# Quantum theory as inductive inference

Ryszard Paweł Kostecki

*Institute of Theoretical Physics, University of Warsaw, Hoża 69, 00-681 Warszawa, Poland*

**Abstract.** We present a new approach to the foundations of quantum theory and information theory which is based on the algebraic approach to integration, information geometry, and maximum relative entropy methods. It enables us to deal with conceptual and mathematical problems of quantum theory without any appeal to Hilbert space framework and without frequentist or subjective interpretation of probability.

**Keywords:** quantum theory, information geometry, inductive inference, maximum entropy
**PACS:** 89.70.Cf 02.50.Cw 03.67.-a 03.65.-w

The notion of information, quantified by entropy, is usually defined using the notion of probability [31]. But Ingarden and Urbanik [20] showed that the notion of information is independent of the notion of probability, and that the latter can be defined by the former. Moreover, probability theory can be considered as a special case of quantum theory [29]. So, what is the relationship between information theory and quantum theory? Considering information theory as more fundamental than quantum theory requires founding it independently of the notion of probability, and equipping it with the dynamical structures sufficient to recover dynamics of quantum theoretic models.

We define the quantitative inductive inference theory as a theory of information dynamics, independent of probability theory. This calls for an independent mathematical formulation of information kinematics and information dynamics, equipped with an independent semantic justification. We replace the *information kinematics* based on probability measures on commutative countably additive algebras of subsets of a given space by one based on spaces of finite integrals on abstract commutative or non-commutative (n-c) algebras. This change has two reasons: (1) the conflict between the Bayes–Laplace (BL) and the Borel–Kolmogorov (BK) approaches to mathematical foundations of the probability theory can be resolved by the Huygens–Whittle (HW) approach, which is based on the Daniell–Stone (DS) theory of integrals on commutative algebras; (2) the analogous approach to mathematical foundations of quantum theory, provided in terms of integrals (normal algebraic states) on n-c $W^*$-algebras avoids several important problems of Hilbert space based approach, including the problem of representation-dependent separation of the Hilbert space vectors into amplitudes and phases that has troubled the attempts to consider the Hilbert space approach to quantum theory as an extension of probability theory. Next we find an information theoretic justification for a unique choice of the measure of relative information distance. This way we derive the constrained maximum relative entropy updating as a unique method of quantitative inductive inference. Previous derivations of uniqueness, based on the work of Shore and Johnson [32], are dependent on the notion of probability which in turn requires its own additional semantic derivation, e.g. following [10], and they do not extend to n-c case.

Our derivation does not refer to the notion of probability, but is instead based on the information geometry on the mathematical level and on the information semantics on the conceptual level. We define the *information dynamics* as a mapping on the space of finite integrals provided by the constrained maximisation of relative entropy. According to *information semantics*, the spaces of finite integrals (models of information), represent *quantified knowledge*, the entropic updating (dynamics of information models) represents *quantitative inductive inference*, while the underlying algebra represents an *abstract qualitative language* subjected to quantitative evaluation.

Finally, we apply this framework to the spaces of integrals on integrable n-c algebras (that is, to the spaces $N_*^+$ of normal finite positive linear functionals on $W^*$-algebras), recovering and generalising the existing settings of quantum theory. In [24] we derive the ordinary description of temporal behaviour of quantum theoretic models specified by the perturbations of generators of unitary dynamics from the entropic updating, using the Tomita–Takesaki theory. This way we show that the quantum theory in its ordinary setting is nothing else than a particular case of quantitative inductive inference (information dynamics). This closely follows the ideas of Jaynes on the character of quantum theory and the ideas of Ingarden on the foundational role of the information theory and information geometry in quantum theory. As opposed to existing 'objective bayesian' (e.g., [8]) and 'subjective bayesian' (e.g., [16]) derivations and/or interpretations of elements of the formalism of quantum theory, our derivation is completely independent of the notions and semantics of probability and Hilbert spaces, as well as of the choice of a particular Hilbert space representation. Our approach to foundations of quantum theory replaces the use of geometry of commutative $L_2$ spaces (which quantitatively represent the abstract Hilbert spaces) by the use of information geometry of n-c $L_p$ spaces (which quantitatively represent the abstract $N_*^+$ spaces). Hence, specification of quantitative results of experimental procedures in terms of the semispectral measures within a fixed Hilbert space representation of the chosen operator algebra is replaced by the specification in terms of algebraic states and their n-c $L_p$ space representations, while the description of temporal behaviour of these quantitative results in terms of unitary or completely positive mappings is replaced by the constrained relative entropic updating of algebraic states. This procedure is described in terms of information geometry as a non-linear projection. Its constraints, specified by nonempty closed convex sets of n-c $L_p$ spaces, define the domain of this projection. Quantum information channels become replaced by the (algebraic) quantum histories of constrained updatings [24]. As opposed to the ordinary settings of quantum information and algebraic quantum theories, our approach can be directly used to construct 'interacting' (predictive, experimentally verifiable) models of quantum field theory. This way it provides the extension of the approaches of Jaynes and Ingarden, who showed that statistical mechanics and thermodynamics can be derived as particular instances of probability (or information) theory.

**The foundations of information theory**

The BL and BK approaches to probability theory can be viewed as extreme cases of an application of two competing principles: evaluational (kinematical) and relational (dynamical). The notion of probability in BL approach is specified by the *conditional probability map* $\mathscr{B} \ni A \mapsto p(A|I) \in [0,1]$, which assigns a quantitative value to an element ('statement') of a Boole or Heyting algebra $\mathscr{B}$ under condition that the reference statement $I \in \mathscr{B}$ is true. This notion is essentially relational, and it is justified by an appeal to relational frameworks, such as Bayes' rule, de Finetti's type derivations (c.f. [16]), or the Cox's type derivations [10]. The notion of probability in BK approach is specified by the *probability measure* on the countably additive algebra of subset of a given space. This notion is essentially evaluational, and it is justified by an appeal to the evaluational framework provided by measure theory. The HW approach [34] to the foundations of probability theory is based on probabilistic expectations *and* conditional expectations on commutative algebras. For a given space $\mathscr{X}$ the *linear lattice* is defined as a commutative algebra $\mathscr{C}$ such that: i) it is a subspace of a space of functions $f : \mathscr{X} \to \mathbb{R} \cup \{\infty\}$, ii) $f \in \mathscr{C} \Rightarrow |f| \in \mathscr{C}$, and iii) $(f \equiv \text{const}) \in \mathscr{C}$, while the DS *integral* is defined as a map $\omega : \mathscr{C} \to \mathbb{R}$ that is positive, linear, and monotone ($\omega(\inf(u_n)) = \inf(\omega(u_n))$ for every convergent monotone sequence $\{u_n\}$ with $\inf u_n \in \mathscr{C}$). When normalised, $\omega$ satisfies all properties of the probabilistic expectation. The *conditional expectation* is then defined as a function $\mathscr{E}_\omega(\cdot|g) \in \mathscr{C}$ such that $\omega((f - \mathscr{E}_\omega(f|g))h(g)) = 0\ \forall f \in \mathscr{C}\ \forall h : \mathbb{R} \to \mathbb{R}$ where $h(g) \in \mathscr{C}$. $\mathscr{E}_\omega$ defines uniquely the 'updated' (conditioned) expectation by $\omega \mapsto \omega_{\text{new}} := \omega(\cdot|\mathscr{E}_\omega) := \omega \circ \mathscr{E}_\omega$. The *probability* and *conditional probability* are defined, respectively, by $p(A) := \omega(\chi_A)$ and $p(A|B) := \omega(\chi_A|\chi_B)$, where $\chi_A(x)$ is a characteristic function on the (finitely or countably) additive algebra. The DS and the Radon–Bourbaki integration theories are equivalent. The setting of BK (resp., BL) approach is recovered by restriction to the evaluational (resp., relational) component, via expectations (resp., conditioned expectations) on $\chi_A$'s on countably (resp., finitely) additive algebras. The HW approach treats both components as equally fundamental.

This shows that the DS integration theory on commutative algebras is an appropriate candidate for the mathematical framework of information theory built without reference to the notion of probability. However, the DS theory lacks the relational notions, and there is also an important disproportion between the kinematic and dynamic structures of HW approach: while the expectation is a normalised canonical DS integral, the definition of a conditional expectation is not canonical. If $\omega(f^2) < \infty$, then the latter is equivalent with $\mathscr{E}_\omega(f|g) := \arg\inf_{f_e \in \mathscr{C}} \omega((f - f_e(f,g))^2)$, but there is still no justification given of a choice of such functional. Considering the updating $[\omega, f] \mapsto [\omega \circ \mathscr{E}_\omega(\cdot|g), f]$ instead of $[\omega, f] \mapsto [\omega, \mathscr{E}_\omega(f|g)]$ makes us wonder to what extent the definition of a map $\omega \mapsto \omega_{\text{new}}$ can be generalized or changed and under which conditions this new map can be unique. We consider the space $M_+$ of all finite DS integrals as a framework for the kinematics of models of commutative information dynamics theory. In order to propose a general framework for dynamics of this theory, we have to replace the relational component of HW approach by minimisation of some uniquely characterised functional on $M_+$.

This is possible for the functionals of the relative distance between the integrals, implemented by the *deviation maps* $D : M_+ \times M_+ \in (\omega, \phi) \mapsto D(\omega, \phi) \in [0, +\infty)$ such that $D(\omega, \phi) \geq 0$ and $D(\omega, \phi) = 0$ iff $\omega = \phi$. Deviation is one of two founding notions of the information geometry theory [1], [2]. The second is the structure of the differential manifold that is imposed on $M_+$ by the local coordinate systems given by embeddings into appropriate Banach spaces (in general, $L^{\Phi_1}$ Orlicz spaces for $\gamma \in \{0,1\}$ and $L_{1/\gamma}$ spaces for $\gamma \in (0,1)$), see [1], [17]. We will use the framework of information geometry to select a unique deviation functional. By convenience, we will pass on description in terms of measures. Eguchi showed, for $\dim \mathscr{X} < \infty$, that any deviation that is symmetric in first derivatives and has a negative definite hessian determines uniquely (up to a scalar factor) the riemannian metric and a pair of two affine connections on $M_+$, defined by $g_\mu(u,v) := -(\partial_u)_\mu(\partial_v)_\mu D(\nu,\mu)|_{\nu=\mu}$, $g_\mu(u\nabla_\mu v,w) := -(\partial_u)_\mu(\partial_v)_\mu(\partial_w)_\nu D(\nu,\mu)|_{\nu=\mu}$, and $g_\mu(v,u\nabla^*_\mu w) := -(\partial_u)_\nu(\partial_w)_\nu(\partial_v)_\mu D(\nu,\mu)|_{\nu=\mu}$. If the torsion and Riemann curvature tensors of both these connections are equal to zero, then $(M_+,g,\nabla,\nabla^*)$ is called a *dually flat manifold* and there exists a pair $(\theta,\eta)$ of coordinate systems on $M_+$, called *dually flat coordinates*, such that $\theta$ consist of $\nabla$-geodesics, while $\eta$ consists of $\nabla^*$-geodesics (they are called *orthogonal* at $q \in M_+$ iff $g_q((\partial_\theta)_q,(\partial_\eta)_q) =: [\theta(\eta),\eta(q)] = 0$). A point $p_Q \in Q \subset M_+$ is called a $\nabla$*-projection* of $p \in M_+$ iff the $\nabla$-geodesic connecting $p$ and $p_Q$ is orthogonal to $Q$ w.r.t. $g$, while $Q \subset M$ is called $\nabla$*-convex* iff $\forall p_1,p_2 \in Q$ $\exists!$ $\nabla$-geodesic connecting $p_1$ and $p_2$ and entirely included in $Q$. If $Q$ is a $\nabla^*$-convex closed set then the unique $\nabla$-projection of $p \in M_+$ on $Q$ is given by $Q \ni p_Q = \arg\inf_{q\in Q} D(p,q)$ [1]. Hence, the dual flatness of the geometry of $M_+$ is a necessary condition for the existence of the unique $\nabla$-projections on the $\nabla^*$-convex subspaces provided by the minimum of the deviation functional. Let $P(p)$ be a probability measure on $M_+$, let $\ell$ be some coordinate map on $M_+$, and let $P(p).f(p) := \int_{p\in M_+} P(p)f(p)$. Define a $(D,P)$*-optimal estimate* w.r.t. $Q \subset M_+$ as $\arg\inf_{q\in Q} P(p).D(p,q) =: q_{e,Q} \in Q$, a $(D,P)$*-ideal estimate* as $(D,P)$-optimal estimate w.r.t. $M_+$, and an $(\ell,P)$*-average* of $p$ as $\ell^{-1}(P(p).\ell(p)) =: \langle p\rangle^\ell_P \in M_+$. If $M_+$ is dually flat and if $(\ell,\ell^*)$ are its dually flat coordinates, then $(D,P)$-ideal estimate is equal to $(\ell,P)$-average [37], [15]. Moreover, if $D$ satisfies the *generalised cosine equation* $D(p_1,p_2)+D(p_2,p_3)-D(p_1,p_3) = [\ell(p_1)-\ell(p_2),\ell^*(p_3)-\ell^*(p_2)]$, then the *error decomposition theorem* $P(p).D(p,q) = P(p).D(p,\langle p\rangle^\ell_P) + D(\langle p\rangle^\ell_P,q)$ holds. Zhu and Rohwer [37] proved this for the family of $\gamma$*-deviations*, defined by [37] $-\mathbf{S}_\gamma(\mu,\nu) := D_\gamma(\mu,\nu) := \int(\mu/(1-\gamma)+\nu/\gamma-\mu^\gamma\nu^{1-\gamma}/(\gamma(1-\gamma)))$ for $\gamma\in(0,1)$ and by $-\mathbf{S}_{KL}(\mu,\nu) := D_1(\mu,\nu) := \int(\nu+\mu+\mu\log(\frac{\mu}{\nu})) = D_0(\nu,\mu)$ for $\gamma\in\{0,1\}$, with the corresponding dually flat coordinates given by *Amari* $\gamma$*-embeddings* $\ell_\gamma : M_+ \mapsto \omega^\gamma/\gamma \in L_{1/\gamma}$ and $\ell^* = \ell_{1-\gamma}$ for $\gamma \in (0,1)$. The setting of [15] enables to generalise this proof to the family of *Bregman deviations* [6], characterised as deviations satisfying the generalised cosine equation and generating the dually flat geometry [2]. The family of *Csiszár deviations* [11] is characterised by nonincreasing under the information loss provided by partitioning of the space $\mathscr{X}$ and invariance under permutations of partitions [12]. Csiszár [13] showed that $\mathbf{S}_{KL}$ are the only permutation of these $\mathscr{P} := M_+ \cap \{\omega|\omega(\mathbb{I}) = 1\}$ that are both Csiszár and Bregman deviations. Amari [3] showed that $D_\gamma$ are the only deviations on $M_+$ satisfying this condition.

This allows us to propose an information theoretic characterisation of the family $D_\gamma$ on $M_+$ for $\gamma \in [0,1]$ (and $\mathbf{S}_{KL}$ on $\mathscr{P}$) based on three conditions: nonincreasing under the loss of information, existence of the unique minimisers for the projections on the nonempty closed convex subspaces, and existence of ideal and optimal estimates that satisfy the error decomposition theorem. On the base of this characterisation, we propose a new approach unifying the probability and information theories in one information dynamics theory. Its evaluational component is given by the finite DS integrals. Its relational component is given by the *constrained maximum relative* $\gamma$*-entropy updating rule*, defined as a map between two finite DS integrals $M_+ \ni p \mapsto p_{\mathrm{new}} := \arg\inf_{q\in Q}\left(\int_{p'\in M_+} E(p')D_\gamma(p',q)+F(q)\right) \in M_+$, where $E$ is a finite integral *on* $M_+$, while $F : Q \to (-\infty,+\infty]$ is a constraining function. This map provides a selection of

$\gamma$-optimal estimates w.r.t. $Q$ that are weighted by $E$ and satisfy the constraints $F$. If $F$ is lower semicontinuous and $(1-\gamma)$-convex, if $E(p')=\delta(p-p')$, and if the above infimum is finite, then the updating procedure selects a unique $p_{\text{new}}$.

Differences between $\gamma$-ideal estimates appear only in the third order of expansion of $D_\gamma$. Up to the second order all $D_\gamma$ are determined by the same riemannian metric, and the inferences based on $D_\gamma$ on $M_+$ agree with the inferences based on the Hilbert space $L_2$ norm (the $D_{1/2}$-estimates). The latter is useful, because the $L_2$ space always contains the $1/2$-ideal estimates, hence the problem of 'best' estimation always has a solution, as opposed to arbitrary $\gamma$-optimal estimates when restricted to arbitrary subspaces of $M_+$. Beyond the second order, the use of the $L_2$ space and its norm is *not* justified. Moreover, the projections in Hilbert space are linear operators, hence they can deal only with linear problems (constraints), as opposed to nonlinear $\gamma$-projections onto convex subspaces of $L_{1/\gamma}$ spaces. The constrained relative $\gamma$-entropy updating incorporates all Hilbert space methods as a special case. Under restriction to $\mathscr{P}$, the only justified method of inference is given by constrained maximisation of $\mathbf{S}_{KL}$. This solves the problem of the relationship between information and probability, made explicit by [20] and widely present in [21] and [22]. The probability theory is just a restriction of the information dynamics theory and requires no additional justification. The HW approach is recovered by the fact [5] that conditional expectations are characterised as minimisers $\mathscr{E}_\omega(f|g) := \arg\inf_{g\in\mathscr{C}} \omega(D_B(f,g))$ for Bregman deviations $D_B$, the BK approach is recovered by forgetting the relational component, and passing to probabilistic measures, while the BL approach is recovered by normalisation and by the fact [36] that the Bayes rule is just a special case of the entropic updating rule with Dirac's delta constraints.

**The foundations of quantum theory**

An algebraic approach to the mathematical framework of quantum theory [30] replaces the consideration of abstract Hilbert space by the consideration of abstract *$C^*$-algebra*, which is a Banach space $\mathscr{A}$ over $\mathbb{C}$ equipped with a map $^*:\mathscr{A}\to\mathscr{A}$ such that $(ab)^*=b^*a^*$, $(a+b)^*=a^*+b^*$, $(\lambda a)^*=\bar{\lambda}a^*$, $(a^*)^*=a$, $||a^*||=||a||$, $||ab||\leq||a||||b||$, $||\mathbb{I}||=1$, and $||a^*a||=||a||^2$. An *algebraic state* is a functional $\omega:\mathscr{A}\to\mathbb{C}$ that is linear, positive and normalised. A functional $\omega$ on $\mathscr{A}$ is called *normal* iff $\omega(\sup K)=\sup_{a\in K}\omega(a)$ for every directed filter $K\subset\mathscr{A}$ with the upper bound $\sup K$. A *$W^*$-algebra* is defined as such $C^*$-algebra $N$ which is a Banach dual to some Banach space $N_*$. This algebra directly generalises the notion of an 'integrable' commutative algebra (a linear lattice in the DS theory) to the n-c case. The space $N_*$ is a space of all normal functionals on a $W^*$-algebra $N$. The spaces $N_*^+ =: N_*\cap\{\omega|\omega(A^*A)\geq 0\}$ (resp., the spaces $N_{*1}^+$ of normal algebraic states) on $N$ generalise the spaces $M_+$ of finite (resp., the spaces $\mathscr{P}$ of normalised) DS integrals on $\mathscr{C}$. For any pair of a $C^*$-algebra $\mathscr{A}$ and linear, positive functional $\omega$ on $\mathscr{A}$, the Gel'fand–Naĭmark–Segal (GNS) theorem associates a unique Hilbert space $H_\omega$ and a unique (up to unitary equivalence) representation $\pi_\omega:\mathscr{A}\to\mathfrak{B}(H_\omega)$ such that $\exists\Omega_\omega\in H_\omega$ that is cyclic for $\pi_\omega(\mathscr{A})$, $||\Omega_\omega||^2=1$, and $\omega(A)=\langle\Omega_\omega,\pi_\omega(A)\Omega_\omega\rangle_\omega$ $\forall A\in\mathscr{A}$. This theorem enables to recover the Hilbert space based framework from the n-c generalisation of DS theory. The evaluation of $\text{tr}(\rho\,\cdot)$ on $\mathfrak{B}(H)$ and the evaluation of probability measure $\text{tr}(\rho\text{E}(\cdot))$ on spectrum of a given operator are replaced by acting with $\omega(\cdot)\in N_*^+$ on $N$. Because semispectral measures $\text{E}(\cdot)$ no longer define the probabilistic evaluations, the von Neumann's spectral theorem loses its importance, and the associated problems of the uniqueness and contextuality of definition of operators corresponding to particular quantitative experimental results and descriptions [7] disappear. The only meaning of elements of $W^*$-algebras is: they are the *statements* of an *abstract language* subjected to quantitative evaluation.

Now we need to equip this general setting with the information geometric and dynamical structures. Falcone and Takesaki [14] formulated the canonical theory of n-c $L_p(N)$ spaces on von Neumann algebras $N$ (i.e., on the concrete instances of abstract $W^*$-algebras). The equivalence relation $(x,\omega)\sim_t(y,\phi)\iff y=x[\mathrm{D}\omega:\mathrm{D}\phi]_t$ defines the space $N(t):=N\times N^+_{*0}/\sim_t$, where $N^+_{*0}:=N^+_*\cap\{\omega|\omega(A^*A)=0\Rightarrow A=0\}$, and $[\mathrm{D}\omega:\mathrm{D}\phi]_t$ is Connes' cocycle (a n-c generalisation of the Radon–Nikodým derivative). The elements of $N(t)$ are denoted by $x\omega^{it}$. The action of an algebra of sections $\tilde{x}:\mathbb{R}\ni t\mapsto x(t)\omega^{it}\in\prod_{t\in\mathbb{R}}N(t)$ on a Hilbert space $H_c$ (which is defined by a suitable equivalence relation involving $N$ and $N^+_*$), defines the *core* von Neumann algebra $N_c$. The definitions of $N_c$ and $H_c$ do not depend on the choice of $\omega$, and the assigment $N\to N_c$ is functorial. The core $N_c$ is equipped with a canonical trace $\tau:N_c\to[0,+\infty]$. The algebraic structure of $L_p(N)$ is defined as a space of all closed densely defined positive $\tau$-measurable operators $A$ affiliated with $N_c$ such that $\theta_s(A)=e^{-s/p}A$, where $\theta_s(x\omega^{it}):=e^{-ist}x\omega^{it}$. Authors of [14] define the integral $\int:L_1(N)\to\mathbb{C}$ such that, in particular, $\int x\omega=\omega(x)\in\mathbb{C}$. This integral is meaningful also for such formulas like $\int\omega^{it}y\phi^\gamma\varphi^\alpha abc$, as long as the powers of functionals sum up to 1. The bilinear map $[x,T]:=\int xT$ gives the pairing between $N$ and $L_1(N)$ which identifies $L_1(N)$ with $N_*$. The duality for $S\in L_p(N)$, $T\in L_q(N)$, $1/p+1/q=1$, reads $[S,T]=\int ST$. The analytic structure of $L_p(N)$ is defined by the Banach norm $||T||_p:=(\int|T|^p)^{1/p}$.

Our novel result is a generalisation of quantum information geometric structures which are based on Falcone–Takesaki theory. In analogy with [37], we define the *quantum $\gamma$-embeddings* ($\gamma$-coordinates) by $\ell^\gamma:N^+_*\ni\omega\mapsto\ell^\gamma(\omega):=\omega^\gamma/\gamma\in L_{1/\gamma}(N)$, and define the *quantum relative $\gamma$-deviation* $D_\gamma:N^+_*\times N^+_*\ni(\omega,\phi)\mapsto D_\gamma(\omega,\phi)\in\mathbb{R}$ by

$$D_\gamma(\omega,\phi):=\mathrm{Re}\int\left(\frac{\omega}{1-\gamma}+\frac{\phi}{\gamma}-\ell^\gamma(\omega)\ell^{1-\gamma}(\phi)\right)=\mathrm{Re}\int\left(\frac{\omega}{1-\gamma}+\frac{\phi}{\gamma}-\frac{\omega^\gamma\phi^{1-\gamma}}{\gamma(1-\gamma)}\right),$$

for $\gamma\in(0,1)$, and by the limits under integral sign for $\gamma\in\{0,1\}$. The *quantum relative $\gamma$-entropy* is defined as $\mathbf{S}_\gamma(\omega,\phi):=-D_\gamma(\omega,\phi)$, and $-\mathbf{S}(\phi,\omega):=D_0(\omega,\phi)=D_1(\phi,\omega)$. This definition reduces to the Jenčová–Ojima [23], [27] quantum $\gamma$-deviation for any particular choice of unitary representation of action of $N_c$ on $H_c$ in terms of action of $N\rtimes_{\sigma^\omega}\mathbb{R}$ on $H_\omega\otimes L_2(\mathbb{R},dt)$, to the Umegaki–Araki quantum relative entropy [33], [4] in the Petz [28] form $\lim_{t\to+0}\frac{i}{t}\phi([\mathrm{D}\omega:\mathrm{D}\phi]_t-\mathbb{I})$ for $\omega,\phi\in N^+_{*1}$, to the Hasegawa quantum $\gamma$-deviation [19] $\mathrm{tr}(\rho_\phi-\rho^\gamma_\omega\rho^{1-\gamma}_\phi)/(\gamma-\gamma^2)$ for $\omega=\mathrm{tr}(\rho_\omega\cdot)$ and $\phi=\mathrm{tr}(\rho_\phi\cdot)$, and the Zhu–Rohwer $\gamma$-deviation $D_\gamma$ in commutative case. We generalise the *$\gamma$-representation* of tangent spaces to $T_\omega N^+_*\ni v\mapsto\ell^\gamma_{*(\omega)}(v):=x\phi^\gamma\in T_{\omega^\gamma}L_{1/\gamma}(N)$, where $T_\omega L_{1/\gamma}(N):=\{x\phi^\gamma\in L_{1/\gamma}(N)|\mathrm{Re}\int\omega^{1-\gamma}x\phi^\gamma=0\}=\{x\phi^\gamma\in L_{1/\gamma}|\mathrm{Re}\omega([\mathrm{D}\omega:\mathrm{D}\phi]_{i\gamma}(x))=0\}$. In the special cases, this definition reduces to those given in [23], [18], and [17]. It allows us to define a *quantum $\gamma$-metric* and a *quantum $\gamma$-connection* for $u,v,w\in N^+_*$ as $g^\gamma_\omega(u,v):=\mathrm{Re}\int\ell^\gamma_{*(\omega)}(u)\ell^{1-\gamma}_{*(\omega)}(v)=\mathrm{Re}\int\omega^{1-\gamma}x\omega^\gamma y\in\mathbb{R}$ and

$\nabla^\gamma_\omega(u,v)w := (\gamma-\alpha)\omega^{1-3\alpha}\ell^\alpha_{*(\omega)}(u)\ell^\alpha_{*(\omega)}(v)\ell^\alpha_{*(\omega)}(w) \in N^+_*$, generalising the Wigner–Yanase–Dyson $\gamma$-metric [35], [19] and quantum $\gamma$-connections of [18], respectively.

The properties of $D_\gamma$ are the same as in [23]. For $\gamma \in (0,1)$, $D_\gamma$ is both Petz' $f$-deviation [28] (which is a n-c generalisation of Csiszár's deviation) and the Jenčová deviation [23] (which is a n-c generalisation of Bregman's deviation). $D_\gamma$ (resp., $\mathbf{S}$) is convex and lower semicontinuous on the space $N^+_* \times N^+_{*0}$ (resp., $N^+_* \times N^+_*$) endowed with the product of norm topologies (resp., the topology of pointwise convergence on $N_*$ [26]). If $Q \subset N^+_*$ is weakly closed and convex in terms of the $\gamma$-coordinates, then $\exists!$ projection $N^+_* \ni \omega \mapsto \omega_Q = \arg\inf_{\phi\in Q} D_{1-\gamma}(\omega,\phi) \in Q$. The space $L_2(N)$ equipped with the inner product $\langle S,T\rangle := \int T^*S$ is a Hilbert space that can be identified with the GNS Hilbert space $H_\omega$ for any choice of $\omega \in N^+_*$, with $1/2$-deviation taking the form $D_{1/2}(x,y) = \frac{1}{2}||x-y||^2$. The $1/2$-projection is a minimiser of the Hilbert space norm. Hence, we recover all Hilbert space geometry as a special case of quantum information geometry. Following Amari's and Csiszár's results, *we conjecture that $D_\gamma$ (resp., $\mathbf{S}$) is a unique deviation on $N^+_*$ (resp., $N^+_{*1}$) that belongs to both Jenčová and Petz families of quantum deviations*. Using Bauer's maximisation principle and the results in [26] and [23], we conclude that if $Q$ is a nonempty $(1-\gamma)$-convex weakly closed subspace of $N^+_*$, $\omega \in N^+_*$, and $F : Q \to (-\infty,+\infty]$ is a weakly-* lower semicontinuous $(1-\gamma)$-convex function, then $\exists!\arg\inf_{\phi\in Q}\{D_\gamma(\omega,\phi)+F(\phi)\} =: \omega_{Q,F}$ if this infimum is finite.

We define the *quantum constrained maximum relative $\gamma$-entropy updating rule* by

$$N^+_* \ni \omega \mapsto \arg \inf_{\phi\in Q\subset N^+_*} \left\{ \int_{\varphi\in N^+_*} E(\varphi) D_\gamma(\varphi,\phi) + F(\phi) \right\} \in N^*_+,$$

for the Borel measure $E$ on $N^+_*$, and constraints given by convex weak-* lower semicontinuous $F : Q \to ]-\infty,+\infty]$, which guarantees the existence and uniqueness of the result of this updating when this infimum takes a finite value and if $E(\varphi) = \delta(\varphi-\omega)$. This rule equips the evaluational theory of integrals on n-c algebras with the relational counterpart, which provides a selection of $\gamma$-optimal estimates that are weighted by $E$, relative to $Q$, and constrained by $F$. This generalises and replaces the Hilbert space approach. The Hilbert spaces are replaced by the convex subspaces of n-c $L_\gamma$ spaces used as the codomain of embeddings of the functionals $\omega \in N^+_*$. If the constraints $F(\phi) = c(t)$ depend on 'time' parameter $t$ (discrete or continuous), then the resulting trajectory on quantum information manifold can be understood as a *dynamics of quantum states*, which is non-linear and non-unitary. Interpretation of these spaces as a general framework for nonlinear quantum theoretic models and interpretation of the above updating rule as a general description of their nonlinear temporal behaviour (quantum dynamics) opens new perspectives for the foundations and applications of quantum theory. Our conjecture equips it with a justification based on simple information theoretic postulates, which provides a concrete reply to requests of [16] and [9]. This way quantum theory becomes a theory of quantitative intersubjective inductive inference provided w.r.t. the qualitative abstract language of n-c algebras. The constraints of this inductive logic are imposed by the intersubjective decisions of choice of $\gamma$, $Q$, $E$ and $F$.

On the conceptual level, our approach strips quantum theory, information theory and probability theory from any ontological assumptions. An abstract algebra $N$ subjected to quantitative evaluation (integration) is understood as a qualitative abstract language used as a common reference in an intersubjective communication. The finite integrals of $N^+_*$ are just the carriers of quantitative intersubjective knowledge. The quantitative information models formed by the choices of $Q$ and $E$ (in particular, an 'initial state' $\omega$) as well as the quantitative dynamical models formed by the additional choices of $\gamma$ and $F$ (in particular, an 'actual history' $F(t)$) are just the results of intersubjective agreement with respect to the method of processing the quantitative results and descriptions of some experimental procedures. The ontological beliefs usually introduced by the use of such terms as 'matter', 'fields', 'particles', 'quantum objects', 'randomness', 'universe', 'nature', etc. are completely irrelevant for this framework. As follows from the above formulation, quantum theory is just a particular instance of intersubjective quantitative inductive inference (information dynamics) theory, and nothing else matters.

## REFERENCES


1. Amari S.-i., 1985, *Differential–geometrical methods in statistics*, Springer, Berlin.
2. Amari S.-i., Nagaoka H., 1993, *Joho kika no hoho*, Iwanami Shoten, Tokyo.
3. Amari S.-i., 2009, IEEE Trans. Inf. Theor. **55**, 4925.
4. Araki H., 1976, Publ. Res. Inst. Math. Sci. Kyōto Univ. **11**, 809; 1977, **13**, 173.
5. Banerjee A., Guo X., Wang H., 2005, IEEE Trans. Inf. Theor. **51**, 2664.
6. Bregman L.M., 1967, Zh. Vychestl. Matem. Matem. Fiz. **7**, 620.
7. Busch P., Grabowski M., Lahti P.J., 1995, *Operational quantum physics*, Springer, Berlin.
8. Caticha A., 1998, Phys. Lett. A **244**, 13; 1998, Phys. Rev. A **57**, 1572; 2000, Found. Phys. **30**, 227.
9. Clifton R., Bub J., Halvorson H., 2003, Found. Phys. **33**, 1561.
10. Cox R.T., 1946, Am. J. Phys. **14**, 1.
11. Csiszár I., 1963, Magyar Tud. Akad. Mat. Kutató Int. Közl. **8**, 85.
12. Csiszár I., 1978, in: Kozesnik J. (ed.), Trans. 7th Prague Conf. Inf. Th., p.73.
13. Csiszár I., 1991, Ann. Stat. **19**, 2032.
14. Falcone T., Takesaki M., 2001, J. Funct. Anal. **182**, 170.
15. Frigyik B.A., Srivastava S., Gupta M.R., 2008, IEEE Trans. Inf. Theor. **54**, 5130.
16. Fuchs C., 2003, J. Mod. Opt. **50**, 987.
17. Gibilisco P., Pistone G., 1998, Inf. Dim. Anal. Quant. Prob. Relat. Top. **1**, 325.
18. Gibilisco P., Isola T., 1999, Inf. Dim. Anal. Quant. Prob. Relat. Top. **2**, 169.
19. Hasegawa H., 1993, Rep. Math. Phys. **33**, 8.
20. Ingarden R.S., Urbanik K., 1961, Bull. Acad. Sci. Pol. Sci. Math. **9**, 313; 1962, Colloq. Math. **9**, 131.
21. Jaynes E.T., 1957, Phys. Rev. **106**, 620.
22. Jaynes E.T., 2003, *Probability theory: the logic of science*, Cambridge University Press, Cambridge.
23. Jenčová A., 2005, Inf. Dim. Anal. Quant. Prob. Relat. Top. **8**, 215; 2006, J. Funct. Anal. **239**, 1.
24. Kostecki R.P., 2010, *Algebraic quantum histories*, *Information dynamics I, II, III*, in preparation.
25. Kullback S., 1959, *Information theory and statistics*, Wiley, New York.
26. Ohya M., Petz D., 1993, *Quantum entropy and its use*, Springer, Berlin.
27. Ojima I. 2004, Publ. Res. Inst. Math. Sci. Kyōto Univ. **40**, 731.
28. Petz D., 1985, Publ. Res. Inst. Math. Sci. Kyōto Univ. **21**, 787; 1986, Acta Math. Hung. **47**, 65.
29. Rédei M., Summers S.J., 2007, Stud. Hist. Phil. Mod. Phys. **38**, 390.
30. Segal I.E., 1947, Ann. Math. **48**, 930; 1953, **57**, 401.
31. Shannon C., 1948, Bell Syst. Tech. J. **27**, 379, 623.
32. Shore J.E., Johnson R.W., 1981, IEEE Trans. Inf. Theory **27**, 472; **27**, 472; 1983, **29**, 942.
33. Umegaki H., 1962, Kôdai Math. Sem. Rep. **14**, 59.
34. Whittle P., 1970, *Probability*, Penguin, Harmondsworth.
35. Wigner E., Yanase M., 1963, Proc. Nat. Acad. Sci. U.S.A. **49**, 910.
36. Williams P.M., 1980, Brit. J. Phil. Sci. **31**, 131.
37. Zhu H., Rohwer R., 1997, in: Ellacott S.W. et al. (eds.), *Mathematics of neural networks*, Kluwer, Dordrecht, p.394; Zhu H., 1998, Santa Fe Inst. Tech. Rep. **98-06-44**, **98-06-45**.