Estimation of the probability distribution function*

Problem. Under the formulation of 5.1, it is required to estimate the probability that a new datum (of the same origin as $z_1, \ldots, z_n$) will be not smaller than a quantity $z_j (z_j \in R_+)$.

Solution. (1) Estimate the scale parameter s according to 5.1.

(2) Substitute the estimate $\tilde{s}$ into (66).

(3) Evaluate the probability as $\tilde{p} = p_{ecj}$ where p_{ecj} is given by (63) using (64)–(68).

5.3. *Estimation of probability density*

Problem. Under the formulation of 5.1, it is required to estimate the density of probability of data at an arbitrary point $z_j (z_j \in R_+)$.

Solution. (1) Differentiate (63) to get the data density

$$\frac{\mathrm{d}p_{ecj}}{\mathrm{d}z_j} = \frac{1}{s z_j} \cdot \frac{(\bar{f})^2 \overline{f^2} + \bar{f} \bar{h}_e (\overline{h_e f})}{((\bar{f})^2 + (\bar{h}_e)^2)^{3/2}}. \tag{69}$$

(2) Estimate the scale parameter according to 5.1.

(3) Substitute the estimate $\tilde{s}$ into (66) and (69).

(4) Evaluate the density (69) using (64)–(68).

Comment. Data density function (69) is not necessarily unimodal for a particular data sample.

5.4. *Robust estimation of a parameter of location of a data sample*

Problem. Under the formulation of 5.1, it is required to find an outlier robust estimate $\tilde{z}_0$ of the location parameter (of the ideal value z_0) maximizing the data density function.

Solution. (1) Estimate the scale parameter s as in 5.1

(2) Find the estimate $\tilde{z}_0$ by solving the equation

$$\tilde{z}_0 = \arg\max_{z_j}\left(\frac{dp_{ecj}}{dz_j}\right), \tag{70}$$

where the function to be maximized is (69).

Comment. The non-linear equation (70) is to be solved iteratively. As a first approximation the arithmetic mean $\bar{z}$ will generally be suitable. Equation (70) can have more than one solution if the density function is not unimodal. In this case, there is more than one cluster within the data sample. Each of the clusters has its parameter of location of the density maximum.

It can be shown that the estimate (70) is outlier robust; even more, the influence of peripheral data and peripheral subclusters of the sample on the location parameter is also suppressed.

5.5. *Estimation of a parameter of location of a symmetrical data sample*

Problem. Under the formulation of 5.1 it is required to find the optimal estimate $\tilde{z}_0$ of the location parameter z_0 under the assumption that the data sample is multiplicatively symmetric. The estimate should be

(a) outlier robust

or (b) inlier robust.

Solution. (1) Estimate the scale parameter s as in 5.1 and substitute the estimate into (42).

(2) Find the estimate $\tilde{z}_0$ by solving the equation

$$\text{(a)} \quad h_{ec}(\tilde{z}_0) = 0 \tag{71a}$$

or

$$\text{(b)} \quad h_{qc}(\tilde{z}_0) = 0, \tag{71b}$$

respectively, using relations (66), (40), (41), (46) and (49).

Comment. An example of the multiplicative symmetry: the values $2z_0$ and $z_0/2$ are symmetrically located with respect to the value z_0. It is obvious that by taking the logarithm of multiplicatively symmetric data, additive symmetry results. The solution of (71b) and (71a) is unique. The optimality condition can be interpreted as a requirement to get zero information loss of the composite event (see (34) and (43)).

5.6. *Testing of unimodality of the density of a sample*

Problem. Under the formulation of 5.1, it is required to test the hypothesis that the density function of the sample is unimodal.

Solution. (1) Estimate the location parameter $(\tilde{z}_0)_1$ according to 5.4 using the initial value of $\tilde{z}_0$ for iterations equalling $\min_i(z_i)$.

(2) Repeat the same, starting the iterations from $\max_i(z_i)$ to get $(\tilde{z}_0)_2$.

(3) Test: if $(\tilde{z}_0)_1 = (\tilde{z}_0)_2$ then the data sample has a unimodal density.

5.7. *Test of a membership of a datum with a data sample*

Problem. Given a unimodal data sample $Z(z_0, s, n)$ of data $z_1 \ldots, z_n$ and a datum z_{n+1}, it is required to test the hypothesis that the datum z_{n+1} could be a member of the sample Z.

Solution. (1) Extend the data sample Z by the datum z_{n+1} to get a data sample $Z'(z_0, s', n+1)$.

(2) Test the unimodality of the density of the sample Z' according to 5.6. The datum z_{n+1} may be considered to be a possible member of the sample Z iff the sample Z' has a unimodal density function.

5.8. *Analysis of gnostical variance*

Gnostical variances appearing in (57) and (58) may be used alternatively to get robust estimates of location (Kovanic, 1984b), to analyse data density, to test the unimodality of data samples and so on, in a manner analogous to procedures mentioned above.

5.9. *Robust filtering of a time series*

Problem. A time series $z_1, \ldots, z_t$ of noisy observations of a process is given, the true value of which can be considered to be a constant equal to z_0. The model of observed data is (38). It is required to propose a filtering procedure treating the last n-tuple of observations $z_{t-n+1}, \ldots, z_t$ the output of which, $\tilde{z}_0$, would be robust with respect to strong noise.

Solution. Proceed in the same way as in 5.4 or 5.5, successively forming and treating the data sample Z from the last n-tuple of observations. Use the estimate $\tilde{z}_0$ as the robustly filtered output.

Comment. Applying the procedure according to 5.2 to Z also, a reliable and adaptive diagnosis of sudden changes of the process can be obtained: the estimate may be compared with a given threshold to activate signalization of appearance of unexpected changes of data.

Another robust filtering procedure may be obtained as a recursive solution of the identification problem dealt with below in 5.11.

5.10. *Robust correlation function of a process*

Problem. A time series of equidistant observations $z_1, \ldots, z_n$ of a process is given. The data model is (38), where z_0 and s are constant. Estimated values of the correlation function of the data series at some given points $k\,\Delta t$ (where $k = 0, 1, 2, \ldots$ and Δt is the distance between two observations) are required. Estimates of correlations should be robust with respect to strong errors.

Solution. (1) Estimate parameters s and z_0 according to 5.1 and 5.4 or 5.5.
(2) Estimate values of the correlation function according to (59) (inlier robust) or (60) (outlier robust) for all required values of k using (40)–(42).

5.11. *Robust identification of parameters of a dynamical system*

Problem. A mathematical model of a multidimensional dynamical system is given in the form of a system of ordinary differential equations of first order (with some given intial conditions), parameterized by an unknown parametric vector $\mathbf{c}$. A set of k observations of values of m variables of the system, obtained in an experiment starting with the given initial conditions, and the estimates $s_1, \ldots, s_m$ of scale parameters, characterizing the random errors of observed variables, are given. Given an *a priori* estimate of the vector $\mathbf{c}$, an improvement of this estimate, in a manner robust with respect to strong observation errors, is required.

Solution. (1) Integrate the equation system using the estimate $\bar{\mathbf{c}}$ resulting from the preceding iteration to get all $m \times k$ theoretical values ${}_tz_{ij}$ of variables in observation points corresponding to experimental data ${}_ez_{ij}$ for $i = 1, \ldots, k$ and $j = 1, \ldots, m$.
(2) Evaluate all $k \times m$ quantities

$$q_{ij} = ({}_tz_{ij}/{}_ez_{ij})^{1/s_j} \qquad (72)$$

and corresponding values h_{eij} (65) of estimating irrelevancies.
(3) Get a new estimate $\bar{\mathbf{c}}'$ by solving the equation

$$\bar{\mathbf{c}}' = \underset{\mathbf{c}}{\operatorname{argmin}} \left(\sum_i^k \sum_j^m h_{eij}^2 \right) \qquad (73)$$

of minimization of the estimating variance (61).
(4) If the distance between $\bar{\mathbf{c}}$ and $\bar{\mathbf{c}}'$ is not sufficiently small then go to (1). Else go to end.

Comment. Using alternatively h_{qij} of the type (40) instead of h_{eij} a solution robust with respect to inliers may be obtained.

5.12. *Identification of a regression model robust to both input and output disturbances*

Problem. Let F be a differentiable function of a known type,

$$F: R_m \times R_m \to R_+. \qquad (74)$$

Let $\mathbf{c} \in R_m$ be a column vector of unknown parameters. Let $\mathbf{x} \in R_m$ denote an "input" vector of a regression model

$$z = F(\mathbf{c}, \mathbf{x}) \qquad (75)$$

the quantity z being the "output". Suppose that neither the input nor output is known precisely. Let the true values $\mathbf{x}_{oi}$ and z_{oi} at an ith observation point be related through the equation

$$z_{oi} = F(\mathbf{c}, \mathbf{x}_{oi}). \qquad (76)$$

Let $\mathbf{x}_i$ and z_i denote the uncertain observations of $\mathbf{x}_{oi}$ and z_{oi}, respectively, for $i = 1, \ldots, n$.

Given a kth iteration $\mathbf{c}_k$ of the estimate $\bar{\mathbf{c}}$ of $\mathbf{c}$, an estimate $\bar{s}$ of the scale parameter s, and a twice differentiable criterion function

$$\varphi = \sum_i^n D(h_{ei}), \qquad (77)$$

where

$$D: = r_1 \to R_+ \qquad (78)$$

$$h_{ei} = ((\bar{z}_{oi}/z_i)^{2/\bar{s}} - (z_i/\bar{z}_{oi})^{2/\bar{s}})/((\bar{z}_{oi}/z_i)^{2/\bar{s}} + (z_i/\bar{z}_{oi})^{2/\bar{s}})$$

$$\bar{z}_{oi} = F(\mathbf{c}_k, x_i), \qquad (79)$$

it is required to find a $(k+1)$th iteration $\mathbf{c}_{k+1}$ maximizing the function φ.

Solution. Introduce the column-vector

$$\mathbf{g}_i = \frac{1}{\bar{z}_{oi}}(F'_{i_1}, \ldots, F'_{im})^{\mathrm{T}}. \tag{80}$$

where

$$F'_{ij} = \left(\frac{\partial F}{\partial c_j}\right)_{\mathbf{c}_k, \mathbf{x}_i} \tag{81}$$

and

$$D'_i = \left(\frac{\partial D}{\partial h_e}\right)_{z_0 = \bar{z}_{oi}}$$

$$D''_i = \left(\frac{\partial^2 D}{\partial h_e^2}\right)_{z_0 = \bar{z}_{oi}}. \tag{82}$$

Then the $(k+1)$th iteration of the solution—if it exists—is approximated by the formula

$$\mathbf{c}_{k+1} = \mathbf{c}_k + \bar{s}\left[\sum_i^n f_i^4 D''_i \mathbf{g}_i \mathbf{g}_i^{\mathrm{T}}\right]^{-1} \times \left[\sum_i^n \frac{1}{2} f_i^2 D'_i \mathbf{g}_i\right], \tag{83a}$$

where $f_i^2 = 1 - h_{ei}^2$ and where all quantities on right-hand side are estimated by substitution of $\mathbf{c}_k$ and $\bar{z}_{oi}$ instead of $\mathbf{c}$ and z_{oi}, respectively.

Proof. The equation

$$(\mathrm{d}\varphi)_{\mathbf{c}} = 0 \tag{84}$$

can be solved iteratively by the Newton–Raphson method. When the kth approximation $\mathbf{c}_k$ is already given then the quantity $(\mathbf{c}_{k+1})$ may be approximated by the first two terms of the Taylor expansion of the function φ at the point $\mathbf{c}_k$. This approximation is taken as a solution. Therefore

$$(\mathrm{d}\varphi + \mathrm{d}^2\varphi)_{\mathbf{c}_k} = 0. \tag{85}$$

Both differentials are functions of the irrelevance h_e, hence

$$\left(\sum_i^n D'_i(\mathrm{d}h_{ei} + \mathrm{d}^2 h_{ei}) + \sum_i^n D''_i \mathrm{d}h_{ei}^2\right)_{\mathbf{c}_k} = 0. \tag{86}$$

The second differential $\mathrm{d}^2 h_{ei}$ is neglected since it is a small quantity with respect to the first one $\mathrm{d}h_{ei}$. The total differential of the variable z_0 as a function of $\mathbf{c}_k$ at the ith point is

$$\mathrm{d}z_{0i} = (\mathrm{d}z_0)_i = \sum_j^n \left(\frac{\partial F(\mathbf{x}, \mathbf{c}_k)}{\partial c_{kj}}\right)_{\mathbf{x}_i} \mathrm{d}c_{kj} = \bar{z}_{oi}\mathbf{g}_i^{\mathrm{T}} \mathrm{d}\mathbf{c}_k, \tag{87}$$

where c_{kj} denotes the jth component of the vector $\mathbf{c}_k$. Thus,

$$\mathrm{d}h_{ei} = 2f_i^2 \mathrm{d}z_{oi}/(\bar{s}\bar{z}_{oi}) = 2f_i^2 \bar{s}^{-1}\mathbf{g}_i^{\mathrm{T}}\mathrm{d}\mathbf{c}_k \tag{88}$$

and

$$\mathrm{d}h_{ei}^2 = 4f_i^4 \bar{s}^{-2} \mathrm{d}\mathbf{c}_k^{\mathrm{T}} \mathbf{g}_i \mathbf{g}_i^{\mathrm{T}} \mathrm{d}\mathbf{c}_k. \tag{89}$$

Equation (86) will hold identically for all $\mathrm{d}\mathbf{c}_k$ when

$$\left[\sum_i^n f_i^4 D''_i \mathbf{g}_i \mathbf{g}_i^{\mathrm{T}}\right]_{\mathbf{c}_k} \mathrm{d}\mathbf{c}_k = \left[\sum_i^n \bar{s} f_i^2 \frac{D'_i}{2} \mathbf{g}_i\right]_{\mathbf{c}_k}. \tag{90a}$$

If the system matrix is regular then for $\mathrm{d}\mathbf{c}_k \doteq \mathbf{c}_{k+1} - \mathbf{c}_k$ equation (83) actually holds. These results can be explained by comparing them with the simplest least-squares identification. Indeed, the expression (83) may be interpreted as the least-squares solution

$$\mathbf{c}_{k+1} - \mathbf{c}_k = \left[\sum_i^n \mathbf{G}_i \mathbf{G}_i^{\mathrm{T}}\right]^{-1} \left[\sum_i^n \mathbf{G}_i E_i\right] \tag{83b}$$

of a system of n linear equations having the form

$$\mathbf{G}_i^{\mathrm{T}}(\mathbf{c}_{k+1} - \mathbf{c}_k) = E_i, \tag{90b}$$

where

$$\mathbf{G}_i = f_i^2 \sqrt{-D''_i}\,\mathbf{g}_i \tag{91}$$

and

$$E_i = \frac{\bar{s}}{2}\frac{D'_i}{\sqrt{-D''_i}}. \tag{92}$$

Vector $\mathbf{G}_i$ plays the role of the input of the linear model (90b) the scalar E_i being the output. Transformations $\mathbf{g}_i \to \mathbf{G}_i$ consisting of multiplying $\mathbf{g}_i$ by a function of the ratio $z_i/\bar{z}_{io}$ may be interpreted as non-linear filtering. Three examples of gnostical criterion functions $D(h_e)$ are given in Table 3 together with corresponding weights $f_i^2\sqrt{-D''_i}$. Graphically this is depicted in Fig. 1. All weights decrease rapidly with a deviation of the observed datum z_i from the predicted "true" value $\bar{z}_{oi}$.

The quantity E_i (92) represents a gnostical error function, an error of the ith datum z_i with respect to the "true" value $\bar{z}_{oi}$. In other words, the expression (92) shows how errors are measured when using a gnostical metric. For a weak uncertainty (the ratio $z_i/\bar{z}_{oi}$ differing from unity only

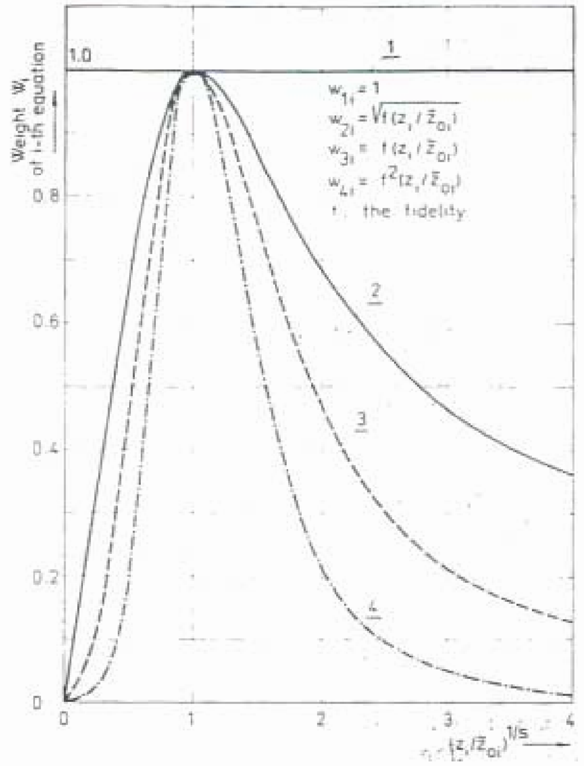

FIG. 1. Weights $w_{ji} = -f_i^4 D_{ji}''$ appearing on the left hand side of (90), which represent the solution of the robust identification problem. Index j denotes the type of criterion function (see Table 3).

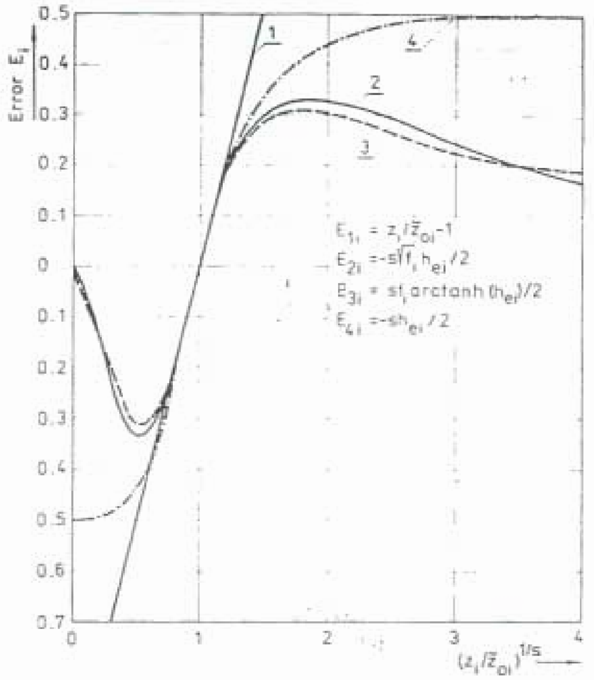

FIG. 2. Weighted errors E_{ji} (92) of the solution of the robust identification problem. Index j identifies the type of criterion function (see Table 3).

slightly) the error E_i approaches $z_i/\bar{z}_{oi} - 1$, which is the "classical" (Euclidean) relative error. However, for a strong uncertainty the behaviour of the error function $E(h_{ei})$ (i.e. $E(z_i/\bar{z}_{oi})$) is far from linear as seen in Table 3 and Fig. 2 for the three types of gnostical criterion functions. Equation (92) may be also interpreted as a non-linear filter. This filter is applied to the output quantity of the indentified model. As seen from Table 3 and Figs 1 and 2, both filters protect the identification process against the influence of outlying (both input and output) data.

5.13. *Robust control*

There are at least three levels of possible applications of results of gnostical theory to feedback control systems with an improved robustness with respect to bad data:

(1) application of a filter of the type described in Section 5.9 or 5.11;
(2) application of an identifier of the type described in Section 5.11;
(3) using a gnostical error function instead of relative difference between reference and actual values of a controlled variable as a control error.

These suggestions open a broad field of theoretical problems. An example of their possible practical impact is given in Section 6.2.

6. EXAMPLES

6.1. *Example 1. A comparison of robust estimators of location*

To compare a gnostical estimator with the stastistical ones a well-known study of Stigler (1977) can be extended (Kovanic and Novovičová, 1986). Two ancient and popular estimators (mean and median) and nine recent robust and adaptive statistical estimators have been tested in Stigler's study using real data from famous physical experiments performed in 18th and 19th century. The set of estimators is extended by the gnostical estimator described in Section 5.4 above using the scale parameter according to Section 5.1. The estimators are applied to 16 independent data series from three experiments:

(1) Short's (1763) determinations of the parallax of the sun (eight data samples);
(2) Newcomb's (1882) measurements of the passage time of light (three samples);
(2) Michelson's (1879) determinations of the velocity of light in air (five samples)

A complete reference for these data is in Stigler (1977).

The size of each of 16 data samples was about 20. The same relative deviation as in Stigler's study has been used for evaluation

$$e_{ij} = |\bar{\theta}_{ij} - \theta_j|/S_j,$$

$$\text{where} \quad S_j = \frac{1}{12} \sum_{i=1}^{12} |\bar{\theta}_{ij} - \theta_j| \qquad (93)$$

TABLE 3. "LINEARIZING" TRANSFORMATION FOR ROBUST IDENTIFICATION

Case j	Uncertainty of data	Type of gnostical criterion function	Weights of summands in (90) $-f_i^4 D''_{ji}$	$-f_i^2 D'_{ji}/2$	Weight of the ith equation	Error function, E_j
1	Weak	$D_1 \approx D_2 \approx D_3$	1	$z_i/\bar{z}_{ei} - 1$	1	$z_i/\bar{z}_{ei} - 1$
2	Strong	$D_1 = \sqrt{1 - h_e^2} = f$	f_i	$\bar{s}f_i h_{ei}/2$	$\sqrt{f_i}$	$-\bar{s}\sqrt{f_i}h_{ei}/2$
3	Strong	$D_2 = -I_e(h_e)$	f_i^2	$\bar{s}f_i^2 \operatorname{arctanh}(h_{ei})/2$	f_i	$-\bar{s}f_i \operatorname{arctanh}(h_{ei})/2$
4	Strong	$D_3 = (1 - h_e^2)/2$	f_i^4	$\bar{s}f_i^2 h_{ei}/2$	f_i^2	$-\bar{s}h_{ei}/2$

and $\bar{\theta}_{ij}$ is the estimate of location parameter of the jth sample ($j = 1, \ldots, 16$) obtained by the ith estimating method ($i = 1, \ldots, 12$). The quantity θ_j is the currently used "true" value of the measured quantity. Stigler's evaluation has been "the less e_{ij}, the better the ith estimator when tested on jth sample". But as shown in his study as well as in the discussion which followed it, the real data under consideration were strongly biased. All tested estimators are asymptotically unbiased, therefore there is a good reason to accept the unit to be the true value of e_{ij} for all tests, at least from the point of view of an "expert board" consisting of all 12 estimating methods together. Then the evaluation should be "the closer is the value of e_{ij} to 1, the better is the ith estimator when tested on jth sample".

The results of such a comparison are summarized in Table 4 for all 12 estimators ordered according to the mean error. A description of all 11 statistical estimators under consideration is in Stigler (1977). The gnostical estimator used algorithms described in 5.4 and 5.1 above. Data were exponentiated before application of this estimator and the logarithm of result was taken. The gnostical estimator appeared to be superior over the statistical ones.

TABLE 4. RESULTS OF THE COMPARISON OF THE GNOSTICAL ESTIMATOR OF LOCATION WITH 11 STATISTICAL ESTIMTORS

Estimator i	Type	Errors: Standard deviation σ_i	Mean error $\lvert\mu_i - 1\rvert$	Range R_i
1	Gnostical	0.038	0.001	0.139
2	Hogg T 1	0.061	0.017	0.261
3	25% Trim	0.070	0.029	0.261
4	Edgeworth	0.079	0.011	0.273
5	15% Trim	0.104	0.032	0.447
6	Tukey Bi-weight	0.131	0.043	0.631
7	Andrews AMT	0.147	0.025	0.660
8	Huber P15	0.210	0.083	0.856
9	10% Trim	0.211	0.097	0.821
10	Mean	0.212	0.078	1.055
11	Median	0.278	0.124	0.962
12	Outmean	0.610	0.086	2.603

m, Number of independent testing data samples ($m = 16$); e_{ij}, Normalized result of estimation defined by (93); μ_1, Mean of e_{ij} over samples $\left(\mu_i = \sum_j^m e_{ij}/m\right)$; σ_i, Standard deviation $\left(\sigma_i = \sqrt{\sum_j^m (e_{ij} - \mu_i)^2/m}\right)$; R_i, Range of errors ($R_i = \operatorname{Max}_j |e_{ij} - 1| - \operatorname{Min}_j |e_{ij} - 1|$).

6.2. *Example 2: Control of a discrete non-linear system*

Consider a Volterra–Lotka discrete system of r_t rabbits and f_t foxes at time t

$$r_{t+1} = \mathrm{INT}((1.05 + 0.05x_{1,t} - (0.03 + 0.02x_{2,t})f_t/f_0 - u_t)r_t) \qquad (94)$$

$$f_{t+1} = \mathrm{INT}((1.15 + 0.05x_{3,t} - (0.135 + 0.045x_{2,t})r_0/r_t - v_t)f_t), \qquad (95)$$

which can be observed only with a delay and with observation errors having intensity c as

$$r'_t = r_{t-1}\exp(cx_{4,t})$$
$$f'_t = f_{t-1}\exp(cx_{5,t}), \qquad (96)$$

where the variables $x_{i,j}$ are independent random variables uniformly distributed between -0.5 and 0.5, and the initial values are $r_0 = 1700$ and $f_0 = r_0/60$. INT denotes the integer part of a real number, and u_t and v_t are the control variables. They should maintain constant levels of both controlled variables r and f, the set points being r_s and f_s. The equations of the controller are:

$$\text{if } r'_t \geq r_s \quad \text{then } u_t = 0.8e(r'_t, r_s) + 0.1e(f'_t, f_s) + 0.5e(F(r'_t), r_s), \quad \text{else } u_t = 0; \qquad (97)$$

$$\text{if } f'_t \geq f_s \quad \text{then } v_t = 1.7e(f'_t, f_s) - 0.7e(r'_t, r_s) + 0.8e(F(f'_t), f_s), \quad \text{else } v_t = 0. \qquad (98)$$

Here $F(y_t)$ represents a filtered value of a variable y at time t obtained recursively as

$$F(y_t) = Y_t/T_t \quad \text{where } Y_t = 0.8Y_{t-1} + y_t \text{ and } T_t = 0.8T_{t-1} + t. \qquad (99)$$

The symbol $e(y'_t, y_s)$ characterizes the control error of the observed value y'_t of a variable y with respect to its reference value y_s. There are two kinds of error functions of interest here. The control error in the *classical* (Euclidean) metric:

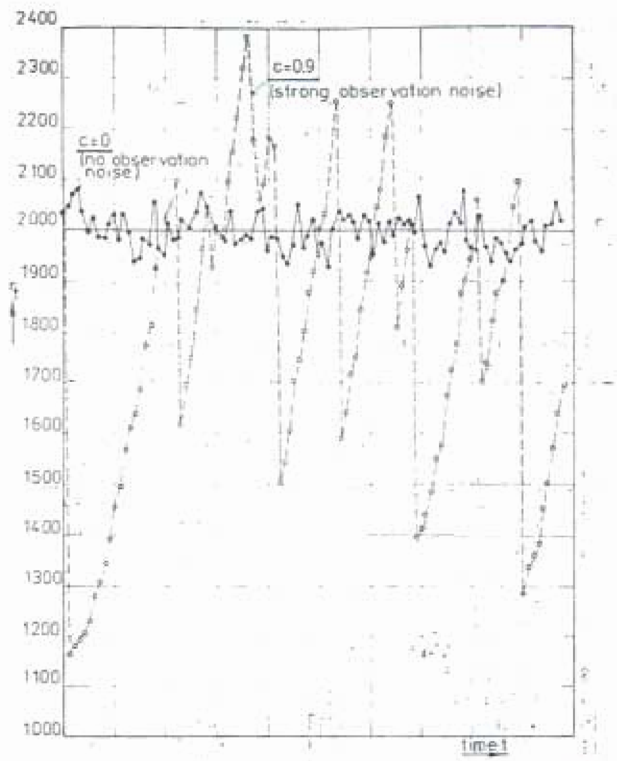

FIG. 3. Automatic control of the system of Example 2 using the ordinary (Euclidean) metric for the evaluation of the control error. Control is sensitive to observation errors.

$$e(y_t', y_s) = y_t'/y_s - 1 \qquad (100)$$

and the control error in *gnostical* metric:

$$e(y_t', y_s) = (\bar{s}/2)(q_t^2 - q^{-2}{}_{1/(q)}{}^2 + q_t^{-2}), \qquad (101)$$

where $q_t = (y_t'/y_s)^{1/3}$

and where $\bar{s}$ is an estimate of the scale parameter of the data y_t. One-way action of the controller corresponds to the nature of the process: the number of rabbits or foxes can be decreased immediately by killing some of them but their increase can be influenced only indirectly.

The parameters of the controller have been estimated experimentally as locally optimal values for the case of zero observation errors ($c = 0$) and for the classical control error function to minimize the mean absolute control error. Figure 3 shows the course of the number of rabbits for this case together with the case of strong observation errors ($c = 0.9$) and for the controller working also with

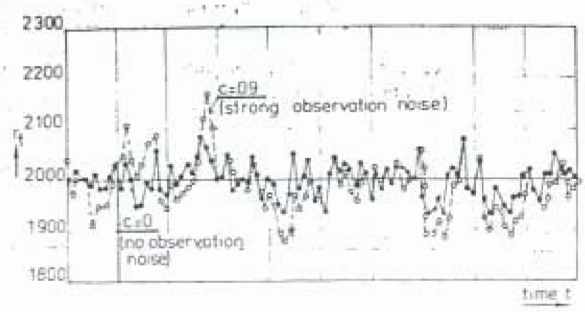

FIG. 4. Automatic control of the system of Example 2 using the gnostical metric for the evaluation of the control error. Application of a non-linear filter of gnostical type increases the robustness of the control with respect to observation errors.

classical error functions. Both cases $c = 0$ and $c = 0.9$ are depicted in Fig. 4 for the gnostical error function, with the same data sequence as in the classical case and the same parameters of the controller. Numerical results are summarized in Table 5. A mere substitution of the gnostical error function instead of the classical one substantially improves the robustness of the system.

It is worth mentioning that the gnostical error function (101) has been obtained rigorously as a very special case of the gnostical identification problem 5.11.

7. CONCLUSIONS

The gnostical theory of real data offers new formulae of characteristics for data samples. These formulae can be directly evaluated from data. They include analogies of classical statistical characteristics of data samples and possess either increased robustness or increased sensitivity with respect to outlying or inlying data. New characteristics of individual data and of small data samples are available, too, such as information, thermodynamical entropy and probability distribution, not based on *a priori* assumptions on a statistical model but using only data. Gnostical characteristics are also suitable as criterion functions for optimization to obtain new efficient and robust algorithms for identification of models, filtering and prediction, decision making, cluster analysis and automatic control.

REFERENCES

Jumarie, G. A. (1975). A relativistic information theory model for general systems. *Int. J. Syst. Sci.*, 6, 865–886.

TABLE 5. MEAN ABSOLUTE CONTROL ERROR FOR EXAMPLE 2

Intensity of observation errors (c in (96))		0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Control error	Classical error function (100)	0.010	0.019	0.029	0.037	0.048	0.061	0.090	0.083	0.094	0.106
	Gnostical error function (101)	0.013	0.015	0.017	0.021	0.023	0.024	0.024	0.024	0.024	0.024

Jumarie, G. A. (1985). *Subjectivity, Information, Systems. Introduction to a Theory of Relativistic Cybernetics.* Gordon and Breach, New York.

Kovanic, P. (1984a). Gnostical theory of individual data. *Probl. Control Inf. Theory*, 13, 259–274.

Kovanic, P. (1984b). Gnostical theory of small samples of real data. *Probl. Control Inf. Theory*, 303–319.

Kovanic, P. (1984c). On relations between information and physics. *Probl. Control Inf. Theory*, 13, 383–399.

Kovanic, P. and J. Novovičová (1986). (to appear).

Krants, D. H., R. D. Luce, P. Suppes and A. Tversky (1971). *Foundations of Measurement.* Academic Press, New York.

Stigler, S. M. (1977). Do robust estimators work with real data? *Ann. Statist.*, 5, 1055–1098.