2/IV-84

объединенный
институт
ядерных
исследований
дубна

1609/84

E17-83-874

E.R.Caianiello*

GEOMETRICAL "IDENTIFICATION" OF QUANTUM AND INFORMATION THEORIES

Submitted to "ТМФ"

* On leave from the University of Salerno, Italy

1983

Quantum mechanics and information theory both deal with "phenomena" (rather than "noumena" as does classical physics) and, in a way or another, with "uncertainty", i.e., lack of information. The issue may be far more general, and indeed of a fundamental nature to all science; e.g., recently R.E.Kalman [1] has shown that "uncertain data" lead to "uncertain models" and that even the simplest linear models must be profoundly modified when this fact is taken into account "in all branches of science, including time-series analysis, economic forecasting, most of econometrics, psychometrics and elsewhere" [2]. The emergence of "discrete levels" in every sort of "structured systems" [3] may point in the same direction.

Connections between statistics, quantum mechanics and information theory have been studied in remarkable papers (notably by M.Tribus [4] and E.T.Jaynes [5]) based on the maximum (Shannon) entropy principle. We propose here to show that such a connection can be obtained in the "natural meeting ground" of geometry. Information theory and several branches of statistics have classic geometric realizations, in which the role of "distance" is played by the infinitesimal difference between two probability distributions: the basic concept is here <u>cross-entropy</u>, i.e., information, which we much prefer to entropy: the latter is essentially "static", while the former correlates a <u>posterior</u> to a <u>prior</u> situation and invites thus to dynamics. The "information distance" is here "complexified" and generalized in a way clearly forced by the formalism adopted. Quantum mechanics, if presented as "quantum geometry" (as proposed in previous occasions by us [6]), appears then to be a particular instance of this wider "complex information geometry".

INFORMATION GEOMETRY

<u>Metric</u>

Let $\mathbf{x} \equiv \{x^{(1)}, x^{(2)}, \dots, x^{(n)}\}$ and $\mathbf{z} \equiv \{z_{(1)}, z_{(2)}, \dots, z_{(m)}\}$ be two sets of real variables; in the current literature of statistics or estimation theory $\mathbf{x}$ is a point in "parameter space" S^n, $\mathbf{z}$ represents random variables. Thus, in the Gaussian distribution

$$\rho_{\text{Gauss}} = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{(z-\mu)^2}{2\sigma^2}\right\} \qquad (1)$$

we read $z \equiv \{z_{(1)}\}$, $x \equiv \{x^{(1)} = f^{(1)}(\mu,\sigma),\ x^{(2)} = f^{(2)}(\mu,\sigma)\}$. This interpretation of the roles of x and z is however not essential, and may be modified whenever convenient (it being a matter of "identification"), provided formal manipulations stay the same. The Kullback-Leibler information [7] (or cross-entropy):

$$I(1,2) = \int \rho(x_1|z)\log\frac{\rho(x_1|z)}{\rho(x_2|z)}\,dz \tag{2}$$

discriminates between distributions at points x_1 and x_2; setting $x_1 = x$, $x_2 = x + dx$ one has from (2)

$$2I(x+dx, x) = ds^2 = g_{hk}(x)\,dx^h dx^k, \tag{3}$$

where $(\partial_h = \frac{\partial}{\partial x^h})$,

$$g_{hk}(x) = g_{kh}(x) = \int \rho(x|z)\,\partial_h \log\rho(x|z)\,\partial_k \log\rho(x|z)\,dz \tag{4}$$

is the Fisher [7,8], or information metric; of course $\rho(x|z) \geq 0$ and

$$\int \rho(x|z)\,dz = 1\,. \tag{5}$$

Connection

The geometrical representation of a model (or theory, according to one's philosophy) requires essentially the choice of a metric G and a connection Γ_μ, and the identification of a reference frame; these depend upon the "universe" one wishes to model. For purposes of statistics and estimation the most general connection is of the form [9]

$$\overset{a}{\Gamma}_{ijk} = [ij,k] - \frac{a}{2}\int \partial_i \log\rho\ \partial_j \log\rho\ \partial_k \log\rho\ dz \tag{6}$$

with a an arbitrary real parameter. The only case of interest to us, in our comparison with quantum geometry, will be $a = 0$ (non-metricity [10] does not appear desirable at this stage). It is instructive however to consider first the general case, on the example of the Gaussian distribution (1), which one may write [11] as

$$\rho_{Gauss}(x|z) = \exp\{zx^{(1)} - z^2x^{(2)} - \frac{1}{4}\frac{[x^{(1)}]^2}{x^{(2)}} + \frac{1}{2}\log x^{(2)} - \frac{1}{2}\log\pi\}$$

$$x^{(1)} = \frac{\mu}{\sigma_1^2}\,, \quad x^{(2)} = \frac{1}{2\sigma^2}\,; \tag{7}$$

then [11] $G \equiv \{g_{hk}\} = \begin{pmatrix} \sigma^2 & 2\mu\sigma^2 \\ 2\mu\sigma^2 & 4\mu^2\sigma^2 + 2\sigma^4 \end{pmatrix}$ and

$$\overset{a}{R}_{1212} = (1 - a^2)\sigma^6\,. \tag{8}$$

Thus, the choice $a = 1$ renders Gaussian distributions as straight lines when an appropriate "natural" frame is taken in "parameter space" [$a = -1$ does the same for "mixture distributions": $\rho(x|z) = (1 - \sum_{h=1}^{n} x^{(h)})\,\rho_0(z) + \sum_{h=1}^{n} x^{(h)}\rho_h(z)$ $(0 \leq x^{(h)} \leq 1)$]. Excepting these cases, and in particular with the metric connection $\Gamma_\mu \equiv [ij;\mu]$ of interest to us, (8) tells that the curvature tensor expresses our lack of information: it vanishes only when $\sigma = 0$, i.e., when the position z is known with absolute certainty. Such possibility is the main intuitive notion we wish to draw from this example.

QUANTUM GEOMETRY

We refer here to the realization of one-particle quantum mechanics treated in earlier works [8]; it will suffice here to recall that its main ingredients are:

1) an Hermitian metric G in 8-dimensional relativistic phase space (which may derive from),

2) an anti-Hermitian connection Γ_μ, such that the ensuing Riemann tensor expresses the Heisenberg commutation relations, i.e., our "uncertainty" or lack of information; both position and momentum operators become covariant derivatives taken along the axes of x.

3) a quantum frame, whose fixing (or "polarization") is a prerequisite and corresponds to "identification" (often ignored by physicists but central to systems theory).

Statistics, etc., use mainly probability distributions ρ, while quantum mechanics prefers their "square roots", i.e., amplitudes. We shall need therefore first of all to re-formulate information geometry in terms of amplitudes. Information and systems theory deal with uncertainties in various ways, none of which excludes the theoretical possibility that they may be all rendered as small as wanted (though this view may have to be changed [1]); quantum physics sets instead theoretical limitations to this possibility. In model-thinking, the difference may appear less relevant, or not relevant at all, depending upon how much of theoretical or technical limitations one wishes to build within the model, or leave out as "error". The development of physics is a paradigm not to be overlooked by sciences which go from "soft" to "hard". We take the information metric (4) as basic; we wish, though, to treat with it also situations in which $\rho(x|z)$ is not given a priori, but has to be determined by means of additional considerations. The meaning of x and z will thus depend on the particular situation: again, "identification". We start with the remark that the metric element (4) can also be written as

$$g_{hk}(x) = 4\int \partial_h \sqrt{\rho(x|z)}\, \partial_k \sqrt{\rho(x|z)}\, dz \tag{9}$$

so that

$$\frac{1}{4} ds^2 = \int(\partial_h \sqrt{\rho}\, dx^h)(\partial_k \sqrt{\rho}\, dx^k)\, dz = \int (d\sqrt{\rho})^2 dz \,. \tag{10}$$

The variable z may of course take also discrete values: $\int \to \Sigma$, so that situations can be conceived in which (9), (10) simply express g_{hk} in terms of a holonomic viel-bein. Thus, if $\phi^{(x)}_{(a)} = \sqrt{\rho(x|z=a)}$, (9) reads in the familiar way

$$\frac{1}{4} g_{hk}(x) = \sum_a \partial_h \phi^{(x)}_{(a)} \partial_k \phi^{(x)}_{(a)} \,. \tag{11}$$

Many other situations can be expressed through the scalar product (9) (hypercomplete sets of states, etc.); the consideration of this typical "system-model" play will not concern us here.

The wanted generalization of the information metric (to include also our quantum metric $g_{hk}(x) = \overline{g_{kh}(x)}$) is now obvious: it reads (neglecting the irrelevant numerical factor)

$$g_{hk}(x) = \int \psi_h(x|z)\, \overline{\psi_k(x|z)}\, dz \tag{12}$$

if $\psi_h(x|z) = \partial_h \phi(x|z)$ (12) yields the general holonomic case; if $\overline{\partial_h \phi} = \partial_h \phi$ we fall back into the standard information metric (9) or (4).

The choice of the connection Γ_μ is the next step in the geometrical construction of a model, or theory, as was evidenced in the discussion of the gaussian and mixture distributions ($a = +1$). Comparison with our work on quantum geometry /8/ requires, as already stated, $a = 0$. It is interesting to see what happens in this case to the connection (6). $\overset{\circ}{\Gamma}_\mu$ reads, in the standard form (4) /11/:

$$\overset{\circ}{\Gamma}_{ijk}(x) = \int \rho\,[\partial^2_{ij} \log\rho\, \partial_k \log\rho + \frac{1}{2}\partial_i \log\rho\, \partial_j \log\rho\, \partial_k \log\rho]\, dz \,. \tag{13}$$

From

$$\prod_{i=1}^{k} \partial_i \log\rho = 2^k \rho^{-k/2} \prod_{i=1}^{k} \partial_i \sqrt{\rho} \tag{14}$$

one finds

$$\overset{\circ}{\Gamma}_{ijk} = 4\int \partial^2_{ij} \sqrt{\rho}\, \partial_k \sqrt{\rho}\, dz \,. \tag{15}$$

This unexpected simplification can hardly be a coincidence; it strongly suggests than (9) is more convenient, or "natural", than (4) for the construction of a wider, complexified geometry. It is important to remind also that the addition (if we start with (12)) of anti-hermitian terms to (15) leaves ds^2 invariant: it amounts to a gauge transformation /12/.

In this perspective, our formulation of quantum geometry obtains by restricting in a suitable way the "complex information geometry" sketched here and, of course, by suitably identifying its mathematical objects with those typical of quantum mechanics (fixing the quantum frame). Obviously, information geometry is another such restriction.

We suggest therefore, in conclusion, that "complex information geometry" may be an apt tool for attempting "technological transfer" among fields of investigation alien thus far to one another. As a final remark, we note that "infinitesimal distance" ds^2 and "infinitesimal cross entropy" dH_c (14) now coincide (to within approximate identifications):

$$ds^2 \equiv dH = g_{hk}\, dx^h\, \overline{dx}^k = \int \psi_h dx^h \bar{\psi}_k\, \overline{dx}^k dz \,. \tag{16}$$

Then, the requirement imposed by relativity for particles to be physical /15/:

$$ds^2 = dt^2 - \frac{1}{c^2} dx^2 \geq 0 \tag{17}$$

coincides with the requirement

$$dH_c \geq 0 \tag{18}$$

which physical phenomena must satisfy. The issue, whether or how $\psi_h(x)$ in (12) can be connected with solutions of (quantum-mechanical) equations, will be considered in a next note.

The author is greatly indebted to Professors S.Amari and R.E.Kalman for kindly communicating to him their important and stimulating results. Warmest thanks are also due to Professors N.N.Bogolubov and N.N.Bogolubov (Jr.) for their so cordial hospitality at JINR in Dubna.

REFERENCES

1. R.E.Kalman. a) Proc.Int.Symp. on Dynamical Systems. (Ed. by A.Bednarek). Academic Press, 1982; b) Current

Developments in the Interface: Economics, Econometrics, Mathematics. (Ed. by S.M.Hazenwinkel and A.H.G.Rinnoy Kan). Reidel Publ.Co., 1982.
2. Kalman R.E. b) above, p.135.
3. Caianiello E.R. Biol.Cybernetics, 1977, 26, p.11; Int. J.Gen.Systems, 1982, 8, p.81 (with G.Scarpetta and G.Simoncelli).
4. Tribus M. Thermostatics and Thermodynamics. Van Nostrand, New York, 1961.
5. Jaynes E.T. In: Brandeis Theor.Phys.Lectures on Statistical Physics, vol.3, New York, 1962.
6. Caianiello E.R. a) Lett.Nuovo Cimento, 1979, 25, p.225; 1980, 27, p.89; 1981, 32, p.65; 1982, 35, p.381; Nuovo Cimento, 1980, 59B, p.350; Proc. VI Int.Conf. on the Problems of Quantum Field Theory, Alushta, 1981. JINR, D2-81-543, Dubna, 1981; Quantization as Geometry in Phase Space. IV and V Conf. on Quantum Theory and the Structure of Time and Space. Tutzing (1980) and 1982; the latter to be published in "Semin.Mat.Fis." Univ.Milano); Lett. Nuovo Cim., 1981, 30, p.469 (with G.Vilasi); 1982, 33, p.555 (with S. de Filippo and S.G.Vilasi); 1982, 34, p.112 (with S. de Filippo, G.Marmo, G.Vilasi), 1982, 34, p.301 (with A.Giovannini); 1983, 36, p.487 (with G.Marmo and G.Scarpetta). b) "Geodesic and Hamiltonian Equations in Quantum Geometry: the Identification Problem. Lett. Nuovo Cim., in print (with G.Marmo and G.Scarpetta).
7. Kullback S.C. Information Theory and Statistics. Wiley, New York, 1959.
8. Fisher R.A. Proc.Cambridge Phil.Soc., 1925, 122, p.700.
9. Chentzov N.N. Statistical Decision Rules and Optimal Conclusions. "Nauka",Moscow, 1972 (in Russian).
10. Schouten J.A. Ricci Calculus. Springer, Berlin, 1954, 2nd ed.
11. Amari S. Raag Reps., 106, Febr.1980; Techn.Rep.Fac. - Engineering, University of Tokyo, METR 81-1, April, 1981.
12. Caianiello E.R. et al.: ref. 6 b) above.
13. In measure-theoretical treatments cross-entropy as used here is often called entropy: cf.ref.5 and 7 above -
14. It becomes $ds^2 = dt^2 - \frac{1}{c^2}dx^2 + \frac{h^2}{\mu^4 c^6}[dE^2 - dp^2] \geq 0$ in our quantum geometry; this leads to a "maximal acceleration". Cf.ref. 6 above.

Received by Publishing Department
on December 22,1983.

SUBJECT CATEGORIES
OF THE JINR PUBLICATIONS

Index	Subject
1.	High energy experimental physics
2.	High energy theoretical physics
3.	Low energy experimental physics
4.	Low energy theoretical physics
5.	Mathematics
6.	Nuclear spectroscopy and radiochemistry
7.	Heavy ion physics
8.	Cryogenics
9.	Accelerators
10.	Automatization of data processing
11.	Computing mathematics and technique
12.	Chemistry
13.	Experimental techniques and methods
14.	Solid state physics. Liquids
15.	Experimental physics of nuclear reactions at low energies
16.	Health physics. Shieldings
17.	Theory of condenced matter
18.	Applied researches
19.	Biophysics

Кайанелло Э.Р. E17-83-874
Геометрическая "идентификация" квантовой и информационной теорий

На основе концепции кросс-энтропии исследована взаимосвязь квантовой и информационной теорий.

Работа выполнена в Лаборатории теоретической физики ОИЯИ.

Препринт Объединенного института ядерных исследований. Дубна 1983

Caianiello E.R. E17-83-874
Geometrical "Identification" of Quantum and Information Theories

The interconnection between quantum and information theories is investigated being based on the conception of the cross-entropy.

The investigation has been performed at the Laboratory of Theoretical Physics, JINR.

Preprint of the Joint Institute for Nuclear Research. Dubna 1983