Quantum information geometry

Ryszard Paweł Kostecki

Perimeter Institute of Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario N2L 2Y5, Canada[*]
Institute of Theoretical Physics, University of Warsaw, Hoża 69, 00-681 Warszawa, Poland

ryszard.kostecki@fuw.edu.pl

December 26, 2014

Abstract

This text is a detailed overview of the quantum and classical information geometry, containing several new concepts and some new results. The key role played by the relative entropy is exposed, and the interconnections between various structures are analysed. We consider the convex/variational nonsmooth part of the theory on the equal footing with the smooth/infinitesimal part, and we also consider the duality principles (as embodied in the concepts of Brègman relative entropy and dually flat smooth geometries) on the equal footing with monotonicity under quantum channels. All results are spelled out with the maximal available generality, so the functional analytic setting of W^*-algebras and Banach spaces is widely used.

This text is a draft (0.89?) version under construction. Several sections contain various mistakes, ommisions, and typos that will be improved in the next version

Contents

1 Introduction

In this paper we will discuss information geometric structures on arbitrary dimensional spaces of quantum states. It is, to a large extent, a systematic review of already known results, but several new results are also proved, and some new concepts are introduced. Our treatment is unique in its perspective, considering the subsets of preduals of W^*-algebras and nonsymmetric distance functions on them as the foundational setting for quantum geometry. Due to out consideration of analytic and geometric aspects of information geometry on the equal footing, this text can be considered as a quantum followup to the works [152, 507].

[*]Current affiliation.

Section 2 is devoted to information models and their categories. In Sections 2.1 and 2.2 we recall main notions of integration theory and of its quantum generalisation, which allows to consider quantum states as integrals over noncommutative W^*-algebras. These Sections are based on the material in [404]. They contain two new results: a construction of a family of $L_p(\mathcal{A})$ spaces canonically associated with any localisable boolean algebra (this is based on an application of Lewin–Lewin generalisation of Segal's generalisation of Radon–Nikodým theorem) and a proof of equivalence of categories of commutative W^*-algebras and proper abstract L_∞ spaces that does not use their topological representation. First of these results sovels a problem posed by Zhu in [788] (see also [635]), while the second provides an elegant justification of the reasoning carried in Section 2.3. In Section 2.3 we define *quantum models* as *arbitrary* subsets $\mathcal{M}(\mathcal{N}) \subseteq \mathcal{N}_\star^+$ of the positive cones $\mathcal{N}_\star^+$ of Banach preduals $\mathcal{N}_\star$ of arbitrary W^*-algebras $\mathcal{N}$, and show that this definition provides a natural generalisation of the notion of a statistical model. In Section 2.4 we discuss the categories of information models (quantum and statistical) that are obtained by using Banach preduals of Markov morphisms (linear normal unital completely positive maps) as arrows.

Section 3 is devoted to the analysis of the fundamental notion of classical and quantum information geometry: the information distance. A *quantum distance* is defined as a function of pairs of elements $\phi, \psi \in \mathcal{M}(\mathcal{N})$ taking the values in $[0, \infty]$, and equal to 0 iff $\phi = \psi$. In Section 3.1 we discuss (as a review) the information distances which are *Markov monotone*, which means that they are nonincreasing under Banach preduals of Markov morphisms. In commutative case these are given by the Csiszár–Morimoto $\mathfrak{f}$-distances [176, 500], while in quantum case they are given by the Kosaki–Petz $\mathfrak{f}$-distances [397, 565]. In Section 3.2 we discuss (as a review) the families of *Brègman functionals* [118, 131, 61], whose characteristic feature is that they satisfy a generalisation of the pythagorean theorem under 'orthogonal' but nonlinear projections onto subsets. In Section 3.3 we introduce the notion of a dualistic Brègman distance, which allows to define the quantum Brègman distance, suitable for the setting of nonlinear infinite dimensional quantum models, with 'orthogonality' implemented by the Young–Fenchel duality associated with any dual pair of vector spaces. In order to obtain well defined uniqueness and composability properties, we further specify this duality to Banach dual pair of reflexive Banach spaces. In Section 3.4 we use the Falcone–Takesaki theory [247] of noncommutative $L_p(\mathcal{N})$ spaces over arbitrary W^*-algebras $\mathcal{N}$ to define the canonical family of quantum γ-distances, which generalises the Jenčová–Ojima family [551, 360] of quantum γ-distances and provides a noncommutative counterpart of the Liese–Vajda family [453] of γ-distances. We prove that this family belongs to an intersection of the Kosaki–Petz $\mathfrak{f}$-distances with the generalised Brègman distances. This is important in face of the result of Amari [19], who showed that in the finite dimensional commutative case the Liese–Vajda family *characterises* an intersection of the Csiszár–Morimoto $\mathfrak{f}$-distances with the Brègman distances. We conjecture that our family of γ-distances shares the same uniqueness property in the noncommutative case. We discuss also the conditions of existence, uniqueness, and stability of the solutions to the corresponding constrained distance minimisation problems.

Section 4 contains a review of those parts of geometry of smooth manifolds that are relevant for discussion of smooth information geometry. Apart from riemannian and affine geometries, we discuss much lesser known Norden–Sen, hessian, Eguchi, and Lauritzen geometries. The elementary setting of smooth manifolds modelled on Banach spaces is introduced.

Section 5 contains a review of smooth information geometry, in commutative and noncommutative cases, both for finite and infinite dimension. This presentation fills the gap in the literature of the subject, which lacks a fairly complete overview of its main results. In particular, we discuss Jenčová's construction [361, 362] of smooth Banach manifold structure on the space $\mathcal{N}_{\star 01}^+$ for an arbitrary W^*-algebra $\mathcal{N}$ (based on a specifically defined noncommutative Orlicz space overn $\mathcal{N}$), and characterisation [358, 359] of dually flat geometries on the spaces $\mathcal{N}_{\star 0}^+$ for type I_n W^*-algebras $\mathcal{N} \cong \mathrm{M}_n(\mathbb{C})$. Our novel contribution discussed in Section 5.1 is a construction of a general noncommutative Orlicz space associated with a n arbitrary W^*-algebra, that is based on the Falcone–Takesaki theory. We also propose an extension of Jenčová's construction to unbounded extended valued operators, which is intended to provide a quantum counterpart of the Pistone–Sempi manifold structure. Section 5.2 contains also a new result: an extension of the Nagaoka–Petz generalised pythagorean theorem [522, 523, 573] for Umegaki distance on $\mathcal{N}^+_{\star 01}$ to quantum γ-distances on $\mathcal{N}^+_{\star 0}$ with $\gamma \in [-1, 2]$. This extension is a straightforward consequence of known results, but it was left unnoticed, and it provides further evidence to our conjecture on the characterisation of quantum γ-distances. Apart from it, we once again apply the Falcone–Takesaki theory in order to generalise the Morozova–Chentsov–Petz family of Markov monotone quantum riemannian metrics to the infinite dimensional case, providing an answer to a question posed in [279].

Conventions

The following conventions are used in this text: 1) ***definitions***, «citations», 'notions subjectable to strict definition', "vague notions", and *attention markers*; 2) the mathematical style of text formatting (definition/theorem/proof) is used only for stating essentially new mathematical results; 3) whenever possible, we refer to original works containing results that are discussed or used; 4) the folk attributions of surnames to mathematical concepts are changed to historically correct ones whenever there is a definite evidence for the latter, and the concepts with attribution of three or four surnames to it are turned to an acronym after first use, while in the case of more authors we use only the descriptive naming of objects;[1] 5) for the Latin transliteration of the Cyrillic script (in references and surnames) we use the following modification of the system GOST 7.79-2000B: ц = c, ч = ch, х = kh, ж = zh, ш = sh, щ = shh, ю = yu, я = ya, ë = ë, ъ = ‘, ь = ’, э = è, й = ĭ, with an exception that surnames beginning with Х are transliterated to H.[2] For Russian texts: ы = y, и = i; for Ukrainian: и = y, і = i, ї = ï. All Cyrillic titles and names were transliterated from the original papers and books.

2 Information models

In Section 2.3 we discuss quantum and statistical information models, in order to show that an abstract definition of a quantum information model as an *arbitrary* subset $\mathcal{M}(\mathcal{N}) \subseteq \mathcal{N}^+_\star$ of the positive cone $\mathcal{N}^+_\star$ of Banach predual $\mathcal{N}_\star$ of arbitrary W^*-algebra $\mathcal{N}$ provides a natural generalisation of the notion of a statistical model. For this purpose, the terminology, notation and results of commutative and noncommutative integration theories will be introduced and discussed, respectively, in Sections 2.1 and 2.2. The exposition in these two sections is based on [404]. Following the principle that a complete specification of a mathematical theory requires a specification of a category that implements it, in Section 2.4 we will discuss the categories of quantum and statistical information models, with arrows given by *coarse grainings*, defined as Banach preduals of Markov maps (that is, normal unital linear completely positive functions). This category determines both the information models (as objects) *and* also the allowed methods of inference (as arrows), which in turn can be used to *characterise* the geometric structure of these models. The last idea was proposed and developed by Chencov and Morozova [151, 152, 153, 501, 502, 503, 504, 154, 505, 506, 507] and was a key principle guiding development of the smooth information geometry in both commutative and noncommutative case (see Section 5 for a detailed discussion). The definition of a quantum model $\mathcal{M}(\mathcal{N})$ does not introduce more specific constraints for the purpose of generality. In our opinion, various additional structures and assumptions (such as smoothness, convexity, finite-dimensionality, etc.) should be introduced in the context of specific subtheories, giving rise to associated categories of quantum models. Such conditions and the geometric structures associated with them will be studied in Sections 3 and 5.

[1]The only definite exception that we have made consciously is the case of objects named by their inventors with someone else's name(s), if this naming convention became influential. The examples include the KMS condition (invented and named by [295]), the BLP space (invented and named by [549]), and the abstract notion of a 'Hilbert space', invented (and named) by von Neumann [749]. Hilbert invented only its special case, an $\boldsymbol{\ell}_2$ space, while an $L_2(\mathbb{R}, \mathrm{d}\lambda)$ space was invented by Riesz. More generally, we have adopted the sequential adaptation of rules: (1) strong relevance of temporal priority, (2) weak relevance of folk popularity, (3) weak avoidance of acronyms. As a borderline example, we speak of the "Morse–Transue–Nakano–Luxemburg norm", abbreviated to the "MTNL norm", despite 5 years of difference between [508, 527] and [472] due to wide popularity of the term "Luxemburg norm" (enforced by [408, 604]), but we speak of the "Csiszár–Morimoto $\mathfrak{f}$-distance" despite 3 years of difference between [176, 500] and the Ali–Silvey paper [11]. On the contrary to this partially restrictive *naming* system, the *references* are made as complete as it is reasonably possible.

[2]This is required for agreement with the widespread practice to transliterate Холево as Holevo, etc.

2.1 Integration on boolean algebras

A ***partially ordered set*** (or a ***poset***) [319] is defined as a pair $(X, \leq)$, where X is a set, and $\leq$ is a relation on X such that

$$x \leq x, \quad (x \leq y,\ y \leq x) \Rightarrow x = y, \quad (x \leq y,\ y \leq z) \Rightarrow x \leq z \ \ \forall x, y, z \in X. \tag{1}$$

If $(X, \leq)$ is a poset and $Y \subseteq X$, then Y is called: ***bounded above*** iff $\exists x \in X\ \forall y \in Y\ y \leq x$; ***bounded below*** iff $\exists x \in X\ \forall y \in Y\ x \leq y$; ***upwards directed*** iff Y is nonempty and every pair of elements of Y is bounded above; ***downwards directed*** iff Y is nonempty and every pair of elements is bounded below. A ***supremum*** (or the ***least upper bound***) of $Y \subseteq X$, denoted by $\sup Y$, is defined as $x \in X$ such that

$$y \leq x, \ \ y \leq z \Rightarrow x \leq z \ \ \forall z \in X\ \forall y \in Y, \tag{2}$$

while an ***infimum*** (or the ***greatest lower bound***) of $Y \subseteq X$, denoted by $\inf Y$, is defined as $x \in X$ such that

$$x \leq y, \ \ z \leq y \Rightarrow z \leq x \ \ \forall z \in X\ \forall y \in Y. \tag{3}$$

If I is a set and $\{x_\iota \mid \iota \in I\} \subseteq X$, then $\sup\{x_\iota \mid \iota \in I\} =: \sup_{\iota \in I}\{x_\iota\} =: \sup_\iota\{x_\iota\}$ (and analogously for inf). If $I = \mathbb{N}$ and $i, n \in \mathbb{N}$, then $\sup_i\{x_i\} =: \bigvee_i x_i$, $\inf_i\{x_i\} =: \bigwedge_i x_i$, $\sup\{x_1, \ldots, x_n\} =: x_1 \vee \ldots \vee x_n$ and $\inf\{x_1, \ldots, x_n\} =: x_1 \wedge \ldots \wedge x_n$. If $(X, \leq)$ is a poset, then $Y \subseteq X$ is called ***order closed*** iff $\sup Z_1 \in Y$ for every nonempty upwards directed $Z_1 \subseteq Y$ such that $\sup Z_1 \in X$ and $\inf Z_2 \in Y$ for every nonempty downwards directed $Z_2 \subseteq Y$ such that $\inf Z_2 \in X$. A poset $(X, \leq)$ is called: ***Dedekind–MacNeille complete*** [209, 477] iff every nonempty bounded above subset of X has a supremum, *or*, equivalently, iff every bounded below subset of X has an infimum; ***countably additive complete*** iff every nonempty bounded above countable subset of X has a supremum *and* every nonempty bounded below countable subset of X has an infimum; ***lattice*** [559, 560, 656, 553] iff every subset of X consisting of two elements has a supremum and infimum. A lattice X is called: ***distributive*** [656] iff

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z) \ \ \forall x, y, z \in X, \tag{4}$$
$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z) \ \ \forall x, y, z \in X; \tag{5}$$

boolean [93, 94, 762] iff it is distributive, contains a ***least element*** $0 \in X$ such that $0 \leq x\ \forall x \in X$ and a ***greatest element*** $1 \in X$ such that $x \leq 1\ \forall x \in X$, and

$$\forall x \in X\ \exists y \in X \ \text{ such that } \ x \wedge y = 0, \ \ x \vee y = 1, \ \text{ and } \ y =: \neg x; \tag{6}$$

Riesz [197, 623, 624, 625] iff it is a vector space over $\mathbb{R}$ such that

$$x \leq y \Rightarrow x + z \leq y + z, \ \ x \geq 0 \Rightarrow \lambda x \geq 0 \ \ \forall \lambda \geq 0\ \forall x, y, z \in X; \tag{7}$$

Banach [373, 374, 78] iff it is a Riesz lattice equipped with a norm $\|\cdot\| : X \to \mathbb{R}^+$ such that $|x| \leq |y| \Rightarrow \|x\| \leq \|y\|$ and it is Cauchy complete with respect to this norm, where $|x| := x \vee (-x)$; an ***f-algebra*** [79] iff it is a Riesz lattice equipped with an associative multiplication $\cdot : X \times X \to X$ such that $(X, +, \cdot)$ is an algebra over $\mathbb{R}$, $x, y \geq 0 \Rightarrow x \cdot y \geq 0$, and $(x \wedge y = 0,\ z \geq 0) \Rightarrow (x \cdot z) \wedge y = 0$. Every Dedekind–MacNeille complete lattice is countably additive complete. Every Riesz lattice is distributive [263]. If X is a Riesz lattice and $x \in X$ then $x^+ := x \vee 0$ and $x^- := (-x) \vee 0$ satisfy $x = x^+ - x^-$ and $|x| = x^+ + x^-$. A Riesz lattice X is called ***archimedean*** iff

$$\{nx \mid n \in \mathbb{N}\} \text{ is bounded above } \Rightarrow x \leq 0 \ \ \forall x \in X. \tag{8}$$

Every countably additive complete Riesz lattice is archimedean. Every Banach lattice is archimedean. Every archimedean f-algebra is commutative [24, 79]. An element $e \in X^+ := \{x \in X \mid x \geq 0\}$ of an archimedean Riesz lattice X is called an ***order unit*** iff $\forall x \in X\ \exists \lambda > 0\ |x| \leq \lambda e$ [263]. If X is an archimedean Riesz lattice with an order unit e, then an ***order unit norm*** on X is defined as a map

$\|\cdot\|_e : X \to \mathbb{R}^+$ such that $\|x\|_e := \min\{\lambda \in \mathbb{R} \mid |x| \leq \lambda e\}$. An ***MI-space*** [410, 371] is defined as a Banach lattice with an order unit norm. An ***abstract*** L_p ***space*** [77, 88, 371] is defined for $p \in [1, \infty[$ as a Banach lattice X with norm such that

$$|x| \wedge |y| = 0 \;\Rightarrow\; \|x+y\|^p = \|x\|^p + \|y\|^p \;\; \forall x, y \in X, \tag{9}$$

and as a countably additive complete MI-space X for $p = \infty$. An abstract L_∞ space will be called ***proper*** iff it is Banach dual to some Banach space. Every abstract L_p space for $p \in [1, \infty[$ is Dedekind–MacNeille complete. A commutative ring $(\mathcal{A}, +, \cdot)$ is called ***boolean*** iff $x^2 = x \; \forall x \in \mathcal{A}$. Every boolean lattice defines a boolean ring with unit by $x + y := (x \wedge \neg y) \vee (\neg x \vee y)$ and $x \cdot y := x \wedge y$, and the converse is also true [693]. By this reason both are referred to as a ***boolean algebra***. A simplest nontrivial example of a boolean algebra is $\mathbf{2}$, consisting of two elements $\{0, 1\}$ such that $0 \leq 1$ and $0 \neq 1$.

An order closed vector subspace Y of a Riesz lattice X is called a ***band*** iff $(x \in Y, \; |y| \leq |x|) \Rightarrow y \in Y$. If X is an archimedean Riesz lattice and $Z \subseteq X$, then

$$Z^\perp := \{x \in X \mid |x| \wedge |y| = 0 \; \forall y \in Z\} \tag{10}$$

is a band and $Z^{\perp\perp} = Z$. A subset Y of an archimedean Riesz lattice X is called a ***projection band*** iff $Y + Y^\perp = X$. If X is archimedean and Dedekind–MacNeille complete, then each band of X is a projection band. The set of all bands of an archimedean Riesz lattice X forms a Dedekind–MacNeille complete boolean algebra $\mathcal{A}$, with $Y \wedge Z := Y \cap Z$, $Y \vee Z := (Y + Z)^{\perp\perp}$, $1 := X$, $0 := \{0\}$, $\neg Y := Y^\perp$, $(Y \leq Z) := (Y \subseteq Z)$, while the set of all projection bands of X forms a boolean subalgebra of $\mathcal{A}$. These two boolean algebras coincide iff X is Dedekind–MacNeille complete.

If $(X_1, \leq_1)$ and $(X_2, \leq_2)$ are partially ordered sets, then a function $f : X_1 \to X_2$ is called: ***order preserving*** iff $x \leq_1 y \Rightarrow f(x) \leq_2 f(y) \; \forall x, y \in X_1$; ***order continuous*** [527] iff it is order preserving, $f(\sup Y) = \sup_{x \in Y}\{f(x)\}$ for every nonempty upwards directed $Y \subseteq X_1$ with $\sup Y \in X_1$, and $f(\inf Y) = \inf_{x \in Y}\{f(x)\}$ for every nonempty downwards directed $Y \subseteq X_1$ with $\inf Y \in X_1$; ***sequentially order continuous*** iff it is order preserving, $f(\sup_i\{x_i\}) = \sup_i\{f(x_i)\}$ for every nondecreasing sequence $\{x_i\} \subseteq X_1$, and $f(\inf_i\{x_i\}) = \inf_i\{f(x)\}$ for every nonincreasing sequence $\{x_i\} \subseteq X_1$. If X_1 and X_2 are lattices, then a ***lattice homomorphism*** is defined as a function $f : X_1 \to X_2$ such that $f(x \vee y) = f(x) \vee f(y)$ and $f(x \wedge y) = f(x) \wedge f(y)$. If $\mathcal{A}_1$ and $\mathcal{A}_2$ are boolean algebras, then a ***boolean homomorphism*** is defined as a ring homomorphism $f : \mathcal{A}_1 \to \mathcal{A}_2$ such that $f(1) = 1$. If X_1 and X_2 are Riesz lattices, then a ***Riesz homomorphism*** is defined as a linear function $f : X_1 \to X_2$ such that any of equivalent conditions holds: $f(x^+) = (f(x))^+$; $f(|x|) = |f(x)|$; $f(x \wedge y) = f(x) \wedge f(y)$; $f(x \vee y) = f(x) \vee f(y)$. If X_1 and X_2 are Banach lattices then a Riesz homomorphism $f : X_1 \to X_2$ is called: ***unit preserving*** iff X_1 has an order unit norm with an order unit e_1, X_2 has an order unit norm with an order unit e_2 and $f(e_1) = e_2$; ***norm preserving*** iff $\|f(x)\|_{X_2} = \|x\|_{X_1}$; ***isometric*** iff it is norm preserving and continuous with respect to norm topologies on X_1 and X_2. A ***boolean isomorphism*** is defined as a bijective boolean homomorphism, while a ***Riesz isomorphism*** is defined as a bijective Riesz homomorphism. Isometric Riesz isomorphisms of Banach lattices coincide with their isometric isomorphisms (surjective isometries). Every isometric Riesz isomorphism is order continuous. Every boolean homomorphism and every Riesz lattice homomorphism is a lattice homomorphism. Every bijective lattice homomorphism is order continuous. Every boolean homomorphism is order preserving. A multiplication in archimedean f-algebra is order continuous.

A ***measure*** on a boolean algebra $\mathcal{A}$ is defined as a function $\mu : \mathcal{A} \to [0, \infty]$ such that $\mu(0) = 0$. It is called: ***countably additive*** iff

$$\mu(\bigvee_i x_i) = \sum_i \mu(x_i) \;\; \text{for} \;\; (i \neq j \Rightarrow x_i \wedge x_j = 0); \tag{11}$$

strictly positive iff $x \neq 0 \Rightarrow \mu(x) > 0$; ***finite*** iff $\mathrm{cod}(\mu) \subseteq \mathbb{R}^+$; ***semi-finite*** iff

$$\forall x \in \mathcal{A} \; \exists y \in \mathcal{A} \;\; \mu(x) = \infty \Rightarrow (y \leq x \text{ and } 0 < \mu(y) < \infty). \tag{12}$$

The space of all semi-finite countably additive measures on a boolean algebra $\mathcal{A}$ will be denoted $\mathcal{W}(\mathcal{A})$, while the subset of strictly positive elements of $\mathcal{W}(\mathcal{A})$ will be denoted $\mathcal{W}_0(\mathcal{A})$. A boolean algebra will be called: ***ccb-algebra*** iff it is countably additive complete; ***mcb-algebra*** iff it allows a semi-finite strictly positive countably additive measure and is Dedekind–MacNeille complete. A pair $(\mathcal{A}, \mu)$ of a ccb-algebra $\mathcal{A}$ and a strictly positive countably additive measure μ on $\mathcal{A}$ is called a ***measure algebra***. A measure algebra $(\mathcal{A}, \mu)$ is called: ***semi-finite*** iff μ is semi-finite; ***localisable*** iff $\mathcal{A}$ is an mcb-algebra and μ is semi-finite. If $(\mathcal{A}, \mu)$ is a measure algebra, then $\mathcal{A}^\mu := \{x \in \mathcal{A} \mid \mu(x) < \infty\}$ is a boolean algebra and an ideal in $\mathcal{A}$. An ***evaluation*** on a boolean algebra $\mathcal{A}$ is defined as a function $\phi : \mathcal{A} \to \mathbb{R}$ satisfying $\phi(0) = 0$ and countably additive in the sense of (11) with μ substituted by ϕ. It is called: ***positive*** iff $\mathrm{cod}(\phi) \subseteq \mathbb{R}^+$; ***strictly positive*** iff $x \neq 0 \Rightarrow \phi(x) > 0$. The set of all evaluations on $\mathcal{A}$ will be denoted $\mathrm{eval}(\mathcal{A})$, and its subsets of all positive (resp., strictly positive) elements will be denoted by $\mathrm{eval}(\mathcal{A})^+$ (resp. $\mathrm{eval}(\mathcal{A})^+_0$). Every positive evaluation is an element of $\mathcal{W}(\mathcal{A})$, hence the diagram

$$\begin{tikzcd} \mathrm{eval}(\mathcal{A})^+_0 \arrow[r, hook] \arrow[d, hook] & \mathcal{W}_0(\mathcal{A}) \arrow[d, hook] \\ \mathrm{eval}(\mathcal{A})^+ \arrow[r, hook] & \mathcal{W}(\mathcal{A}) \end{tikzcd} \tag{13}$$

is commutative.

If $\mathcal{A}$ is a ccb-algebra, then $L_0(\mathcal{A})$ is defined as a set of all functions $f : \mathbb{R} \to \mathcal{A}$ such that

$$f(\lambda_1) = \sup_{\lambda_2 > \lambda_1} f(\lambda_2)\ \forall \lambda_1 \in \mathbb{R}, \quad \inf_{\lambda \in \mathbb{R}} f(\lambda) = 0, \quad \sup_{\lambda \in \mathbb{R}} f(\lambda) = 1. \tag{14}$$

The $L_0(\mathcal{A})$ space can be equipped with an f-algebra structure, provided by

$$(x \cdot y)(\lambda_1) := \sup \left\{ x(\lambda_2) \wedge y \left(\frac{\lambda_1}{\lambda_2} \right) \mid \lambda_2 \in \mathbb{Q},\ \lambda_2 > 0 \right\} \quad \forall x, y \geq 0, \tag{15}$$

and

$$x \cdot y := x^+ \cdot x^+ - x^+ \cdot y^- - x^- \cdot y^+ + x^- \cdot y^- \quad \forall x, y \in L_0(\mathcal{A}). \tag{16}$$

For any measure algebra $(\mathcal{A}, \mu)$, the map

$$\|\cdot\|_1 : L_0(\mathcal{A}) \ni f \mapsto \int_0^\infty \mathrm{d}\lambda\, \mu(|f(\lambda)|) \in [0, \infty], \tag{17}$$

where $\mathrm{d}\lambda$ is a Lebesgue measure on $\mathbb{R}$, allows to define

$$L_1(\mathcal{A}, \mu) := \{ f \in L_0(\mathcal{A}) \mid \|f\|_1 < \infty \}. \tag{18}$$

Moreover, for $p \in]1, \infty[$,

$$|f(\lambda)|^p := \begin{cases} \left| f(\lambda^{1/p}) \right| & : \lambda \geq 0 \\ 1 & : \lambda < 0 \end{cases} \tag{19}$$

allows to define

$$L_p(\mathcal{A}, \mu) := \{ f \in L_0(\mathcal{A}) \mid |f|^p \in L_1(\mathcal{A}, \mu) \} \tag{20}$$

and

$$\|\cdot\|_p : L_p(\mathcal{A}, \mu) \ni f \mapsto \||f|^p\|_1^{1/p} \in \mathbb{R}^+. \tag{21}$$

For $p \in [1, \infty[$ the maps $\|\cdot\|_p$ are norms on $L_p(\mathcal{A}, \mu)$ under which $L_p(\mathcal{A}, \mu)$ are Cauchy complete.

If $(\mathcal{A}, \mu)$ is a measure algebra and $x \in L_1(\mathcal{A}, \mu)$, then the function $\int \mu : L_1(\mathcal{A}, \mu) \to \mathbb{R}$, defined by

$$\int \mu x := \|x^+\|_1 - \|x^-\|_1 = \int_0^\infty \mathrm{d}\lambda\, \mu(x(\lambda)) - \int_0^\infty \mathrm{d}\lambda\, \mu(-x(\lambda)), \tag{22}$$

is linear and order continuous, and satisfies

$$\|x\|_1 = \int \mu|x| \;\; \forall x \in L_1(\mathcal{A},\mu), \tag{23}$$

$$\|x\|_p = \left(\int \mu|x|^p\right)^{1/p} = \||x|^p\|_1^{1/p} \;\; \forall x \in L_p(\mathcal{A},\mu) \; \forall p \in [1,\infty[. \tag{24}$$

Let $\mathcal{A}$ be an arbitrary boolean algebra, let X be a vector space of all sums $\sum_{i=1}^n \lambda_i x_i$ with $\{\lambda_i\} \subseteq \mathbb{R}$ and $\{x_i\} \subseteq \mathcal{A}$, and let Y be a vector subspace of X spanned by the elements of X of the form $(x_1 \vee x_2) - x_1 - x_2$ for $x_1, x_2 \in \mathcal{A}$ such that $x_1 \wedge x_2 = 0$. The space X/Y can be equipped with the norm

$$\|f\|_\infty := \min\{\lambda \geq 0 \mid |f| \leq \lambda\chi(1)\} \;\; \forall f \in X/Y, \tag{25}$$

where $\chi : \mathcal{A} \to X/Y$ is defined as a map from $x \in \mathcal{A}$ to an image of $x \in X$ in X/Y. The space $L_\infty(\mathcal{A})$ is defined as a Cauchy completion of X/Y in $\|\cdot\|_\infty$.[3] The order unit of $L_\infty(\mathcal{A})$ is given by the constant function taking the value 1 everywhere. The projection band algebra of $L_\infty(\mathcal{A})$ is boolean isomorphic to $\mathcal{A}$. $L_\infty(\mathcal{A})$ is Dedekind–MacNeille complete iff $\mathcal{A}$ is, and is countably additive complete iff $\mathcal{A}$ is. If $f : \mathcal{A}_1 \to \mathcal{A}_2$ is a boolean homomorphism, then the formula

$$L_\infty(f)(\chi(x)) = \chi(f(x)) \;\; \forall x \in \mathcal{A}_1 \tag{26}$$

determines a unique Riesz homomorphism $L_\infty(f) : L_\infty(\mathcal{A}_1) \to L_\infty(\mathcal{A}_2)$ which is unit preserving, and is surjective (resp.: injective; order continuous) iff f is surjective (resp.: injective; order continuous).

The spaces $L_p(\mathcal{A},\mu)$ inherit an f-algebra structure from $L_0(\mathcal{A})$ and are Dedekind–MacNeille complete. If $\mathcal{A}$ is a ccb-algebra and $\mu_1, \mu_2 \in \mathcal{W}(\mathcal{A})$, then $L_p(\mathcal{A},\mu_1)$ and $L_p(\mathcal{A},\mu_2)$ are isometrically Riesz isomorphic. If $(\mathcal{A},\mu)$ is a localisable measure algebra, then the band algebra of $L_p(\mathcal{A},\mu)$ is boolean isomorphic to $\mathcal{A}$. The space $\mathrm{eval}(\mathcal{A})$ is an abstract L_1 space, and if $(\mathcal{A},\mu)$ is a semi-finite measure algebra, then there exists a bijective Riesz isomorphism between $\mathrm{eval}(\mathcal{A})$ and $L_1(\mathcal{A},\mu)$. Hence, there exists a bijection between $L_1(\mathcal{A},\mu)^+$ and $\mathrm{eval}(\mathcal{A})^+$. For any measure algebra $(\mathcal{A},\mu)$ and $\gamma \in]0,1[$ there is a Banach space duality $L_{1/\gamma}(\mathcal{A},\mu)^\star \cong L_{1/(1-\gamma)}(\mathcal{A},\mu)$ determined by the map

$$L_{1/\gamma}(\mathcal{A},\mu) \times L_{1/(1-\gamma)}(\mathcal{A},\mu) \ni (x,y) \mapsto \int \mu xy \in \mathbb{R}. \tag{27}$$

The space $L_\infty(\mathcal{A})$ can be identified with the linear subspace of $L_0(\mathcal{A})$ generated by $\chi(1)$, and in such case $L_1(\mathcal{A},\mu) \times L_\infty(\mathcal{A}) \ni (x,y) \mapsto x \cdot y \in L_1(\mathcal{A},\mu)$ is a bilinear maps, while

$$L_1(\mathcal{A},\mu) \times L_\infty(\mathcal{A}) \ni (x,y) \mapsto \int \mu xy \in \mathbb{R} \tag{28}$$

is a bilinear functional. According to Segal's theorem [658], the space $L_1(\mathcal{A},\mu)^\star$ is isometrically Riesz isomorphic to $L_\infty(\mathcal{A})$ iff $(\mathcal{A},\mu)$ is localisable, and in such case all Banach preduals of $L_\infty(\mathcal{A})$ are isometrically (and Riesz) isomorphic. According to the Bohnenblust–Kakutani–Nakano theorem [88, 371, 370, 89, 27]:

(i) every abstract L_p space X for $p \in [1,\infty[$ is isometrically Riesz isomorphic to some $L_p(\mathcal{A},\mu)$ space, where $\mathcal{A}$ is uniquely determined as an mcb-algebra of projection bands of X, while $\mu \in \mathcal{W}(\mathcal{A})$ is (nonuniquely) determined by $\mathcal{A}$ and a norm of X, so that $(\mathcal{A},\mu)$ is a localisable measure algebra;

(ii) every abstract L_∞ space X determines a ccb-algebra $\mathcal{A}$ of its projection bands, and X is isometrically Riesz isomorphic to $L_\infty(\mathcal{A})$. Hence, every Dedekind–MacNeille complete abstract L_∞ space X is isometrically Riesz isomorphic to $L_\infty(\mathcal{A})$, where $\mathcal{A}$ is a Dedekind–MacNeille complete boolean algebra.

[3] Equivalently, one can define (real or complex) Banach lattice $L_\infty(\mathcal{A})$ as the space of all (real or complex) continuous functions on the Stone spectrum $\mathrm{sp}_{\mathrm{S}}(\mathcal{A})$, endowed with its multiplication, linear and order structures, and norm given by $\|f\| := \sup_{\chi \in \mathrm{sp}_{\mathrm{S}}(\mathcal{A})}\{|f(\chi)|\}$.

By Segal's theorem [658], this implies that

(iii) every proper abstract L_∞ space X is isometrically Riesz isomorphic to $L_\infty(\mathcal{A})$, where $\mathcal{A}$ is an mcb-algebra.

For a given set $\mathcal{X}$, a ***countably additive algebra*** on $\mathcal{X}$ is defined as a family $\mho(\mathcal{X})$ of subsets of $\mathcal{X}$ such that

$$\varnothing \in \mho(\mathcal{X}),\ \ \mathcal{Y} \in \mho(\mathcal{X}) \Rightarrow \mathcal{X} \setminus \mathcal{Y} \in \mho(\mathcal{X}),\ \ \bigcup_i \mathcal{X}_i \in \mho(\mathcal{X}) \ \text{ for any sequence } \{\mathcal{X}_i\} \subseteq \mho(\mathcal{X}). \tag{29}$$

A ***countably additive ideal*** [734] of a countably additive algebra $\mho(\mathcal{X})$ on $\mathcal{X}$ is defined as a family $\mho^0(\mathcal{X})$ of subsets of $\mho(\mathcal{X})$ such that

1) $\varnothing \in \mho^0(\mathcal{X})$,
2) $(\mathcal{X}_1 \in \mho^0(\mathcal{X}),\ \mathcal{X}_2 \in \mho(\mathcal{X}),\ \mathcal{X}_2 \subseteq \mathcal{X}_1) \Rightarrow \mathcal{X}_2 \in \mho^0(\mathcal{X})$,
3) $\bigcup_i \mathcal{X}_i \in \mho^0(\mathcal{X})$ for any countable set $\{\mathcal{X}_i\} \subseteq \mho^0(\mathcal{X})$.

A ***premeasurable space*** is defined as a pair $(\mathcal{X}, \mho(\mathcal{X}))$, while a ***measurable space*** is defined as a triple $(\mathcal{X}, \mho(\mathcal{X}), \mho^0(\mathcal{X}))$, where $\mho(\mathcal{X})$ is any countable additive algebra on $\mathcal{X}$, while $\mho^0(\mathcal{X})$ any countably additive ideal of $\mho(\mathcal{X})$. A ***complete morphism*** of premeasurable spaces, $(\mathcal{X}_1, \mho_1(\mathcal{X}_1), \mho^0_1(\mathcal{X}_1)) \to (\mathcal{X}_2, \mho_2(\mathcal{X}_2), \mho^0_2(\mathcal{X}_2))$, is defined as a map $f : \mathcal{X}_1 \to \mathcal{X}_2$ such that $f^{-1}(\mathcal{Y}) \in \mho_1(\mathcal{X}_1)\ \forall \mathcal{Y} \in \mho_2(\mathcal{X}_2)$ and $f^{-1}(\mathcal{Z}) \in \mho^0_1(\mathcal{X}_1)\ \forall \mathcal{Z} \in \mho^0(\mathcal{X}_2)$. A ***measure*** on a premeasurable space $(\mathcal{X}, \mho(\mathcal{X}))$ is defined as a function $\tilde{\mu} : \mho(\mathcal{X}) \to [0, \infty]$ such that $\tilde{\mu}(\varnothing) = 0$. A measure is called ***countably additive*** iff $\tilde{\mu}(\bigcup_i \mathcal{X}_i) = \sum_i \tilde{\mu}(\mathcal{X}_i)$ for any countable sequence $\{\mathcal{X}_i\} \subseteq \mho(\mathcal{X})$ satisfying $i \neq j \Rightarrow \mathcal{X}_i \cap \mathcal{X}_j = \varnothing$. A set of all countably additive measures on $(\mathcal{X}, \mho(\mathcal{X}))$ will be denoted $\mathrm{Meas}^+(\mathcal{X}, \mho(\mathcal{X}))$. As set of finite elements of $\mathrm{Meas}^+(\mathcal{X}, \mho(\mathcal{X}))$ will be denoted $\mathrm{Meas}^+_{\mathrm{fin}}(\mathcal{X}, \mho(\mathcal{X}))$. Moreover, $\mathrm{Meas}(\mathcal{X}, \mho(\mathcal{X})) := \{\tilde{\mu} := \tilde{\mu}_1 - \tilde{\mu}_2 \mid \tilde{\mu}_1, \tilde{\mu}_2 \in \mathrm{Meas}^+(\mathcal{X}, \mho(\mathcal{X}))\}$. A ***measure space*** is defined as a triple $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$, where $(\mathcal{X}, \mho(\mathcal{X}))$ is a premeasurable space, and $\tilde{\mu}$ is countably additive measure on it. Given a measure space $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$, a set $\mathcal{Y} \subseteq \mathcal{X}$ is called ***$\tilde{\mu}$-null*** iff there exists $\mathcal{Z} \subseteq \mho(\mathcal{X})$ such that $\mathcal{Y} \subseteq \mathcal{Z}$ and $\tilde{\mu}(\mathcal{Z}) = 0$. A family of all $\tilde{\mu}$-null subsets of $\mathcal{X}$ is denoted by $\mathrm{null}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$. One says that the property $Q(\chi)$ holds for ***$\tilde{\mu}$-almost every*** $\chi \in \mathcal{X}$ iff $\{\chi \in \mathcal{X} \mid Q(\chi) \text{ is false}\} \in \mathrm{null}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$. The set

$$\mho^{\tilde{\mu}}(\mathcal{X}) := \mho(\mathcal{X}) \cap \mathrm{null}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) = \{\mathcal{Y} \in \mho(\mathcal{X}) \mid \tilde{\mu}(\mathcal{Y}) = 0\} \tag{30}$$

is a countably additive ideal of $\mho(\mathcal{X})$, hence, every measure space $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ determines a corresponding measurable space $(\mathcal{X}, \mho(\mathcal{X}), \mho^{\tilde{\mu}}(\mathcal{X}))$. A measure space $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ is called: ***semi-finite*** iff

$$\forall \mathcal{X}_1 \in \mho(\mathcal{X})\ \exists \mathcal{X}_2 \in \mho(\mathcal{X})\ \ \tilde{\mu}(\mathcal{X}_1) = \infty \Rightarrow (\mathcal{X}_2 \subseteq \mathcal{X}_1 \text{ and } 0 < \tilde{\mu}(\mathcal{X}_2) < \infty); \tag{31}$$

localisable iff it is semi-finite and for all $Y \subseteq \mho(\mathcal{X})$ there exists $\mathcal{X}_1 \in \mho(\mathcal{X})$ such that

1) $\mathcal{Y} \setminus \mathcal{X}_1 \in \mathrm{null}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})\ \forall \mathcal{Y} \in Y$,
2) $(\mathcal{X}_2 \in \mho(\mathcal{X}),\ \mathcal{Y} \setminus \mathcal{X}_2 \in \mathrm{null}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})\ \forall \mathcal{Y} \in Y) \Rightarrow \mathcal{X}_1 \setminus \mathcal{X}_2 \in \mathrm{null}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$.

A measurable space $(\mathcal{X}, \mho(\mathcal{X}), \mho^0(\mathcal{X}))$ will be called ***localisable*** iff there exists a measure $\tilde{\mu}$ on $(\mathcal{X}, \mho(\mathcal{X}))$ such that $\mho^0(\mathcal{X}) = \mho^{\tilde{\mu}}(\mathcal{X})$ and $(\mathcal{X}, \mho(\mathcal{X}), \mho^{\tilde{\mu}}(\mathcal{X}))$ is localisable. If $(\mathcal{X}, \mho(\mathcal{X}))$ is a premeasurable space and $\mathcal{Y} \subseteq \mathcal{X}$, then $\mho^{\mathcal{X}}(\mathcal{Y}) := \{\mathcal{Z} \cap \mathcal{Y} \mid \mathcal{Z} \in \mho(\mathcal{X})\}$ is a countably additive algebra on $\mathcal{Y}$. A function $f : \mathcal{Y} \to \mathbb{R}$ is called ***$\mho(\mathcal{X})$-measurable*** iff $\{\chi \in \mathcal{X} \mid f(\chi) \leq \lambda\} \subseteq \mho^{\mathcal{X}}(\mathcal{Y})\ \forall \lambda \in \mathbb{R}$.

By Wecken's theorem [757], every measurable space $(\mathcal{X}, \mho(\mathcal{X}), \mho^0(\mathcal{X}))$ determines a ccb-algebra $\mathcal{A}$ by $\mathcal{A} := \mho(\mathcal{X})/\mho^0(\mathcal{X})$, and, in particular, every measure space $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ determines a ccb-algebra

$$\mathcal{A}_{\tilde{\mu}} := \mho(\mathcal{X})/\mho^{\tilde{\mu}}(\mathcal{X}) = \mho(\mathcal{X})/\{\mathcal{Y} \in \mho(\mathcal{X}) \mid \tilde{\mu}(\mathcal{Y}) = 0\}, \tag{32}$$

and a measure algebra $(\mathcal{A}_{\tilde{\mu}}, \mu)$, with

$$\mu([\mathcal{Z}]_{\mathcal{A}_{\tilde{\mu}}}) := \tilde{\mu}(\mathcal{Z})\ \ \forall \mathcal{Z} \in \mho(\mathcal{X}), \tag{33}$$

where the map $\mho(\mathcal{X}) \ni \mathcal{Z} \mapsto [\mathcal{Z}]_{\mathcal{A}_{\tilde{\mu}}} \in \mathcal{A}_{\tilde{\mu}}$ is defined by (32), and is sequentially order continuous. On the other hand, for every ccb-algebra $\mathcal{A}$ the Loomis–Sikorski theorem [468, 683] provides an explicit construction of a measurable space $(\mathrm{sp}_{\mathrm{S}}(\mathcal{A}), \mho_{\mathrm{LS}}(\mathrm{sp}_{\mathrm{S}}(\mathcal{A})), \mho^0_{\mathrm{LS}}(\mathrm{sp}_{\mathrm{S}}(\mathcal{A})))$, such that $\mho_{\mathrm{LS}}(\mathrm{sp}_{\mathrm{S}}(\mathcal{A}))/\mho^0_{\mathrm{LS}}(\mathrm{sp}_{\mathrm{S}}(\mathcal{A}))$ is boolean isomorphic to $\mathcal{A}$.[4] As a consequence, one can show that for every measure algebra $(\mathcal{A}, \mu)$ there exists a measure preserving isomorphism to a measure algebra of some measure space. By Kelley–Namioka theorem [387], the measure space is localisable iff the corresponding ccb-algebra is an mcb-algebra.

A function $f : \mathcal{X} \to \mathbb{R}$ is called ***$\tilde{\mu}$-simple*** iff $f = \sum_{i=1}^n \lambda_i \chi_{\mathcal{Y}_i}$, where $n \in \mathbb{N}$, $\{\lambda_i\} \subseteq \mathbb{R}$, $\{\mathcal{Y}_i\} \subseteq \mathcal{X}$ are $\mho(\mathcal{X})$-measurable sets with $\tilde{\mu}(\mathcal{Y}_i) < \infty$, and $\chi_{\mathcal{Y}_i}$ are characteristic functions of $\mathcal{Y}_i$. A ***$\tilde{\mu}$-integral*** of a $\tilde{\mu}$-simple f is defined as $\int \tilde{\mu} f := \sum_{i=1}^n \lambda_i \mu(\mathcal{Y}_i)$. A function $f : \mathcal{X} \to \mathbb{R}$ is called ***$\tilde{\mu}$-integrable*** iff $f = f_a - f_b$, where $f_o \in \{f_a, f_b\}$ satisfy

1) $\mathcal{X} \setminus \mathrm{dom} f_o$ is $\tilde{\mu}$-null,
2) $f_o(\chi) \in \mathbb{R}^+ \; \forall \chi \in \mathrm{dom} f_o$,
3) there exists a nondecreasing sequence $\{f_i\}$ of simple functions $f_i : \mathcal{X} \to \mathbb{R}^+$ such that $\sup_i\{\int \tilde{\mu} f_i\} < \infty$ and $\lim_{i\to\infty} f_i(\chi) = f_o(\chi)$ holds $\tilde{\mu}$-almost everywhere.

A ***$\tilde{\mu}$-integral*** of $\tilde{\mu}$-integrable f is defined as $\int \tilde{\mu} f := \int \tilde{\mu} f_a - \int \tilde{\mu} f_b$. If $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ is a measure space, then the set of functions $f : \mathcal{X} \to \mathbb{R}$ such that

i) $\mathcal{X} \setminus \mathrm{dom} f$ is $\tilde{\mu}$-null,
ii) $\exists \mathcal{Y} \subseteq \mathcal{X}$ such that $\mathcal{X} \setminus \mathcal{Y}$ is $\tilde{\mu}$-null and $f|_{\mathcal{Y}}$ is $\mho^{\mathcal{X}}(\mathcal{Y})$-measurable,

is denoted by $\mathcal{L}_0(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$. A space $\mathcal{L}_\infty(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ is defined as a set of $f \in \mathcal{L}_0(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ such that

$$\exists \lambda \geq 0 \;\; \mathcal{X} \setminus \{\chi \in \mathrm{dom} f \mid |f(\chi)| \leq \lambda\} \in \mathrm{null}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}). \tag{34}$$

A space $\mathcal{L}_p(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$, for $p \in]1, \infty[$, is defined as a set of all $f \in \mathcal{L}_0(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ such that $|f|^p$ is $\tilde{\mu}$-integrable. For $p \in [1, \infty] \cup \{0\}$ [621, 599, 237]

$$L_p(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) := \mathcal{L}_p(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) / =_{\tilde{\mu}}, \tag{35}$$

where $=_{\tilde{\mu}}$ is an equivalence relation on elements of $\mathcal{L}_0(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ such that $f_1 =_{\tilde{\mu}} f_2$ iff $f_1 = f_2$ holds $\tilde{\mu}$-almost everywhere.

Every $L_p(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ space for $p \in [1, \infty[$ is a Banach lattice that is isometrically Riesz isomorphic to $L_p(\mathcal{A}_{\tilde{\mu}}, \mu)$ with $(\mathcal{A}_{\tilde{\mu}}, \mu)$ determined by (32) and (33). All mutually isometrically Riesz isomorphic $L_p(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ spaces constructed over various measure spaces $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ can be identified with a single $L_p(\mathcal{A}, \mu)$ space, with $\mathcal{A} \cong \mho(\mathcal{X})/\mho^{\tilde{\mu}}(\mathcal{X})$ and $\mu([\cdot]_{\mathcal{A}}) = \tilde{\mu}$. Finally, for any $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ one has an isometric Riesz isomorphism $L_\infty(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \cong L_\infty(\mho(\mathcal{X})/\mho^{\tilde{\mu}}(\mathcal{X})) \cong L_\infty(\mathcal{A})$. The restriction of validity of isometric isomorphism $L_1(\mathcal{A}, \mu)^\star \cong L_\infty(\mathcal{A})$ to localisable measure algebras $(\mathcal{A}, \mu)$ is equivalent with restriction of validity of the Steinhaus–Nikodým theorem [688, 538] $L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \cong L_\infty(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ to localisable measure spaces, which was established by Segal [658]. The relationships between Riesz lattice theoretic, boolean algebra theoretic, and measure space theoretic approaches to integration theory can be summarised by the following theorem: `[the categories of (1)(2)(3) are equivalent]`. For a detailed discussion see [404] and [262].

The key role played by mcb-algebras in the BKN theorem and the above equivalences suggests that it might be possible to deal with Banach lattice isomorphic $L_p(\mathcal{A}, \mu)$ spaces for $p \in [1, \infty]$ without specifying any particular measure $\mu \in \mathcal{W}_0(\mathcal{A})$ associated with a given mcb-algebra $\mathcal{A}$. In what follows,

---

[4] The ***Stone spectrum*** [693] of a boolean algebra $\mathcal{A}$ is defined as a set $\mathrm{sp}_{\mathrm{S}}(\mathcal{A})$ of nonzero boolean homomorphisms from $\mathcal{A}$ to $\mathbf{2}$, $\mathrm{sp}_{\mathrm{S}}(\mathcal{A}) := \mathrm{Hom}_{\mathbf{B}}(\mathcal{A}, \mathbf{2}) \setminus \{0\}$, equipped with a topology of open sets given by $\{\mathcal{Y} \subseteq \mathrm{sp}_{\mathrm{S}}(\mathcal{A}) \mid \forall \chi \in \mathcal{Y} \; \exists x \in \mathcal{A} \;\; \chi \in \hat{x} \subseteq \mathcal{Y}\}$, where $\hat{\cdot} : \mathcal{A} \to \mathrm{Hom}_{\mathbf{Top}}(\mathrm{sp}_{\mathrm{S}}(\mathcal{A}), \mathbf{2})$ is the ***Stone representation map*** defined by $\hat{x} := \{\chi \in \mathrm{sp}_{\mathrm{S}}(\mathcal{A}) \mid \chi(x) = 1\}$. The set $\{\hat{x} \subseteq \mathrm{sp}_{\mathrm{S}}(\mathcal{A}) \mid x \in \mathcal{A}\}$ consists of all subsets of $\mathrm{sp}_{\mathrm{S}}(\mathcal{A})$ that are open and closed, and is boolean isomorphic to $\mathcal{A}$. The algebra $\mho_{\mathrm{LS}}(\mathrm{sp}_{\mathrm{S}}(\mathcal{A}))$ consists of all open-and-closed subsets of $\mathrm{sp}_{\mathrm{S}}(\mathcal{A})$, while the ideal $\mho^0_{\mathrm{LS}}(\mathrm{sp}_{\mathrm{S}}(\mathcal{A}))$ consists of all subsets of $\mathrm{sp}_{\mathrm{S}}(\mathcal{A})$ that are unions of sequences of such subsets $\mathcal{Y} \subseteq \mathrm{sp}_{\mathrm{S}}(\mathcal{A})$ that $\mathrm{int}(\bar{\mathcal{Y}}) = \varnothing$.

we will construct a family of canonical $L_p(\mathcal{A})$ spaces that are associated functorially to mcb-algebras $\mathcal{A}$. This construction is new for $p \in [1, \infty[$, and it is aimed to provide a commutative counterpart to the Kosaki and Falcone–Takesaki constructions of canonical noncommutative $L_p(\mathcal{N})$ spaces. In principle, one could try to define the space $L_p(\mathcal{A})$ as an equivalence class of $L_p(\mathcal{A}, \mu)$ spaces divided by the isometric Riesz isomorphisms generated by varying μ within $\mathcal{W}_0(\mathcal{A})$. However, this would remove too much structure, making category theoretic description inapplicable (or less applicable). Hence, instead of 'isomorphism invariant' definition, we will provide 'isomorphism covariant' construction, which follows the ideas of Neveu [534] and Zhu [788]. This will enable us to provide an explicit description of the relationship between the canonical integration theory and the commutative case of canonical noncommutative integration theory, *without passing to representations* in terms of measure algebras or measure spaces.

For any countably additive measures μ_1, μ_2 on a ccb-algebra $\mathcal{A}$, μ_1 is called: ***absolutely continuous*** with respect to μ_2 (denoted by $\mu_1 \ll \mu_2$) iff

$$\forall \epsilon_1 > 0 \; \exists \epsilon_2 > 0 \; \forall x \in \mathcal{A} \quad \mu_2(x) \leq \epsilon_2 \;\Rightarrow\; \mu_1(x) \leq \epsilon_1, \tag{36}$$

or, equivalently, iff

$$\mu_2(x) = 0 \quad \Rightarrow \quad \mu_1(x) = 0 \;\; \forall x \in \mathcal{A}; \tag{37}$$

compatible [448] with μ_2 iff

$$\forall x \in \mathcal{A} \;\; 0 < \mu_1(x) < \infty \quad \Rightarrow \quad (\exists y \in \mathcal{A} \;\; y \leq x, \; \mu_1(y) > 0, \; \mu_2(y) < \infty). \tag{38}$$

If $\mu_1, \mu_2 \in \mathcal{W}_0(\mathcal{A})$, then $\mu_1 \ll \mu_2 \ll \mu_1$. If $f \in L_1(\mathcal{A}, \mu)$ and $x \in \mathcal{A}$, then

$$\int_x \mu f := \int_0^\infty \mathrm{d}\lambda \, \mu(x \wedge f(\lambda)). \tag{39}$$

If $\mathcal{A}$ is an mcb-algebra and $\mu_1, \mu_2 \in \mathcal{W}(\mathcal{A})$, then the Segal–Lewin–Lewin theorem [658, 448] (see also [783, 494, 386, 746, 138]) states that for $\mu_2 \ll \mu_1$ and μ_2 compatible with μ_1

$$\exists! f \in L_1(\mathcal{A}, \mu_1) \; \forall x \in \mathcal{A} \;\; \mu_2(x) = \int_x \mu_1 f. \tag{40}$$

Such f will be called a ***Radon–Nikodým quotient*** and denoted by $\frac{\mu_2}{\mu_1}$. This theorem is a generalisation of the Lebesgue–Radon–Daniell–Nikodým theorem [439, 599, 198, 537] (which holds for countably additive finite measures). For $\mu_1, \mu_2 \in \mathcal{W}_0(\mathcal{A})$ the compatibility of μ_1 with μ_2 is equivalent with

$$\forall x \in \mathcal{A}^{\mu_1} \;\; \exists \mathcal{A} \ni y \leq x \;\; y \in \mathcal{A}^{\mu_2}. \tag{41}$$

If $\mu_1, \mu_2, \mu_3 \in \mathcal{W}_0(\mathcal{A})$ are mutually compatible, then their Radon–Nikodým quotients satisfy $\frac{\mu_1}{\mu_2}\frac{\mu_2}{\mu_3} = \frac{\mu_1}{\mu_3}$ and

$$\left(\frac{\mu_i}{\mu_j}\right)^{-1} = \left(\frac{\mu_j}{\mu_i}\right) \quad \forall i, j \in \{1, 2, 3\}. \tag{42}$$

As a consequence, for $\gamma \in]0, 1]$, $\mu_1, \mu_2 \in \mathcal{W}_0(\mathcal{A})$ such that μ_1 and μ_2 are mutually compatible, $f_1 \in L_{1/\gamma}(\mathcal{A}, \mu_1)$, $f_2 \in L_{1/\gamma}(\mathcal{A}, \mu_2)$, the formula

$$(f_1, \mu_1) \sim_{1/\gamma} (f_2, \mu_2) \quad :\Longleftrightarrow \quad f_1 = f_2 \left(\frac{\mu_2}{\mu_1}\right)^{1/\gamma} \tag{43}$$

determines an equivalence relation on $L_0(\mathcal{A}) \times \mathcal{W}_0(\mathcal{A})$, which defines the family of equivalence classes

$$\{ f\mu^\gamma := (f, \mu)/\sim_{1/\gamma} \mid \mu \in \mathcal{W}_0(\mathcal{A}), \; f \in L_{1/\gamma}(\mathcal{A}, \mu) \}. \tag{44}$$

Let $L_{1/\gamma}(\mathcal{A})$ denote the set $\{f\mu^\gamma \mid \mu \in \mathcal{W}_0(\mathcal{A}),\ f \in L_{1/\gamma}(\mathcal{A},\mu)\}$ equipped with the operations

$$f_1\mu_1^\gamma + f_2\mu_2^\gamma := \left(f_1\left(\frac{\mu_1}{\mu_4}\right)^\gamma + f_2\left(\frac{\mu_2}{\mu_4}\right)^\gamma\right)\mu_4^\gamma, \tag{45}$$

$$\lambda(f\mu^\gamma) := (\lambda f)\mu^\gamma, \tag{46}$$

$$\|f\mu^\gamma\|_{1/\gamma} := \left(\int \mu|f|^{1/\gamma}\right)^\gamma, \tag{47}$$

$$f_1\mu_1^\gamma \wedge f_2\mu_2^\gamma := \left(f_1\left(\frac{\mu_1}{\mu_4}\right)^\gamma \wedge f_2\left(\frac{\mu_2}{\mu_4}\right)^\gamma\right)\mu_4^\gamma, \tag{48}$$

$$f_1\mu_1^\gamma \vee f_2\mu_2^\gamma := \left(f_1\left(\frac{\mu_1}{\mu_4}\right)^\gamma \vee f_2\left(\frac{\mu_2}{\mu_4}\right)^\gamma\right)\mu_4^\gamma, \tag{49}$$

where $\mu_4 \in \mathcal{W}_0(\mathcal{A})$ is an arbitrary element of $\mathcal{W}_0(\mathcal{A})$ providing representation of an equivalence class $f\mu^\gamma$ (hence, it is compatible with $\mu_1, \mu_2 \in \mathcal{W}_0(\mathcal{A})$).

Proposition 2.1. *$L_{1/\gamma}(\mathcal{A})$ is an abstract $L_{1/\gamma}$ space for $\gamma \in]0,1]$.*

Proof. We need to check that $L_{1/\gamma}(\mathcal{A})$ satisfies the following properties: 1) it is a lattice; 2) it is a vector space over $\mathbb{R}$; 3) $x \leq y \Rightarrow x+z \leq y+z$; 4) $x \geq 0 \Rightarrow \lambda x \geq 0\ \forall\lambda \geq 0$; 5) $|x| \leq |y| \Rightarrow \|x\| \leq \|y\|$; 6) $\|\cdot\|_{1/\gamma}$ is a norm; 7) it is Cauchy complete in $\|\cdot\|_{1/\gamma}$; 8) $|x| \wedge |y| = 0 \Rightarrow \|x+y\|_{1/\gamma}^{1/\gamma} = \|x\|_{1/\gamma}^{1/\gamma} + \|y\|_{1/\gamma}^{1/\gamma}$. We begin by noting that 2) follows directly from (45), (46) and the vector space structure of $L_{1/\gamma}(\mathcal{A},\mu)$, 6) and 7) follow directly from 2), (47) and the Banach space structure of $L_{1/\gamma}(\mathcal{A},\mu)$, while 1) follows directly from (48), (49) and the lattice structure of $L_{1/\gamma}(\mathcal{A},\mu)$. Hence, it remains to prove 3), 4), 5), and 8).

3) $f_1\mu_1^\gamma \leq f_2\mu_2^\gamma \iff f_1\left(\frac{\mu_1}{\mu_4}\right)^\gamma\mu_4^\gamma \leq f_2\left(\frac{\mu_2}{\mu_4}\right)^\gamma\mu_4^\gamma$,
so $f_1\mu_1^\gamma + f_3\mu_3^\gamma = f_1\left(\frac{\mu_1}{\mu_4}\right)^\gamma\mu_4^\gamma + f_3\left(\frac{\mu_3}{\mu_4}\right)^\gamma\mu_4^\gamma \leq f_2\left(\frac{\mu_2}{\mu_4}\right)^\gamma\mu_4^\gamma + f_3\left(\frac{\mu_3}{\mu_4}\right)^\gamma\mu_4^\gamma = f_2\mu_2^\gamma + f_2\mu_3^\gamma$.

4) $f\mu^\gamma \geq 0 \iff f \geq 0 \Rightarrow \lambda f \geq 0 \iff (\lambda f)\mu^\gamma \geq 0 \iff \lambda(f\mu^\gamma) \geq 0$.

5) Using the f-algebra structure of $L_{1/\gamma}(\mathcal{A},\mu_4)$, we obtain

$$|f\mu^\gamma| = (f\mu^\gamma) \vee (-f\mu^\gamma) = \left(f\left(\frac{\mu}{\mu_4}\right)^\gamma \vee -f\left(\frac{\mu}{\mu_4}\right)^\gamma\right)\mu_4^\gamma = \left|f\left(\frac{\mu}{\mu_4}\right)^\gamma\right|\mu_4^\gamma. \tag{50}$$

This allows us to write

$$|f_1\mu_1^\gamma| \leq |f_2\mu_2^\gamma|, \tag{51}$$

$$\left|f_1\left(\frac{\mu_1}{\mu_4}\right)^\gamma\right|\mu_4^\gamma \leq \left|f_2\left(\frac{\mu_2}{\mu_4}\right)^\gamma\right|\mu_4^\gamma, \tag{52}$$

$$\left(\left|f_2\left(\frac{\mu_2}{\mu_4}\right)^\gamma\right| - \left|f_1\left(\frac{\mu_1}{\mu_4}\right)^\gamma\right|\right)\mu_4^\gamma \geq 0, \tag{53}$$

$$\left|f_2\left(\frac{\mu_2}{\mu_4}\right)^\gamma\right| \geq \left|f_1\left(\frac{\mu_1}{\mu_4}\right)^\gamma\right|, \tag{54}$$

$$\left\|f_2\left(\frac{\mu_2}{\mu_4}\right)^\gamma\right\| \geq \left\|f_1\left(\frac{\mu_1}{\mu_4}\right)^\gamma\right\|, \tag{55}$$

$$\left(\int \mu_4\left|f_2\left(\frac{\mu_2}{\mu_4}\right)^\gamma\right|^{1/\gamma}\right)^\gamma \geq \left(\int \mu_4\left|f_1\left(\frac{\mu_1}{\mu_4}\right)^\gamma\right|^{1/\gamma}\right)^\gamma, \tag{56}$$

$$\left(\int \mu_2|f_2|^{1/\gamma}\right)^\gamma \geq \left(\int \mu_1|f_1|^{1/\gamma}\right)^\gamma, \tag{57}$$

$$\|f_2\mu_2^\gamma\|_{1/\gamma} \geq \|f_1\mu_1^\gamma\|_{1/\gamma}. \tag{58}$$

8) We have

$$\|f_1\mu_1^\gamma + f_2\mu_2^\gamma\|_{1/\gamma}^{1/\gamma} = \|f_1 + f_2\|_{1/\gamma}^{1/\gamma}, \tag{59}$$

$$\|f_1\mu_1^\gamma\|_{1/\gamma}^{1/\gamma} + \|f_2\mu_2^\gamma\|_{1/\gamma}^{1/\gamma} = \|f_1\|_{1/\gamma}^{1/\gamma} + \|f_2\|_{1/\gamma}^{1/\gamma}. \tag{60}$$

In order to prove $|f_1\mu_1^\gamma| \wedge |f_2\mu_2^\gamma| = 0 \iff |f_1| \wedge |f_2| = 0$, we need to use $(x \wedge y = 0,\ z \geq 0) \Rightarrow (x \cdot z) \wedge y = 0$ in $L_{1/\gamma}(\mathcal{A}, \mu_4)$, and the positivity of Radon–Nikodým quotient, which gives us

$$0 = |f_1\mu_1^\gamma| \wedge |f_2\mu_2^\gamma| = \left|f_1\left(\frac{\mu_1}{\mu_4}\right)^\gamma\right|\mu_4^\gamma \wedge \left|f_2\left(\frac{\mu_2}{\mu_4}\right)^\gamma\right|\mu_4^\gamma \quad \iff \tag{61}$$

$$0 = \left|f_1\left(\frac{\mu_1}{\mu_4}\right)^\gamma\left(\left(\frac{\mu_1}{\mu_4}\right)^{-1}\right)^\gamma\right|\mu_4^\gamma \wedge \left|f_2\left(\frac{\mu_2}{\mu_4}\right)^\gamma\left(\left(\frac{\mu_2}{\mu_4}\right)^{-1}\right)^\gamma\right|\mu_4^\gamma = (|f_1| \wedge |f_2|)\,\mu_4^\gamma \quad \iff \tag{62}$$

$$0 = |f_1| \wedge |f_2|. \tag{63}$$

Thus, an abstract $L_{1/\gamma}$ space structure of $L_{1/\gamma}(\mathcal{A})$ follows from an abstract $L_{1/\gamma}$ space structure of $L_{1/\gamma}(\mathcal{A}, \mu_4)$ for $\mu_4 \in \mathcal{W}_0(\mathcal{A})$. □

Hence, every mcb-algebra $\mathcal{A}$ allows to construct a family of canonical commutative $L_p(\mathcal{A})$ spaces over $\mathcal{A}$, with $p \in [1, \infty]$, which are abstract L_p spaces and do not depend on the choice of measure on $\mathcal{A}$. This assignment is functorial, with boolean homomorphisms $f : \mathcal{A}_1 \to \mathcal{A}_2$ mapped to the unit preserving Riesz homomorphisms $\widetilde{f} : L_p(\mathcal{A}_1) \to L_p(\mathcal{A}_2)$, and with boolean isomorphisms mapped to the unit preserving Riesz isomorphisms.

Proposition 2.2. *The map $[\cdot]_\mu : L_{1/\gamma}(\mathcal{A}) \ni x\mu^\gamma \mapsto x \in L_{1/\gamma}(\mathcal{A}, \mu)$ is an isometric Riesz isomorphism.*

Proof. Linearity follows from (45) and (46), isometry follows from (47), while the property $[|x|]_\mu = |[x]_\mu|$ follows from (50). □

Hence, for $\mu \in \mathcal{W}_0(\mathcal{A})$ the function $[\cdot]_\mu$ provides an isometrically Riesz isomorphic representation of $L_{1/\gamma}(\mathcal{A})$ space in terms of the $L_{1/\gamma}(\mathcal{A}, \mu)$ space.

Corollary 2.3. *For any mcb-algebra $\mathcal{A}$ there exists a bijective Riesz homomorphism $L_1(\mathcal{A}) \cong \mathrm{eval}(\mathcal{A})$, and the diagram*

$$\begin{tikzcd} L_1(\mathcal{A})_0^+ \arrow[r, hook] \arrow[d, hook] & \mathcal{W}_0(\mathcal{A}) \arrow[d, hook] \\ L_1(\mathcal{A})^+ \arrow[r, hook] & \mathcal{W}(\mathcal{A}) \end{tikzcd} \tag{64}$$

commutes.

This is a strict analogue of (68) for mcb-algebras. Moreover, if $\mu_1, \mu_2 \in \mathcal{W}_0(\mathcal{A})$ and μ_1 is compatible with μ_2, then $\int \mu_1 f = \int \mu_2 \frac{\mu_1}{\mu_2} f\ \forall f \in L_1(\mathcal{A}, \mu_1)$. This allows us to define a ***canonical integral***,

$$\int : L_1(\mathcal{A}) \ni x \mapsto \int x := \int \mu f \in \mathbb{R}, \tag{65}$$

where $[x]_\mu = f \in L_1(\mathcal{A}, \mu)$, which is independent of the choice of an arbitrary $\mu \in \mathcal{W}_0(\mathcal{A})$. As a result, we obtain a bilinear functional

$$L_{1/\gamma}(\mathcal{A}) \times L_{1/(1-\gamma)}(\mathcal{A}) \ni (x, y) \mapsto \int xy = \int yx \in \mathbb{R}, \tag{66}$$

which sets up a canonical Banach space duality between $L_{1/\gamma}(\mathcal{A})$ and $L_{1/(1-\gamma)}(\mathcal{A})$ spaces for $\gamma \in]0, 1]$, and satisfies $\|x\|_{1/\gamma} = (\int |x|^{1/\gamma})^\gamma$.

2.2 Integration on W^*-algebras

A C^****-algebra*** is a Banach space $\mathcal{C}$ over $\mathbb{C}$ with unit $\mathbb{I}$ that is also an algebra over $\mathbb{C}$ and is equipped with an operation $^*:\mathcal{C}\to\mathcal{C}$ satisfying $(xy)^*=y^*x^*$, $(x+y)^*=x^*+y^*$, $x^{**}=x$, $(\lambda x)^*=\lambda^*x^*$, and $\|x^*x\|=\|x\|^2$, where λ^* is a complex conjugation of $\lambda\in\mathbb{C}$. A W^****-algebra*** is defined as such C^*-algebra that has a Banach predual. If a predual of C^*-algebra exists then it is unique. Given a W^*-algebra $\mathcal{N}$, we will denote its predual by $\mathcal{N}_\star$. Moreover, $\mathcal{N}_\star^+:=\{\phi\in\mathcal{N}_\star\mid\phi(x^*x)\geq 0\;\forall x\in\mathcal{N}\}$, $\mathcal{N}_{\star 0}^+:=\{\phi\in\mathcal{N}_\star^+\mid\omega(x^*x)=0\Rightarrow x=0\quad\forall x\in\mathcal{N}\}$, $\mathcal{N}_{\star 1}^+:=\{\phi\in\mathcal{N}_\star^+\mid\|\phi\|=1\}$, $\mathcal{N}^{\mathrm{sa}}:=\{x\in\mathcal{N}\mid x^*=x\}$, $\mathcal{N}^+:=\{x\in\mathcal{N}\mid\exists y\in\mathcal{N}\quad x=y^*y\}$, $\mathrm{Proj}(\mathcal{N}):=\{x\in\mathcal{N}^{\mathrm{sa}}\mid xx=x\}$. An element $x\in\mathcal{N}$ is called: ***partial isometry*** iff $x^*x\in\mathrm{Proj}(\mathcal{N})$; ***absolute value*** of $y\in\mathcal{N}$, denoted $x=|y|$, iff $y^*y=x^2$. The elements of $\mathcal{N}_\star^+$ will be called ***quantum states*** or ***states***. For $\mathcal{N}=\mathfrak{B}(\mathcal{H})$, where $\mathfrak{B}(\mathcal{H})$ is defined as the space of all bounded linear operators on the Hilbert space $\mathcal{H}$, $\mathcal{N}_\star=\mathfrak{G}_1(\mathcal{H}):=\{x\in\mathfrak{B}(\mathcal{H})\mid\|x\|_{\mathfrak{G}_1(\mathcal{H})}:=\mathrm{tr}(\sqrt{x^*x})<\infty\}$. If $(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$ is a localisable measure space, then $L_\infty(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$ is a commutative W^*-algebra, and $L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$ is its predual. Every commutative W^*-algebra can be represented in this form. This indicates how the theory of W^*-algebras generalises both the localisable measure theory and the theory of bounded operators over Hilbert spaces.

A ***weight*** on a W^*-algebra $\mathcal{N}$ is defined as a function $\omega:\mathcal{N}^+\to[0,+\infty]$ such that $\omega(0)=0$, $\omega(x+y)=\omega(x)+\omega(y)$, and $\lambda\geq 0\Rightarrow\omega(\lambda x)=\lambda\omega(x)$, with the convention $0\cdot(+\infty)=0$. A weight is called: ***faithful*** iff $\omega(x)=0\Rightarrow x=0$; ***finite*** iff $\omega(\mathbb{I})<\infty$; ***semi-finite*** iff a left ideal in $\mathcal{N}$ given by

$$\mathfrak{n}_\phi:=\{x\in\mathcal{N}\mid\phi(x^*x)<\infty\}\tag{67}$$

is weakly-$\star$ dense in $\mathcal{N}$; ***trace*** iff $\omega(xx^*)=\omega(x^*x)\;\forall x\in\mathcal{N}$; ***normal*** iff $\omega(\sup\{x_\iota\})=\sup\{\omega(x_\iota)\}$ for any uniformly bounded increasing net $\{x_\iota\}\subseteq\mathcal{N}^+$. A space of all normal semi-finite weights on a W^*-algebra $\mathcal{N}$ is denoted $\mathcal{W}(\mathcal{N})$, while the subset of all faithful elements of $\mathcal{W}(\mathcal{N})$ is denoted $\mathcal{W}_0(\mathcal{N})$. Every state is a finite normal weight, and every faithful state is a finite faithful normal state, hence the diagram

$$\begin{tikzcd}
\mathcal{N}_{\star 0}^+ \arrow[r, hook] \arrow[d, hook] & \mathcal{W}_0(\mathcal{N}) \arrow[d, hook] \\
\mathcal{N}_\star^+ \arrow[r, hook] & \mathcal{W}(\mathcal{N})
\end{tikzcd}\tag{68}$$

commutes. The domain of a weight ω can be extended by linearity to the topological $*$-algebra

$$\mathfrak{m}_\omega:=\mathrm{span}_{\mathbb{C}}\{x^*y\mid x,y\in\mathcal{N},\;\omega(x^*x)<\infty,\;\omega(y^*y)<\infty\}=\mathrm{span}_{\mathbb{C}}\{x\in\mathcal{N}^+\mid\omega(x)<\infty\}\subseteq\mathcal{N},\tag{69}$$

while ω can be extended to a positive linear functional on $\mathfrak{m}_\omega$, which coincides with ω on $\mathfrak{m}_\omega\cap\mathcal{N}^+$.

W^*-algebras for which there exists at least one faithful normal state are called ***countably finite***, while these for which there exists at least one faithful normal semi-finite trace are called ***semi-finite***. Every W^*-algebra admits at least one faithful normal semi-finite weight. A W^*-algebra is called: ***type I*** iff it is $*$-isomorphic to $\mathfrak{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$; ***type III*** iff it is not semi-finite; ***type II*** iff it is neither type I nor type III.

For $\psi\in\mathcal{W}(\mathcal{N})$,

$$\mathrm{supp}(\psi)=\mathbb{I}-\sup\{P\in\mathrm{Proj}(\mathcal{N})\mid\psi(P)=0\}.\tag{70}$$

For $\omega,\phi\in\mathcal{N}_\star^+$ we will write $\omega\ll\phi$ iff $\mathrm{supp}(\omega)\leq\mathrm{supp}(\phi)$.[5] An element $\omega\in\mathcal{N}^{\star+}$ is faithful iff $\mathrm{supp}(\omega)=\mathbb{I}$. If ϕ is a normal weight on a W^*-algebra $\mathcal{N}$ (which includes $\omega\in\mathcal{N}_\star^+$ as a special case), then the restriction of ϕ to a ***reduced*** W^*-algebra,

$$\mathcal{N}_{\mathrm{supp}(\phi)}:=\{x\in\mathcal{N}\mid\mathrm{supp}(\phi)x=x=x\,\mathrm{supp}(\phi)\}=\bigcup_{x\in\mathcal{N}}\{\mathrm{supp}(\phi)x\,\mathrm{supp}(\phi)\},\tag{71}$$

[5] If $\mathcal{N}=\mathfrak{B}(\mathcal{H})$ and $\omega=\mathrm{tr}(\rho_\omega\cdot)$ for $\rho_\omega\in\mathfrak{G}_1(\mathcal{H})^+$, then $\mathrm{supp}(\omega)=\mathrm{ran}(\rho_\omega)$, so for any $\phi=\mathrm{tr}(\rho_\phi\cdot)$ with $\rho_\phi\in\mathfrak{G}_1(\mathcal{H})^+$ one has $\omega\ll\phi$ iff $\mathrm{ran}(\rho_\omega)\subseteq\mathrm{ran}(\rho_\phi)$.

is a faithful normal weight (respectively, an element of $(\mathcal{N}_{\mathrm{supp}(\phi)})^+_{\star 0}$). If ϕ is semi-finite, then $\phi|_{\mathcal{N}_{\mathrm{supp}(\phi)}} \in \mathcal{W}_0(\mathcal{N}_{\mathrm{supp}(\phi)})$. Hence, given $\psi \in \mathcal{W}(\mathcal{N})$ and $P \in \mathrm{Proj}(\mathcal{N})$, $P = \mathrm{supp}(\psi)$ iff $\psi|_{\mathcal{N}_P} \in \mathcal{W}_0(\mathcal{N}_P)$ and $\psi(P) = \psi(PxP)$ $\forall x \in \mathcal{N}^+$. In particular, for $\omega, \phi \in \mathcal{N}_\star^+$ and $\omega \ll \phi$, we have $\omega|_{\mathcal{N}_{\mathrm{supp}(\phi)}} \in \mathcal{W}_0(\mathcal{N}_{\mathrm{supp}(\phi)})$.

A ***$*$-homomorphism*** of C^*-algebras $\mathcal{C}_1$ and $\mathcal{C}_2$ is defined as a map $\varsigma : \mathcal{C}_1 \to \mathcal{C}_2$ such that $\varsigma(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 \varsigma(x_1) + \lambda_2 \varsigma(x_2)$, $\varsigma(x_1 x_2) = \varsigma(x_1)\varsigma(x_2)$, $\varsigma(x^*) = \varsigma(x)^*$ for all $x, x_1, x_2 \in \mathcal{C}_1$. A $*$-homomorphism $\varsigma : \mathcal{C}_1 \to \mathcal{C}_2$ of C^*-algebras $\mathcal{C}_1$ and $\mathcal{C}_2$ is called: ***unital*** iff $\varsigma(\mathbb{I}) = \mathbb{I}$; a ***$*$-isomorphism*** iff $0 = \ker(\varsigma) := \{x \in \mathcal{C}_1 \mid \varsigma(x) = 0\}$. A ***representation*** of a C^*-algebra $\mathcal{C}$ is defined as a pair $(\mathcal{H}, \pi)$ of a Hilbert space $\mathcal{H}$ and a $*$-homomorphism $\pi : \mathcal{C} \to \mathfrak{B}(\mathcal{H})$. A representation $\pi : \mathcal{C} \to \mathfrak{B}(\mathcal{H})$ is called: ***nondegenerate*** iff $\{\pi(x)\xi \mid (x, \xi) \in \mathcal{C} \times \mathcal{H}\}$ is dense in $\mathcal{H}$; ***normal*** iff it is continuous with respect to the weak-$\star$ topologies of $\mathcal{C}$ and $\mathfrak{B}(\mathcal{H})$; ***faithful*** iff $\ker(\pi) = \{0\}$. An element $\xi \in \mathcal{H}$ is called ***cyclic*** for a C^*-algebra $\mathcal{C} \subseteq \mathfrak{B}(\mathcal{H})$ iff $\mathcal{C}\xi := \bigcup_{x \in \mathcal{C}} \{x\xi\}$ is norm dense in $\mathfrak{B}(\mathcal{H})$. A representation $\pi : \mathcal{C} \to \mathfrak{B}(\mathcal{H})$ of a C^*-algebra $\mathcal{C}$ is called ***cyclic*** iff there exists $\Omega \in \mathcal{H}$ that is cyclic for $\pi(\mathcal{C})$. According to the Gel'fand–Naĭmark–Segal theorem [274, 657] for every pair $(\mathcal{C}, \omega)$ of a C^*-algebra $\mathcal{C}$ and $\omega \in \mathcal{C}^{\star +}$ there exists a triple $(\mathcal{H}_\omega, \pi_\omega, \Omega_\omega)$ of a Hilbert space $\mathcal{H}_\omega$ and a cyclic representation $\pi_\omega : \mathcal{C} \to \mathfrak{B}(\mathcal{H})$ with a cyclic vector $\Omega_\omega \in \mathcal{H}_\omega$, and this triple is unique up to unitary equivalence. It is constructed as follows. For a C^*-algebra $\mathcal{C}$ and $\omega \in \mathcal{C}^{\star +}$, one defines the scalar form $\langle \cdot, \cdot \rangle_\omega$ on $\mathcal{C}$,

$$\langle x, y \rangle_\omega := \omega(x^* y) \;\; \forall x, y \in \mathcal{C}, \tag{72}$$

and the ***Gel'fand ideal***

$$\mathcal{I}_\omega := \{x \in \mathcal{C} \mid \omega(x^* x) = 0\} = \{x \in \mathcal{C} \mid \omega(x^* y) = 0 \; \forall y \in \mathcal{C}\}, \tag{73}$$

which is a left ideal of $\mathcal{C}$, closed in the norm topology (it is also closed in the weak-$\star$ topology if $\omega \in \mathcal{C}_\star^{\star +}$). The form $\langle \cdot, \cdot \rangle_\omega$ is hermitean on $\mathcal{C}$ and it becomes a scalar product $\langle \cdot, \cdot \rangle_\omega$ on $\mathcal{C}/\mathcal{I}_\omega$. The Hilbert space $\mathcal{H}_\omega$ is obtained by the completion of $\mathcal{C}/\mathcal{I}_\omega$ in the topology of norm generated by $\langle \cdot, \cdot \rangle_\omega$. Consider the morphisms

$$[\cdot]_\omega : \mathcal{C} \ni x \longmapsto [x]_\omega \in \mathcal{C}/\mathcal{I}_\omega, \tag{74}$$

$$\pi_\omega(y) : [y]_\omega \longmapsto [xy]_\omega. \tag{75}$$

The element $\omega \in \mathcal{C}^{\star +}$ is uniquely represented in terms of $\mathcal{H}_\omega$ by the vector $[\mathbb{I}]_\omega =: \Omega_\omega \in \mathcal{H}_\omega$, which is cyclic for $\pi_\omega(\mathcal{C})$ and satisfies $\|\Omega_\omega\| = \|\omega\|$. Hence

$$\omega(x) = \langle \Omega_\omega, \pi_\omega(x) \Omega_\omega \rangle_\omega \;\;\; \forall x \in \mathcal{C}, \tag{76}$$

An analogue of this theorem for weights follows the similar construction, but lacks cyclicity. If $\mathcal{N}$ is a W^*-algebra, and ω is a weight on $\mathcal{N}$, then there exists the Hilbert space $\mathcal{H}_\omega$, defined as the completion of $\mathfrak{n}_\omega / \ker(\omega)$ in the topology of a norm generated by the scalar product $\langle \cdot, \cdot \rangle_\omega : \mathfrak{n}_\omega \times \mathfrak{n}_\omega \ni (x, y) \mapsto \omega(x^* y) \in \mathbb{C}$,

$$\mathcal{H}_\omega := \overline{\mathfrak{n}_\omega / \ker(\omega)} = \overline{\{x \in \mathcal{N} \mid \omega(x^* x) < \infty\} / \{x \in \mathcal{N} \mid \omega(x^* x) = 0\}} = \overline{\mathfrak{n}_\omega / \mathcal{I}_\omega}, \tag{77}$$

and there exist the maps

$$[\cdot]_\omega : \mathfrak{n}_\omega \ni x \mapsto [x]_\omega \in \mathcal{H}_\omega, \tag{78}$$

$$\pi_\omega : \mathcal{N} \ni x \mapsto ([y]_\omega \mapsto [xy]_\omega) \in \mathfrak{B}(\mathcal{H}_\omega), \tag{79}$$

such that $[\cdot]_\omega$ is linear, $\mathrm{ran}([\cdot]_\omega)$ is dense in $\mathcal{H}_\omega$, and $(\mathcal{H}_\omega, \pi_\omega)$ is a representation of $\mathcal{N}$. If $\omega \in \mathcal{W}(\mathcal{N})$ then $(\mathcal{H}_\omega, \pi_\omega)$ is nondegenerate and normal. It is also faithful if $\omega \in \mathcal{W}_0(\mathcal{N})$.

The ***commutant*** of a subalgebra $\mathcal{N}$ of any algebra $\mathcal{C}$ is defined as

$$\mathcal{N}^\bullet := \{y \in \mathcal{C} \mid xy = yx \; \forall x \in \mathcal{N}\}, \tag{80}$$

while the ***center*** of $\mathcal{N}$ is defined as $\mathfrak{Z}_\mathcal{N} := \mathcal{N} \cap \mathcal{N}^\bullet$. A unital $*$-subalgebra $\mathcal{N}$ of an algebra $\mathfrak{B}(\mathcal{H})$ is called: a ***factor*** iff $\mathfrak{Z}_\mathcal{N} = \mathbb{C}\mathbb{I}$; the ***von Neumann algebra*** [750, 514] iff $\mathcal{N} = \mathcal{N}^{\bullet\bullet}$. An image

$\pi(\mathcal{N})$ of any representation $(\mathcal{H},\pi)$ of a W^*-algebra $\mathcal{N}$ is a von Neumann algebra iff π is normal and nondegenerate.

A subspace $\mathcal{D}\subseteq\mathcal{H}$ of a complex Hilbert space $\mathcal{H}$ is called a ***cone*** iff $\lambda\xi\in\mathcal{D}\;\forall\xi\in\mathcal{D}\;\forall\lambda\geq 0$. A cone $\mathcal{D}\subseteq\mathcal{H}$ is called ***self-polar*** iff

$$\mathcal{D}=\{\zeta\in\mathcal{H}\mid\langle\xi,\zeta\rangle_{\mathcal{H}}\geq 0\;\forall\xi\in\mathcal{D}\}. \tag{81}$$

Every self-polar cone $\mathcal{D}\subseteq\mathcal{H}$ is pointed ($\mathcal{D}\cap(-\mathcal{D})=\{0\}$), spans linearly $\mathcal{H}$ ($\mathrm{span}_{\mathbb{C}}\mathcal{D}=\mathcal{H}$), and determines a unique conjugation[6] J in $\mathcal{H}$ such that $J\xi=\xi\;\forall\xi\in\mathcal{H}$ [296], as well as a partial order on the set $\mathcal{H}^{\mathrm{sa}}:=\{\xi\in\mathcal{H}\mid J\xi=\xi\}$ given by,

$$\xi\leq\zeta\iff\xi-\zeta\in\mathcal{D}\;\;\forall\xi,\zeta\in\mathcal{H}^{\mathrm{sa}}. \tag{82}$$

If $\mathcal{N}$ is a W^*-algebra, $\mathcal{H}$ is a Hilbert space, $\mathcal{H}^\natural\subseteq\mathcal{H}$ is a self-polar cone, π is a nondegenerate faithful normal representation of $\mathcal{N}$ on $\mathcal{H}$, and J is conjugation on $\mathcal{H}$, then the quadruple $(\mathcal{H},\pi,J,\mathcal{H}^\natural)$ is called ***standard representation*** of $\mathcal{N}$ and $(\mathcal{H},\pi(\mathcal{N}),J,\mathcal{H}^\natural)$ is called ***standard form*** of $\mathcal{N}$ iff the conditions [297]

$$J\pi(\mathcal{N})J=\pi(\mathcal{N})^\bullet,\;\;\xi\in\mathcal{H}^\natural\Rightarrow J\xi=\xi,\;\;\pi(x)J\pi(x)J\mathcal{H}^\natural\subseteq\mathcal{H}^\natural,\;\;\pi(x)\in\mathfrak{Z}_{\pi(\mathcal{N})}\Rightarrow J\pi(x)J=\pi(x)^*. \tag{83}$$

hold. For any standard representation

$$\forall\phi\in\mathcal{N}_\star^+\;\exists!\xi_\pi(\phi)\in\mathcal{H}^\natural\;\forall x\in\mathcal{N}\;\;\phi(x)=\langle\xi_\pi(\phi),\pi(x)\xi_\pi(\phi)\rangle_{\mathcal{H}} \tag{84}$$

holds. The map $\xi_\pi:\mathcal{N}_\star^+\to\mathcal{H}^\natural$ is order preserving.

For a given W^*-algebra $\mathcal{N}$, $\phi\in\mathcal{W}(\mathcal{N})$, and $\omega\in\mathcal{W}_0(\mathcal{N})$ the map

$$R_{\phi,\omega}:[x]_\omega\mapsto[x^*]_\phi\;\;\forall x\in\mathfrak{n}_\omega\cap\mathfrak{n}_\phi^* \tag{85}$$

is a densely defined, closable antilinear operator. Its closure admits a unique polar decomposition

$$\overline{R}_{\phi,\omega}=J_{\phi,\omega}\Delta_{\phi,\omega}^{1/2}, \tag{86}$$

where $J_{\phi,\omega}$ is a conjugation operator, called ***relative modular conjugation***, while $\Delta_{\phi,\omega}$ is a positive self-adjoint operator on $\mathrm{dom}(\Delta_{\phi,\omega})\subseteq\mathcal{H}_\omega$ with $\mathrm{supp}(\Delta_{\phi,\omega})=\mathrm{supp}(\phi)\mathcal{H}_\omega$, called a ***relative modular operator*** [33, 166, 217]. We define $\Delta_\phi:=\Delta_{\phi,\phi}$. If $\mathcal{N}\cong\mathfrak{B}(\mathcal{H})$, $\phi=\mathrm{tr}(\rho_\phi\,\cdot\,)$, $\omega=\mathrm{tr}(\rho_\omega\,\cdot\,)$, $\mathfrak{L}_\rho$ denotes left multiplication by ρ, $\mathfrak{R}_\rho^{-1}$ denotes right multiplication by ρ^{-1}, then $\Delta_{\phi,\omega}=\mathfrak{L}_{\rho_\phi}\mathfrak{R}_{\rho_\omega}^{-1}$. The relative modular operators allow to define a one-parameter family of partial isometries in $\mathrm{supp}(\phi)\mathcal{N}$, called ***Connes' cocycle*** [165],

$$\mathbb{R}\ni t\mapsto[\phi:\omega]_t:=\Delta_{\phi,\psi}^{\mathrm{i}t}\Delta_{\omega,\psi}^{-\mathrm{i}t}=\Delta_{\phi,\omega}^{\mathrm{i}t}\Delta_{\omega,\omega}^{-\mathrm{i}t}\in\mathrm{supp}(\phi)\mathcal{N}, \tag{87}$$

where $\psi\in\mathcal{W}_0(\mathcal{N})$ is arbitrary, so it can be set equal to ω. Connes showed that $[\phi:\omega]_t$ can be characterised as a canonical object associated to any pair $(\phi,\omega)\in\mathcal{W}(\mathcal{N})\times\mathcal{W}_0(\mathcal{N})$ on any W^*-algebra $\mathcal{N}$, independently of any representation. As shown by Araki and Masuda [42] (see also [486]), the definition of $\Delta_{\phi,\omega}$ and $[\phi:\omega]_t$ can be further extended to the case when $\phi,\omega\in\mathcal{W}(\mathcal{N})$, by means of a densely defined closable antilinear operator

$$R_{\phi,\omega}:[x]_\omega+(\mathbb{I}-\mathrm{supp}(\overline{[\mathfrak{n}_\phi]_\omega}))\zeta\mapsto\mathrm{supp}(\omega)[x^*]_\phi\;\;\forall x\in\mathfrak{n}_\omega\cap\mathfrak{n}_\phi^*\;\forall\zeta\in\mathcal{H}, \tag{88}$$

where $(\mathcal{H},\pi,J,\mathcal{H}^\natural)$ is a standard representation of a W^*-algebra $\mathcal{N}$, and $\mathcal{H}_\phi\subseteq\mathcal{H}\supseteq\mathcal{H}_\omega$. For $\phi,\omega\in\mathcal{N}_\star^+$ this becomes a closable antilinear operator [40, 396]

$$R_{\phi,\omega}:x\xi_\pi(\omega)+\zeta\mapsto\mathrm{supp}(\omega)x^*\xi_\pi(\phi)\;\;\forall x\in\pi(\mathcal{N})\;\forall\zeta\in(\pi(\mathcal{N})\xi_\pi(\omega))^\perp, \tag{89}$$

[6]A linear operator $J:\mathrm{dom}(J)\to\mathcal{H}$, where $\mathrm{dom}(J)\subseteq\mathcal{H}$, is called a ***conjugation*** iff it is antilinear, isometric, and involutive ($J^2=\mathbb{I}$).

acting on a dense domain $(\pi(\mathcal{N})\xi_\pi(\omega))\cup(\pi(\mathcal{N})\xi_\pi(\omega))^\perp\subseteq\mathcal{H}$, where $(\pi(\mathcal{N})\xi_\pi(\omega))^\perp$ denotes a complement of the closure in $\mathcal{H}$ of the linear span of the action $\pi(\mathcal{N})$ on $\xi_\pi(\omega)$. In both cases, the relative modular operator is determined by the polar decomposition of the closure $\overline{R}_{\phi,\omega}$ of $R_{\phi,\omega}$,

$$\Delta_{\phi,\omega}:=R^*_{\phi,\omega}\overline{R}_{\phi,\omega}. \tag{90}$$

If (88) or (89) is used instead of (85), then the formula (87) has to be replaced by

$$\mathbb{R}\ni t\mapsto[\phi:\omega]_t\mathrm{supp}(\overline{[\mathfrak{n}_\phi]_\psi}):=\Delta^{\mathrm{i}t}_{\phi,\psi}\Delta^{-\mathrm{i}t}_{\omega,\psi}, \tag{91}$$

and $[\phi:\omega]_t$ is a partial isometry in $\mathrm{supp}(\phi)\mathcal{N}\mathrm{supp}(\omega)$ whenever $[\mathrm{supp}(\phi),\mathrm{supp}(\omega)]=0$. For $\phi,\psi\in\mathcal{W}_0(\mathcal{N})$ Connes' theorem [164, 165] states that the following conditions are equivalent:

i) $\exists\lambda>0\ \ \psi\leq\lambda\phi$,

ii) $x\in\mathfrak{n}_\phi\ \Rightarrow\ x\in\mathfrak{n}_\psi$,

iii) $t\mapsto[\psi:\phi]_t$ can be extended to a map that is valued in $\mathcal{N}$, bounded (by $\lambda^{1/2}$) and weakly-$\star$ continuous on a strip $\{z\in\mathbb{C}\mid\mathrm{im}(z)\in[-\frac{1}{2},0]\}$, holomorphic in interior of this strip, and satisfying the boundary condition

$$\psi(x)=\phi\left([\psi:\phi]_{-\mathrm{i}/2}{}^*x[\psi:\phi]_{-\mathrm{i}/2}\right)\ \ \forall x\in\mathfrak{m}_\psi. \tag{92}$$

This theorem extends to $\psi\in\mathcal{W}(\mathcal{N})$, with $\mathbb{R}\ni t\mapsto[\psi:\phi]_t\in\mathrm{supp}(\psi)\mathcal{N}\ \forall t\in\mathbb{R}$ [396]. Thus, whenever the condition i) is satisfied, the analytic continuation of Connes' cocycle

$$h^{1/2}=[\psi:\phi]_{-\mathrm{i}/2} \tag{93}$$

plays the role of a noncommutative (square root of) Radon–Nikodým quotient.

Recall that any weight on a W^*-algebra $\mathcal{N}$ can be uniquely extended to a linear functional on $\mathfrak{m}_\phi$ which coincides with ϕ on $\mathcal{N}^+\cap\mathfrak{m}_\phi$. Given a semi-finite trace $\tau:\mathcal{N}^+\to[0,\infty]$ on a semi-finite W^*-algebra $\mathcal{N}$, its extension to a two-sided ideal $\mathfrak{m}_\tau$ of $\mathcal{N}$ satisfies

$$\tau(yx)=\tau(xy)\ \ \forall x\in\mathfrak{m}_\tau\ \forall y\in\mathcal{N}. \tag{94}$$

In addition, if τ is normal, then for any $x\in\mathfrak{m}_\tau$ the map

$$y\mapsto\omega_x(y):=\tau(xy) \tag{95}$$

is an element of $\mathcal{N}^+_\star$ [238]. Moreover,

$$\tau(yx)=\tau(x^{1/2}yx^{1/2})=\tau(y^{1/2}xy^{1/2})\ \ \forall x\in\mathfrak{m}^+_\tau\ \forall y\in\mathcal{N}^+. \tag{96}$$

So, the formula

$$\omega_x(y):=\tau(x^{1/2}yx^{1/2})\ \ \forall y\in\mathcal{N} \tag{97}$$

gives rise to $\omega_x\in\mathcal{N}^+_\star$ with $\|\omega_x\|=\tau(|x|)$ for each $x\in\mathfrak{m}_\tau$. Let τ be a faithful normal semi-finite trace on a W^*-algebra $\mathcal{N}$. The map

$$\|\cdot\|_p:\mathcal{N}\ni x\mapsto\|x\|_p:=\tau(|x|^p)^{1/p}\in[0,\infty] \tag{98}$$

for $p\in[1,\infty[$ is a norm on a vector space $\{x\in\mathcal{N}\mid\|x\|_p<\infty\}$. Denote the Cauchy completion of this normed vector space by $L_p(\mathcal{N},\tau)$. Equivalently, $L_p(\mathcal{N},\tau)$ can be defined as a Cauchy completion of $\{x\in\mathcal{N}\mid\tau(|x|)<\infty\}$ in the norm given by $\|\cdot\|_p$ [533], or as a Cauchy completion of $\mathrm{span}_{\mathbb{C}}\{x\in\mathcal{N}^+\mid\tau(\mathrm{supp}(x))<\infty\}$ in $\|\cdot\|_p$ [588]. The space $L_1(\mathcal{N},\tau)$ can be equivalently defined also as a Cauchy completion of $\mathfrak{m}_\tau$ in $\|\cdot\|_1$, while $L_2(\mathcal{N},\tau)$ as a Cauchy completion of $\mathfrak{n}_\tau$ in $\|\cdot\|_2$ [226, 708]. The property

$|\tau(x)| \leq \|x\|_1\ \forall x \in \mathfrak{m}_\tau$ allows the unique continuous extension of τ from a linear functional on $\mathfrak{m}_\tau$ to a linear functional on $L_1(\mathcal{N},\tau)$. This extends a bilinear form

$$\mathfrak{m}_\tau \times \mathcal{N} \ni (h,x) \mapsto \tau(h^{1/2}xh^{1/2}) \in \mathbb{C} \tag{99}$$

to the bilinear form $L_1(\mathcal{N},\tau) \times \mathcal{N} \to \mathbb{C}$, which defines a duality between $L_1(\mathcal{N},\tau)$ and $\mathcal{N}$, and makes $L_1(\mathcal{N},\tau)$ isometrically isomorphic to $\mathcal{N}_\star$ [226]. Extending the notation ω_x of (95) to all elements of $\mathcal{N}_\star$ corresponding to $x \in L_1(\mathcal{N},\tau)$, we have

$$\omega_x(y) = \tau(yx) = \tau(xy)\ \ \forall y \in \mathcal{N}\ \forall x \in L_1(\mathcal{N},\tau), \tag{100}$$

and [238, 659]

$$\forall \omega \in \mathcal{N}_\star^+\ \ \exists! x \in L_1(\mathcal{N},\tau)^+\ \ \forall y \in \mathcal{N}\ \ \omega(y) = \tau(xy) = \tau(x^{1/2}yx^{1/2}). \tag{101}$$

Such x will be called a ***Dye–Segal density*** of ω with respect to τ.

A closed densely defined linear operator $x : \mathrm{dom}(x) \to \mathcal{H}$ with $\mathrm{dom}(x) \subseteq \mathcal{H}$ and polar decomposition $x = v|x|$ will be called ***affiliated*** with a von Neumann algebra $\mathcal{C}$ acting on $\mathcal{H}$ iff $v \in \mathcal{C}$ and all spectral projections of $|x|$ belong to $\mathcal{C}$. Let τ be a fixed faithful normal semi-finite trace on a W^*-algebra $\mathcal{N}$. Using the notion of measurability with respect to a trace τ, the above range of $L_p(\mathcal{N},\tau)$ spaces can be represented in terms of operators affiliated to a von Neumann algebra $\pi_\tau(\mathcal{N})$ acting on $\mathcal{H}_\tau$, where $(\mathcal{H}_\tau,\pi_\tau)$ is the GNS Hilbert space of $(\mathcal{N},\tau)$. A closed densely defined linear operator $x : \mathrm{dom}(x) \to \mathcal{H}$ is called ***τ-measurable*** [659, 533] iff $\exists \lambda > 0\ \ \tau(P^{|x|}(]\lambda,+\infty[)) < \infty$. The space of all τ-measurable operators affiliated with $\pi_\tau(\mathcal{N})$ will be denoted by $\mathscr{M}(\mathcal{N},\tau)$. For $x,y \in \mathscr{M}(\mathcal{N},\tau)$ the algebraic sum $x+y$ and algebraic product xy may not be closed, hence in general they do not belong to $\mathscr{M}(\mathcal{N},\tau)$. However, their closures (denoted with the abuse of notation by the same symbol) belong to $\mathscr{M}(\mathcal{N},\tau)$. See [513] for further discussion of $\mathscr{M}(\mathcal{N},\tau)$ and its topologies. Consider the extension of a trace τ from $\mathcal{N}^+$ to $\mathscr{M}(\mathcal{N},\tau)^+$ given by

$$\tau : \mathscr{M}(\mathcal{N},\tau)^+ \ni x \mapsto \tau(x) := \sup_{n\in\mathbb{N}} \left\{ \tau\left(\int_0^n P^x(\lambda)\lambda \right) \right\} \in [0,\infty], \tag{102}$$

the map

$$\|\cdot\|_p : \mathscr{M}(\mathcal{N},\tau) \ni x \mapsto \|x\|_p := (\tau(|x|^p))^{1/p} \in [0,\infty], \tag{103}$$

and the family of vector spaces

$$L_p(\mathcal{N},\tau) := \{x \in \mathscr{M}(\mathcal{N},\tau) \mid \|x\|_p < \infty\}, \tag{104}$$

where $p \in [1,\infty[$. The map (103) is a norm on (104) [775], and $L_p(\mathcal{N},\tau)$ are Cauchy complete with respect to the topology of this norm. In addition, one defines $L_\infty(\mathcal{N}) := \mathcal{N}$. The Banach spaces $L_p(\mathcal{N},\tau)$ defined this way coincide with the $L_p(\mathcal{N},\tau)$ spaces defined before. The spaces $L_p(\mathcal{N},\tau)$ embed continuously into $\mathscr{M}(\mathcal{N},\tau)$ [533]. For all $\gamma \in]0,1]$, the duality

$$L_{1/\gamma}(\mathcal{N},\tau) \times L_{1/(1-\gamma)}(\mathcal{N},\tau) \ni (x,y) \mapsto [\![x,y]\!] := \tau(xy) \in \mathbb{R} \tag{105}$$

determines an isometric isomorphism of Banach spaces

$$L_{1/\gamma}(\mathcal{N},\tau)^\star \cong L_{1/(1-\gamma)}(\mathcal{N},\tau). \tag{106}$$

Now we will consider the special case of the above spaces. The space of ***Riesz–Schauder*** [622, 654] (or ***compact***) operators over a Hilbert space $\mathcal{H}$,

$$\mathfrak{G}_0(\mathcal{H}) := \overline{\{x \in \mathfrak{B}(\mathcal{H}) \mid \dim\mathrm{ran}(x) \leq \infty\}}, \tag{107}$$

where bar denotes the Cauchy completion in the norm of $\mathfrak{B}(\mathcal{H})$, allows to define the space $\mathfrak{G}_1(\mathcal{H})$ of ***trace class*** (or ***nuclear***) operators [651, 652] and the space $\mathfrak{G}_2(\mathcal{H})$ of ***Hilbert–Schmidt*** operators

[655, 747, 692] as a Cauchy completion of $\mathfrak{G}_0(\mathcal{H})$ in the norm $\|x\|_1 := \mathrm{tr}(|\sqrt{x^*x}|)$ and $\|x\|_2 := \mathrm{tr}(x^*x)$, respectively. More generally, the spaces $\mathfrak{G}_p(\mathcal{H})$ of ***von Neumann–Schatten p-class*** operators over a Hilbert space $\mathcal{H}$ are defined as [751, 648, 651, 652, 649]

$$\mathfrak{G}_p(\mathcal{H}) := \{x \in \mathfrak{G}_0(\mathcal{H}) \mid \|x\|_p := \mathrm{tr}((x^*x)^{p/2})^{1/p} < \infty\}, \tag{108}$$

for $p \in [1, \infty[$, and they are Banach spaces with respect to the norm $\|\cdot\|_p$ for $p \in [1, \infty[$. In addition, one sets $\mathfrak{G}_\infty(\mathcal{H}) := \mathfrak{B}(\mathcal{H})$ with $\|x\|_\infty := \|x\|_{\mathfrak{B}(\mathcal{H})}$. The spaces $\mathfrak{G}_p(\mathcal{H})$ are uniformly convex and uniformly Fréchet differentiable for $p \in]1, \infty[$ [226, 493, 425], and the following Banach space dualities hold [650, 493]:

$$\mathfrak{G}_0(\mathcal{H})^\star \cong \mathfrak{G}_1(\mathcal{H}), \quad \mathfrak{G}_1(\mathcal{H})^\star \cong \mathfrak{G}_\infty(\mathcal{H}), \quad \mathfrak{G}_{1/\gamma}(\mathcal{H})^\star \cong \mathfrak{G}_{1/(1-\gamma)}(\mathcal{H}), \tag{109}$$

for $\gamma \in]0, 1]$. If $\mathcal{N} \subseteq \mathfrak{B}(\mathcal{H})$, then [226, 227]

$$\forall \omega \in \mathcal{N}^\star \quad \big(\omega \in \mathcal{N}_\star \quad \iff \quad \exists x \in \mathfrak{G}_1(\mathcal{H}) \quad \omega(\cdot) = \mathrm{tr}_{\mathfrak{B}(\mathcal{H})}(x\,\cdot)\big)\,. \tag{110}$$

In such case $\|\omega\| = \mathrm{tr}(x)$. This theorem holds also for $(\omega, x) \in \mathcal{N}_\star^+ \times \mathfrak{G}_1(\mathcal{H})^+$, as well as for $(\omega, x) \in \mathcal{N}_{\star 1}^+ \times \mathfrak{G}_1(\mathcal{H})_1^+$. However, the uniqueness of x in (110), as well as in its positive and normalised cases, holds only for $\mathcal{N} = \mathfrak{B}(\mathcal{H})$, because in such case (110) defines a linear isometry $L_1(\mathfrak{B}(\mathcal{H}), \mathrm{tr}) \cong \mathfrak{G}_1(\mathcal{H}) \cong \mathfrak{B}(\mathcal{H})_\star$ [225, 649]. More generally, if $\mathcal{N} \subseteq \mathfrak{B}(\mathcal{H})$, then [226]

$$\mathcal{N}_\star \cong \mathfrak{G}_1(\mathcal{H})/\{x \in \mathfrak{G}_1(\mathcal{H}) \mid \mathrm{tr}(xy) = 0 \;\forall y \in \mathcal{N}\}. \tag{111}$$

The space $\mathfrak{G}_2(\mathcal{H})$ can be equipped with the inner product

$$\langle x, y\rangle_{\mathfrak{G}_2(\mathcal{H})} := \mathrm{tr}(y^*x) \;\; \forall x, y \in \mathfrak{G}_2(\mathcal{H}), \tag{112}$$

which turns it into a Hilbert space, called the ***Hilbert–Schmidt space***. The von Neumann–Schatten $\mathfrak{G}_p(\mathcal{H})$ spaces can be characterised by

$$\mathfrak{G}_p(\mathcal{H}) = L_p(\mathfrak{B}(\mathcal{H}), \mathrm{tr}) \;\; \forall p \in [1, \infty]. \tag{113}$$

Falcone and Takesaki [247] have constructed a family of noncommutative $L_p(\mathcal{N})$ spaces that are canonically associated to every W^*-algebra, including also those that do not admit faithful normal semi-finite traces. For a detailed review of this construction, see [404]. Here we will need only several facts about them. Its key feature is a construction of a semi-finite von Neumann algebra $\widetilde{\mathcal{N}}$ and a faithful normal semi-finite trace $\widetilde{\tau} : \widetilde{\mathcal{N}} \to [0, \infty]$ that are uniquely defined for any W^*-algebra $\mathcal{N}$, with no dependence of an additional weight or state on $\mathcal{N}$. Using these objects, a topological $*$-algebra $\mathscr{M}(\widetilde{\mathcal{N}}, \widetilde{\tau})$ of $\widetilde{\tau}$-measureable operators is defined, as well as a canonical integral $\int : \mathscr{M}(\widetilde{\mathcal{N}}, \widetilde{\tau}) \to \mathbb{C}$. All spaces $L_p(\mathcal{N})$ for $p \in \mathbb{C}$ and such that $\mathrm{re}\,(p) > 0$ are Banach spaces (the corresponding norms will be denoted $\|\cdot\|_p$) and their Banach duals are given by $L_q(\mathcal{N})$ spaces with $\frac{1}{p} + \frac{1}{q} = 1$. The space $L_\infty(\mathcal{N})$ is defined as $\mathcal{N}$, and an isometric isomorphism $\mathcal{N}_\star \cong L_1(\mathcal{N})$ holds. All $L_p(\mathcal{N})$ spaces with $p \in \mathbb{C} \cup \{+\infty\}$ and $\mathrm{re}\,(p) > 0$ embed into $\mathscr{M}(\widetilde{\mathcal{N}}, \widetilde{\tau})$. It is equipped with a function $\mathrm{grad} : \mathscr{M}(\widetilde{\mathcal{N}}, \widetilde{\tau}) \to \mathbb{C}$ satisfying

$$\mathrm{grad}(x^*) = (\mathrm{grad}(x))^*, \tag{114}$$
$$\mathrm{grad}(|x|) = \mathrm{re}\,(\mathrm{grad}(x)) = \tfrac{1}{2}(\mathrm{grad}(x) + \mathrm{grad}(x)^*), \tag{115}$$
$$\mathrm{grad}(\overline{xy}) = \mathrm{grad}(x) + \mathrm{grad}(y), \tag{116}$$
$$\mathrm{re}\,(\mathrm{grad}(x)) \geq 0 \Rightarrow |x|^{1/\mathrm{re}\,(\mathrm{grad}(x))} \in \mathcal{N}_\star^+, \tag{117}$$

where $\overline{xy}$ is the closure of xy. The Falcone–Takesaki ***canonical integral*** $\int : (\widetilde{\mathcal{N}}, \widetilde{\tau}) \to \mathbb{C}$ satifies $\int : L_1(\mathcal{N}) \ni \phi \mapsto \int \phi = \phi(\mathbb{I}) \in \mathbb{C}$, the norms $\|\cdot\|_p$ for $p \in \mathbb{C}$ and $\mathrm{re}\,(p) \geq 1$ read

$$\|\cdot\|_p : L_p(\mathcal{N}) \ni x \mapsto \|x\|_p := \left(\int |x|^{\mathrm{re}\,(p)}\right)^{1/\mathrm{re}\,(p)} \in \mathbb{R}^+, \tag{118}$$

while the Banach space duality between $L_p(\mathcal{N})$ and $L_q(\mathcal{N})$ for $1/p+1/q=1$ and $p\in\{\lambda\in\mathbb{C}\mid\mathrm{re}\,(\lambda)>0\}$ reads

$$L_p(\mathcal{N})\times L_q(\mathcal{N})\ni(x,y)\mapsto[\![x,y]\!]_{\widetilde{\mathcal{N}}}:=\int xy\in\mathbb{C}. \tag{119}$$

Moreover, the space $L_2(\mathcal{N})$ is also a Hilbert space with respect to the inner product

$$L_2(\mathcal{N})\times L_2(\mathcal{N})\ni(x_1,x_2)\mapsto\langle x_1,x_2\rangle_{L_2(\mathcal{N})}:=\int x_2^*x_1\in\mathbb{C}. \tag{120}$$

If $\{x_i\}_{i=1}^n\subseteq\mathscr{M}(\widetilde{\mathcal{N}},\widetilde{\tau})$, $\sum_{i=1}^n\mathrm{grad}(x_i)=:r\leq 1$ and $\mathrm{re}\,(\mathrm{grad}(x_i))\geq 0$ $\forall i\in\{1,\ldots,n\}$, then the noncommutative analogue of the Rogers–Hölder inequality holds [399],

$$\|x_1\cdots x_n\|_{1/r}\leq\|x_1\|_{1/\mathrm{re}\,(\mathrm{grad}(x_1))}\cdots\|x_n\|_{1/\mathrm{re}\,(\mathrm{grad}(x_n))}. \tag{121}$$

The stronger condition $\sum_{i=1}^n\mathrm{grad}(x_i)=1$ implies that $x_1\cdots x_n\in L_1(\mathcal{N})$, and in such case

$$\int x_1\cdots x_n=\int x_nx_1\cdots x_{n-1}. \tag{122}$$

Consider a W^*-algebra $\mathcal{N}$ and a relation $\sim_t$ on $\mathcal{N}\times\mathcal{W}_0(\mathcal{N})$ defined by [247]

$$(x,\psi)\sim_t(y,\phi)\iff y=x[\psi:\phi]_t\ \ \forall x,y\in\mathcal{N}\ \forall\psi,\phi\in\mathcal{W}_0(\mathcal{N}). \tag{123}$$

The property $[\omega_1:\omega_2]_t[\omega_2:\omega_3]_t=[\omega_1:\omega_3]_t$ $\forall\omega_1,\omega_2,\omega_3\in\mathcal{W}_0(\mathcal{N})$ $\forall t\in\mathbb{R}$ of Connes' cocycle implies that $\sim_t$ is an equivalence relation in $\mathcal{N}\times\mathcal{W}_0(\mathcal{N})$. The equivalence class $(\mathcal{N}\times\mathcal{W}_0(\mathcal{N}))/\sim_t$ is denoted by $\mathcal{N}(t)$, and its elements are denoted by $x\psi^{\mathrm{i}t}$. The operations

$$x\psi^{\mathrm{i}t}+y\psi^{\mathrm{i}t}:=(x+y)\psi^{\mathrm{i}t}, \tag{124}$$
$$\lambda(x\psi^{\mathrm{i}t}):=(\lambda x)\psi^{\mathrm{i}t}\ \forall\lambda\in\mathbb{C}, \tag{125}$$
$$\left\|x\psi^{\mathrm{i}t}\right\|:=\|x\|, \tag{126}$$

equip $\mathcal{N}(t)$ with the structure of the Banach space, which is isometrically isomorphic to $\mathcal{N}$, considered as a Banach space. By definition, $\mathcal{N}(0)$ a W^*-algebra that is trivially $*$-isomorphic to $\mathcal{N}$. For $t\neq 0$ the spaces $\mathcal{N}(t)$ are not W^*-algebras, however

$$L_{1/\mathrm{i}t}(\mathcal{N})=\mathcal{N}(t)\ \ \forall t\in\mathbb{R}. \tag{127}$$

This suggests to use the symbolic notation $y=x\phi^{\mathrm{grad}(y)}=x\phi^{\gamma}$ with $(x,\phi)\in\mathcal{N}\times\mathcal{W}_0(\mathcal{N})$ for a generic element y of the space $L_{1/\gamma}(\mathcal{N})$ with $\mathrm{re}\,(\gamma)\in]0,1[$, with boundary cases given by $x\in L_\infty(\mathcal{N})=\mathcal{N}$ and $\phi\in L_1(\mathcal{N})\cong\mathcal{N}_\star$. For $x_i=y_i\phi^{z_i}$ with a fixed $\phi\in\mathcal{N}_{\star 0}^+$, the equation (122) turns to the Araki multiple KMS condition for σ^ϕ and $\beta=1$ [31, 33, 42, 485]. More generally, the function

$$\mathbb{C}^n\ni(z_1,\ldots,z_n)\mapsto\phi_1^{z_1}y_1\cdots\phi_n^{z_n}y_n\phi_{n+1}^{1-z_1-\ldots-z_n}\in\mathcal{N}_\star \tag{128}$$

is a bounded holomorphic function on the tube

$$\{(z_1,\ldots,z_n)\in\mathbb{C}^n\mid\mathrm{re}\,(z_i)>0\ \ \forall i\in\{1,\ldots,n\},\ \sum_{i=1}^n\mathrm{re}\,(z_i)\leq 1\}, \tag{129}$$

with respect to the norm topology of $\mathcal{N}_\star$ [773]. The above algebraic relations can be used in order to rewrite Connes' cocycle as

$$[\omega:\phi]_t=\Delta_{\omega,\phi}^{\mathrm{i}t}\Delta_\phi^{-\mathrm{i}t}=\omega^{\mathrm{i}t}\phi^{-\mathrm{i}t}, \tag{130}$$

which holds for all $\phi,\omega\in\mathcal{W}_0(\mathcal{N})$, and for all $\phi,\omega\in\mathcal{N}_\star^+$ provided $\mathrm{supp}(\omega)\leq\mathrm{supp}(\phi)$, and to rewrite the Tomita–Takesaki modular automorphism as

$$\sigma_t^\phi(x)=\Delta_\phi^{\mathrm{i}t}x\Delta_\phi^{-\mathrm{i}t}=\phi^{\mathrm{i}t}x\phi^{-\mathrm{i}t}, \tag{131}$$

which holds for all $\phi \in \mathcal{W}_0(\mathcal{N})$, and for all $\phi \in \mathcal{W}(\mathcal{N})$, provided $x \in \mathcal{N}_{\mathrm{supp}(\phi)}$. These remarkable algebraic properties were observed by Woronowicz [771] and were later developed by Connes [168, 169, 170] and Yamagami [773, 774]. For the negative powers of weights, ϕ^{-p} for $p > 0$, there are no corresponding $L_{-p}(\mathcal{N})$ spaces. However, as shown in [674], the right and left multiplications, $\mathfrak{R}(\phi^{-p})$ and $\mathfrak{L}(\phi^{-p})$, for $\phi \in \mathcal{W}_0(\mathcal{N})$ are well defined[7] and satisfy $\mathfrak{R}(\phi^{-p}) = (\mathfrak{R}(\phi^p))^{-1}$, $\mathfrak{L}(\phi^{-p}) = (\mathfrak{L}(\phi^p))^{-1}$, $\mathfrak{R}(\phi^{-p})\mathfrak{R}(\phi^p) = \mathbb{I}$, as well as

$$\Delta^{1/p}_{\phi,\psi} = \mathfrak{R}(\phi^{-1/p})\mathfrak{L}(\psi^{1/p}), \tag{132}$$

where $\psi \in \mathcal{W}(\mathcal{N})$. This gives

$$\int \psi^\gamma \phi^{1-\gamma} = \int \psi^\gamma \phi^{-\gamma}\phi = \int (\mathfrak{R}(\phi^{-\gamma})\mathfrak{L}(\psi^\gamma)\mathbb{I})\phi = \phi(\mathfrak{R}(\phi^{-\gamma})\mathfrak{L}(\psi^\gamma)\mathbb{I}) = \left\langle \xi_\pi(\phi), \Delta^\gamma_{\psi,\phi}\xi_\pi(\phi)\right\rangle_{\mathcal{H}} \tag{133}$$

for any standard representation $(\mathcal{H}, \pi, J, \mathcal{H}^\natural)$. In analogy with the equations (130) and (131), the equation (133) holds also when $\phi, \psi \in \mathcal{N}^+_\star$ and $\psi \ll \phi$, because in such case ϕ is faithful on $\mathcal{N}_{\mathrm{supp}(\phi)}$ and this algebra contains the support of ϕ.

The $L_p(\mathcal{N})$ spaces defined above are isometrically isomorphic with the $L_p(\mathcal{N}, \psi)$ spaces of Haagerup–Terp [298, 713], Araki–Masuda [42, 485], Kosaki–Terp [398, 714], and Kosaki [396], which are all uniformly convex and uniformly Fréchet differentiable for $p \in]1, \infty[$ (for proofs, see [713], [42, 485], [398], and [396], respectively).[8]

The canonical (representation independent) character of the Falcone–Takesaki 'noncommutative integral' $\int$ corresponds to the canonical (representation independent) character of Connes' cocycle as the noncommutative analogue of the Radon–Nikodým quotient.

The roles played in the commutative integration theory by mcb-algebras $\mathcal{A}$ and their representations in terms of measurable spaces $(\mathcal{X}, \mho(\mathcal{X}), \mho^0(\mathcal{X}))$ or measure spaces $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ are analogous to the roles played in the noncommutative integration theory by, respectively, W^*-algebras $\mathcal{N}$ and their standard representations $(\mathcal{H}, \pi(\mathcal{N}), J, \mathcal{H}^\natural)$ or the GNS representations $(\mathcal{H}_\omega, \pi_\omega, \Omega_\omega)$. In particular, if $\mathcal{N}$ is commutative, then each $\mu_\psi \in \mathcal{W}(\mathcal{A})$ determines a normal semi-finite trace on $\mathcal{N}$ by

$$\psi(x) = \int \mu_\psi x \;\; \forall x \in \mathcal{N}^+. \tag{134}$$

If $\mu_\phi \in \mathcal{W}_0(\mathcal{A})$ corresponds to $\phi \in \mathcal{W}_0(\mathcal{N})$ and $f \in L_1(\mathcal{A}, \mu_\phi)$ is its Radon–Nikodým quotient with respect to $\mu_\psi \in \mathcal{W}(\mathcal{A})$, $f = \frac{\mu_\psi}{\mu_\phi}$, which means

$$\psi(x) = \int \mu_\psi x = \int \mu_\phi \frac{\mu_\psi}{\mu_\phi} x = \phi\left(\frac{\mu_\psi}{\mu_\phi} x\right) \;\; \forall x \in L_\infty(\mathcal{A})^+, \tag{135}$$

then the faithfulness of ψ corresponds to strict positivity of μ_ψ and implies $\frac{\mu_\psi}{\mu_\phi} > 0$. In such case, the map $\mathbb{R} \ni t \mapsto \left(\frac{\mu_\psi}{\mu_\phi}\right)^{\mathrm{i}t} \in \mathcal{N}$ satisfies

$$\left(\frac{\mu_\psi}{\mu_\phi}\right)^{\mathrm{i}t} = [\psi : \phi]_t \;\; \forall t \in \mathbb{R}. \tag{136}$$

The boolean ideals $\mathcal{A}^\mu \subseteq \mathcal{A}$ for $\mu \in \mathcal{W}(\mathcal{A})$ play the role analogous to the ideals $\mathfrak{n}_\psi \subseteq \mathcal{N}$ for $\psi \in \mathcal{W}(\mathcal{N})$. In particular, the compatibility condition $x \in \mathcal{A}^{\mu_1} \Rightarrow \exists y \leq x \; y \in \mathcal{A}^{\mu_2}$ plays a crucial role in the definition of the Radon–Nikodým quotient $\frac{\mu_1}{\mu_2}$ of $\mu_1, \mu_2 \in \mathcal{W}_0(\mathcal{A})$, which corresponds to the crucial role

[7]More precisely, let the adjective 'strong' refers to the topological closure of some algebraic operation in $\mathscr{M}(\widetilde{\mathcal{N}}, \widetilde{\tau})$. For any $\lambda \geq 0$, $t > 0$, $\phi \in \mathcal{N}^+_{\star 0}$, the map $\mathfrak{R}(\phi^t) : L_{1/\lambda}(\mathcal{N}) \to L_{1/(\lambda+t)}(\mathcal{N})$, defined as a strong composition with ϕ^t from right, is everywhere defined, bounded, and injective with dense range. Moreover, the maps $\mathfrak{R}(\phi^t)^{-1}$ and $\mathfrak{R}(\phi^{-t})$ have the same range and agree (from this it follows that they are equal). The map $\mathfrak{R}(\phi^{-t})$ is closed, and is understood as a strong product, defined only when the closure is $\widetilde{\tau}$-measurable. The same holds for $\mathfrak{R}$ replaced by $\mathfrak{L}$. If $\phi \in \mathcal{N}^+_{\star 0}$ is replaced by $\phi \in \mathcal{W}_0(\mathcal{N})$, then all those properties hold except that $\mathfrak{R}(\phi^t)$ and $\mathfrak{L}(\phi^t)$ are no longer everywhere defined or bounded.

[8]See Section 3.2 for the definitions of uniform convexity and uniform Fréchet differentiability.

played by the condition $x \in \mathfrak{n}_{\psi_1} \Rightarrow x \in \mathfrak{n}_{\psi_2}$ in Connes' theorem on the extension of Connes' cocycle $[\psi_1 : \psi_2]_t$ of $\psi_1, \psi_2 \in \mathcal{W}_0(\mathcal{N})$ to the (square root of) noncommutative analogue of the Radon–Nikodým quotient, $[\psi : \phi]_{-\mathrm{i}/2}$.

In order to keep the algebraic representation-independent formulation, we will prove that proper abstract L_∞ spaces coincide with the commutative W^*-algebras without refering to measure spaces. This defines an equivalence between the category of mcb-algebras (with boolean isomorphisms), commutative W^*-algebras (with normal unital $*$-isomorphisms), and proper abstract L_∞ spaces (with unit preserving isometric Riesz isomorphisms). Together with the full and faithful functor from the category of $L_p(\mathcal{A})$ spaces with isometric Riesz isomorphisms to the category of $L_p(\mathcal{N})$ spaces with isometric isomorphisms, this shows that canonical commutative integration theory is precisely a commutative sector of canonical noncommutative integration theory.

Proposition 2.4. *The categories of commutative W^*-algebras with unital normal $*$-homomorphisms and proper L_∞ spaces with order continuous unit preserving Riesz homomorphisms are equivalent, and the same holds for mutual restriction of homomorphisms to isomorphisms.*

Proof. Using Freudenthal's spectral theorem [263], Lyubovin [474, 475] and Vulikh [752, 753] proved that each commutative von Neumann algebra is a Dedekind–MacNeille complete Banach lattice (for earlier proofs of this result, depedending on Gel'fand's representation theorem, see [375, 250]), while Luxemburg and Zaanen [473] proved that each commutative von Neumann algebra is an MI-space. Both proofs hold for arbitrary W^*-algebra. Taking into account that each W^*-algebra has a unique Banach predual, we conclude that each commutative W^*-algebra is a complex *proper* abstract L_∞ space. Conversely, each real abstract L_∞ space X has a form $L_\infty(\mathcal{A})$ over a ccb-algebra $\mathcal{A}$ of projection bands of X. Hence (see e.g. [262]) X is a real commutative Banach algebra and an archimedean real f-algebra. As an f-algebra, it satisfies $\left|y^2\right| = |y|^2$, where $y^2 := y \cdot y$. As an archimedean f-algebra it satisfies [333]

$$\forall x \in X \quad x \geq 0 \iff \exists! y \in X \quad x = y^2, \tag{137}$$

while as a Banach lattice it satisfies $|x| \leq |y| \Rightarrow \|x\|_X \leq \|y\|_X$. Hence, X satisfies $\left\|y^2\right\|_X = \|y\|_X^2$. Its Banach algebra complexification $X_\mathbb{C} := X + \mathrm{i}X$, equipped with multiplication, involution, and norm:

$$(x_1 + \mathrm{i}x_2) \cdot (y_1 + \mathrm{i}y_2) := (x_1 y_1 - x_2 y_2) + \mathrm{i}(x_1 y_1 + x_2 y_2), \tag{138}$$

$$(x_1 + \mathrm{i}x_2)^* := x_1 - \mathrm{i}x_2, \tag{139}$$

$$\|x + \mathrm{i}y\|_{X_\mathbb{C}} := \left\|x^2 + y^2\right\|_X^{1/2}, \tag{140}$$

is a commutative C^*-algebra (see e.g. [742]). On the other hand, the Banach lattice complexification $\tilde{X}_\mathbb{C}$ of X is equipped with the norm [71]

$$\|x + \mathrm{i}y\|_{\tilde{X}_\mathbb{C}} := \||x + \mathrm{i}y|\|_X = \left\|\sqrt{x^2 + y^2}\right\|_X. \tag{141}$$

These two complexifications coincide, because $\left\|x^{1/2}\right\|_X = \|x\|_X^{1/2}$ $\forall x = y^2 \geq 0$. Thus, every abstract L_∞ space is a commutative C^*-algebra with the multiplicative unit $\mathbb{I}$ given by the order unit. The Dedekind–MacNeille completeness of $X_\mathbb{C}$ implies that its boolean algebra $\mathcal{A}$ of projection bands is a Dcb-algebra, hence $X_\mathbb{C}$ is a commutative AW^*-algebra [376]. Finally, the existence of a unique predual turns X into a commutative W^*-algebra. Every algebra homomorphism f of f-algebras with multiplicative unit element is a Riesz homomorphism iff it satisfies $f(|x|) = |f(x)|$ [334]. But this is equivalent to a condition that f is a $*$-homomorphism, since $f(x^* x) = f(x)^* f(x) = |f(x)|^2$, hence $f(|x|) = |f(x)|$, which follow from (137). From the equality of multiplicative unit $\mathbb{I}$ with an order unit, and coincidence of definitions of normality and order continuity, it follows that a function $f : X_1 \to X_2$ between two commutative unital W^*-algebras X_1 and X_2 is a normal (resp., unital) $*$-homomorphism iff it is order continuous (resp., unit preserving). Finally, the surjective isometries of commutative W^*-algebras coincide with their $*$-isomorphisms (and are normal), while the surjective isometries of Banach lattices coincide with their isometric Riesz isomorphisms (and are order continuous). □

This establishes direct analogy between the properties of the family $L_{1/\gamma}(\mathcal{A})$ spaces over mcb-algebras $\mathcal{A}$ and the properties of the family of $L_{1/\gamma}(\mathcal{N})$ spaces over W^*-algebras $\mathcal{N}$. In what follows, we will see that those two settings coincide in the case when W^*-algebra is commutative.

2.3 Statistical and quantum models

The Borel–Steinhaus–Kolmogorov [95, 96, 689, 394] approach to mathematical foundations of probability theory and statistics is developed within the frames of measure theory on abstract sets. For a given choice of a 'background' premeasurable space $(\mathcal{X}, \mho(\mathcal{X}))$, a statistical model is defined [83] as a subset

$$\mathcal{M}(\mathcal{X}, \mho(\mathcal{X})) \subseteq \mathrm{Meas}^+(\mathcal{X}, \mho(\mathcal{X})) \tag{142}$$

of the set $\mathrm{Meas}^+(\mathcal{X}, \mho(\mathcal{X}))$ of all countably additive measures on $(\mathcal{X}, \mho(\mathcal{X}))$. In order to deal with the elements of $\mathcal{M}(\mathcal{X}, \mho(\mathcal{X}))$, it is assumed that there exists a countably additive measure $\tilde{\mu}$ on $(\mathcal{X}, \mho(\mathcal{X}))$ such that $\mathcal{M}(\mathcal{X}, \mho(\mathcal{X}))$ can be represented as a set of Radon–Nikodým quotients of elements of $\mathrm{Meas}^+(\mathcal{X}, \mho(\mathcal{X}))$ with respect to $\tilde{\mu}$,

$$\mathcal{M}(\mathcal{X}, \mho(\mathcal{X})) \cong \mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \subseteq L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+, \tag{143}$$

where $\cong$ denotes the bijection between sets. This assumption requires absolute continuity of elements of $\mathcal{M}(\mathcal{X}, \mho(\mathcal{X}))$ with respect to $\tilde{\mu}$. Among all statistical models of this form, the probabilistic models are defined as subsets

$$\mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \subseteq L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+_1 := \{p \in L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+ \mid \int \tilde{\mu} p = 1\}. \tag{144}$$

The set $\mathcal{X}$ is called a 'sample space' and is in principle arbitrary, the choice of $\mho(\mathcal{X})$ is provided by some additional principle (e.g., one chooses an algebra $\mho_{\mathrm{Borel}}(\mathcal{X})$ of Borel subsets of $\mathcal{X}$ if $\mathcal{X}$ is a topological space), while the choice of element $\tilde{\mu}$ of $\mathrm{Meas}^+(\mathcal{X}, \mho(\mathcal{X}))$ is also arbitrary (e.g., one chooses some element in the set $\mathrm{Meas}^+_\star(\mathcal{X}, \mho_{\mathrm{Borel}}(\mathcal{X}))$ of normal Radon measures on $\mho_{\mathrm{Borel}}(\mathcal{X})$). However, as follows from the discussion in Section 2.1, there exist many different choices of $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ that lead to mutually isometrically isomorphic $L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ spaces. This suggests that one should be able to reformulate the above setting in a more concise form. By Segal's theorem [658], validity of the Radon–Nikodým theorem, which is necessary and sufficient to guarantee that statistical models allow representation of their elements in terms of the Radon–Nikodým quotients, is equivalent to requirement that the Steinhaus–Nikodým [688, 538] isometric isomorphism holds,

$$L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^\star \cong L_\infty(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}), \tag{145}$$

and is equivalent to the condition that the measure space $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ is localisable. But in view of the results discussed in Section 2.1, this means that the notion of statistical model can be made independent of the choice of a 'sample space' $\mathcal{X}$ and a measure space $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$. It depends only on the choice of an mcb-algebra $\mathcal{A}$, which canonically determines associated family of $L_p(\mathcal{A})$ spaces with $p \in [1, \infty]$. In consequence, we define: a ***statistical model*** as a subset $\mathcal{M}(\mathcal{A}) \subseteq L_1(\mathcal{A})^+$; a ***probabilistic model*** as a subset

$$\mathcal{M}(\mathcal{A}) \subseteq L_1(\mathcal{A})^+_1 := \{p \in L_1(\mathcal{A})^+ \mid \|p\| = 1\}. \tag{146}$$

An element $\phi \in L_1(\mathcal{A})^+$ will be called a ***statistical state***, while an element $\phi \in L_1(\mathcal{A})^+_1$ will be called a ***probabilistic state*** (or ***probabilistic expectation***). If some representation of $L_1(\mathcal{A})$ in terms of $L_1(\mathcal{A}, \mu)$ or $L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ is chosen, then the element representing a statistical state (resp., probabilistic state) will be called ***statistical density*** (resp., ***probability density***) with respect to μ or $\tilde{\mu}$ [467].

The postulate of expressibility of elements of $\mathcal{M}(\mathcal{X}, \mho(\mathcal{X}))$ in terms of the Radon–Nikodým quotients can be justified in various ways, and in particular by referring to the notion of *sufficiency*. This idea of Fisher [254] was translated into mathematical terms by Neyman [536], and obtained an abstract measure theoretic formulation due to Halmos and Savage [304] in terms of the Radon–Nikodým quotients of the elements of the model $\mathcal{M}(\mathcal{X},\mho(\mathcal{X}))\subseteq\mathrm{Meas}^+(\mathcal{X},\mho(\mathcal{X}))$ with respect to a single dominating[9] measure $\tilde{\mu}$ on $\mho(\mathcal{X})$. However, this formulation covered only such models for which the dominating measure $\tilde{\mu}$ was finite countably additive. The direct extension to more general case is involved in pathologies [595, 129], which were shown to be solvable only for a class of models—called 'compact' [596, 213], 'weakly dominated' [517], or 'coherent' [318]—for which there exists a dominating measure $\tilde{\mu}$ such that $(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$ is localisable (see [471, 91] for reviews).[10]

Let us compare our proposal with other approaches to the mathematical foundations of probability theory. Carathéodory's approach to integration[11] has led Kappos [377, 378, 379, 380] to develop foundations of probability theory based on countably additive finite measures on ccb-algebras.[12] Our approach can be considered as a restriction of Kappos' approach on the level of admitted class of boolean algebras, but as an extension on the level of admitted measures.[13] This change allows us to define the canonical range of $L_p(\mathcal{A})$ spaces that are independent of the choice of 'reference' measure, and which precisely correspond to the case when the Steinhaus–Nikodým duality and the Radon–Nikodým theorem (for not necessarily finite measures) hold. On the other hand, the foundations of probability theory based on Riesz lattices emerged from Daniell's approach to integration[14], and were developed by Le Cam [435, 436, 437] (see also [695, 719, 740]) and Whittle [763, 764]. The main difference between Whittle's and Le Cam's approach is that the former starts from Daniell's integral ω over a Daniell lattice over a given 'sample space' $\mathcal{X}$, while the latter starts from an abstract Banach lattice and recovers a 'sample space' by means of the Bohnenblust–Kakutani–Nakano representation theorem.[15] More precisely, the approach of Le Cam is based on consideration of an abstract L_1 space X as a fundamental entity of the theory. The probabilistic model (an 'experiment' in Le Cam's terminology) is defined as a subset of $\mathcal{M}(X)\subseteq\{x\in X^+\mid\|x\|=1\}$. However, in order to guarantee the well-behavedness of inferences on such model, it is necessary to restrict models under consideration to the class of 'coherent' models, which is equivalent to the assumption that $X^\star$ is a proper abstract L_∞ space, which is equivalent to assuming that the projection bands in $X^\star$ form an mcb-algebra $\mathcal{A}$ such that X is isometrically Riesz isomorphic to $L_1(\mathcal{A})$, see [91, 719, 92]. Thus, Le Cam's approach restricts to ours in all cases when it is fully applicable. Finally, Whittle's approach is based on consideration of normalised Daniell–Stone integrals ('probabilistic expectations') on a given Daniell lattice of functions $f:\mathcal{X}\to\mathbb{R}\cup\{+\infty\}$. In order to equip Whittle's approach with the structure allowing the Steinhaus–Nikodým duality, one needs to pass to Banach lattice setting and impose the existence of strong order unit as well as Dedekind–MacNeille completeness of $L_\infty(X,\omega)$ space, which amounts to recovering Le Cam's setting, but in a representation that is dependent on the choice of a 'sample space' $\mathcal{X}$ and integral ω, or, more generally, Daniell system $(\hat{X},\hat{\omega})$. Our approach removes this representation dependence.

In the BSK and Kappos' approaches ***probability*** is defined as a normalised countably additive measure, and is considered as an elementary notion. In Whittle's and Le Cam's approaches ***probability*** is considered as a derived notion, defined by $p(\mathcal{Y}):=\omega(\chi_{\mathcal{Y}})$, where $\chi_{\mathcal{Y}}(x)$ is a characteristic function

[9]A measure $\tilde{\mu}\in\mathrm{Meas}(\mathcal{X},\mho(\mathcal{X}))$ is called to be ***dominating*** with respect to a given set $X\subseteq\mathrm{Meas}(\mathcal{X},\mho(\mathcal{X}))$ iff $\widetilde{\nu}\ll\tilde{\mu}$ $\forall\widetilde{\nu}\in X$. The definition of this term for measures on boolean algebras is analogous.

[10]This has in turn rendered the notion of *pairwise sufficiency* [304] equally fundamental as sufficiency, because they coincide precisely for coherent models [682, 772].

[11]Von Neumann also gave lectures on this topic, but they were not published, cf. [663, 638].

[12]Such approach was suggested also by Weil [759].

[13]On the noncommutative level, Kappos' approach corresponds to the theory of states and weights on Rickart C^*-algebras (which in principle allows to obtain *some* interesting results, because the elements of Rickart C^*-algebras admit unique polar decompositions [30, 283]), while our approach corresponds to the theory of states and weights on W^*-algebras.

[14]Quite ironically, Daniell has not used his approach to integration to make any foundational claims in his own work on probability theory [199]. Wiener [767, 768] used Daniell's approach to define Wiener's integral (which obtained its measure theoretic implementation much later), but he also had not provided any suggestion that Daniell's approach should be used in foundations of probability theory. First functional analytic foundational approach based on expectation was proposed by Segal [660, 662], but it was not developed to full theory. Expectations were considered as more fundamental than probabilities also by de Finetti [206]. See [691, 741, 671] for some additional historical comments.

[15]For the notions of Daniell integral, Daniell lattice, and Daniell system, see e.g. [404].

of $\mathcal{Y} \subseteq \mathcal{X}$, while ω is a probabilistic expectation (Le Cam's approach requires to provide first the BKN representation). In our approach probability is also a derived notion, and is constructed either by passing through the BKN representation, and then using Whittle's or the BSK definition, or by choosing any $\mu \in \mathcal{W}(\mathcal{A})$ on a mcb-algebra $\mathcal{A}$ and using Kappos' definition.

The setting of Banach lattices $L_p(\mathcal{A})$ canonically associated with mcb-algebras $\mathcal{A}$ provides this way a foundational approach that is free from the notions of 'probability' and 'sample space', is independent of the choice of any 'reference' measure on $\mathcal{A}$ or a 'reference' integral on $L_\infty(\mathcal{A})$, and is equivalent to the BSK, Kappos', Le Cam's and Whittle's approaches precisely on the range where these approaches are fully applicable. These equivalences follow from equivalence of categories of: (1) proper abstract L_∞ spaces with order continuous unit preserving Riesz homomorphisms, (2) mcb-algebras with order continuous boolean homomorphisms, (3) localisable measurable spaces with categorical duals of complete morphisms.

The equivalence of these categories with the category of commutative W^*-algebras with unital weak-$\star$ $*$-homomorphisms allow us to exploit the above insights in a straightforward way, using the Falcone–Takesaki theory of canonical $L_p(\mathcal{N})$ spaces over *arbitrary* W^*-algebras $\mathcal{N}$. For a given W^*-algebra $\mathcal{N}$, we define a ***quantum model*** (or a ***quantum information model***) as a subset $\mathcal{M}(\mathcal{N}) \subseteq L_1(\mathcal{N})^+ \cong \mathcal{N}_\star^+$. In agreement with terminology of Section 2.2, the elements of $\mathcal{M}(\mathcal{N})$ are called ***states*** (or ***quantum states***, or ***quantum information states***, or ***quantum expectations***). The ***normalised quantum model*** is defined as a subset $\mathcal{M}(\mathcal{N}) \subseteq \mathcal{N}_{\star 1}^+ = \mathcal{S}(\mathcal{N}) \cap \mathcal{N}_\star$, and its elements are called ***normalised states*** (or ***normalised quantum expectations***). If $\mathcal{N}$ admits a faithful normal semi-finite trace τ, and a representation of $\mathcal{N}_\star^+$ in terms of $L_1(\mathcal{N},\tau)^+$ is considered, then $\rho_\phi \in L_1(\mathcal{N},\tau)^+$ representing $\phi \in \mathcal{N}_\star^+$ will be called a ***density operator*** [426, 748]. From Section 2.2 it follows that for commutative W^*-algebras $\mathcal{N}$ the quantum information models $\mathcal{M}(\mathcal{N})$ turn into statistical models $\mathcal{M}(\mathcal{A})$, where $\mathcal{N} \cong L_\infty(\mathcal{A})$ is an isometric isomorphism and a $*$-isomorphism, while $L_\infty(\mathcal{A}) \cong L_1(\mathcal{A})^\star$ is an isometric isomorphism and a Riesz isomorphism. Restriction to normalised states in this case gives $\mathcal{N}_{\star 1}^+ \cong L_1(\mathcal{A})_1^+$.

If $\mathcal{N}$ is any W^*-algebra and $\psi \in \mathcal{W}_0(\mathcal{N})$, then $\mathcal{M}(\mathcal{N})$ can be represented as a subset (see [404])

$$\mathcal{M}(\mathcal{N},\psi) \subseteq L_1(\mathcal{N},\psi)^+ \cong \mathscr{M}^1(\mathcal{N} \rtimes_{\sigma^\psi} \mathbb{R}, \tau_\psi)^+. \tag{147}$$

This provides a noncommutative counterpart of representation of $\mathcal{M}(\mathcal{A})$ in terms of a subset $\mathcal{M}(\mathcal{A},\mu) \subseteq L_1(\mathcal{A},\mu)^+$ for any choice of a measure $\mu \in \mathcal{W}_0(\mathcal{A})$ on an mcb-algebra $\mathcal{A}$. Finally, if $\mathcal{N}$ does not contain any type III factor and if $\mathcal{N}_\star^+$ contains at least one faithful element ω, then $\mathcal{M}(\mathcal{N})$ can be represented in terms of the space $\mathcal{M}(\mathcal{H}_\omega) \subseteq \mathfrak{G}_1(\mathcal{H}_\omega)^+$ where $\mathfrak{G}_1(\mathcal{H}_\omega)^\star \cong \mathfrak{B}(\mathcal{H}_\omega)$. This provides a noncommutative analogue of representation of $\mathcal{M}(\mathcal{A})$ in terms of a subset $\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu}) \subseteq L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+$ over such measure space $(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$ that $(\mathcal{X},\mho(\mathcal{X}),\mho^{\tilde{\mu}}(\mathcal{X}))$ is localisable and $\tilde{\mu}$ is strictly positive. This analogy becomes strict if $\mathcal{N}$ is commutative and $\tilde{\mu}$ is finite: every $\phi \in L_1(\mathcal{A})^+$ defines an element $\mu \in \mathcal{W}(\mathcal{A})$ by means of (64).

2.4 Markovian categories

Following Wald's [754, 756] and Blackwell's [83, 84] works, Chencov [148, 151, 152] and Morse & Sacksteder [509, 643, 644] have introduced the category of statistical models with Banach preduals of Markov maps as morphisms. (Further early works on this topic are [463, 484, 637, 636, 429, 483, 153].) The underlying idea was to consider this category as an underlying structure of mathematical foundations of statistical theory.[16] This expresses the central role that Markov morphisms began to play at that time in statistical theory [435, 534, 122].

Given any set $\mathcal{X}$, the set of all subsets of $\mathcal{X}$ will be denoted $\wp(\mathcal{X})$. If $\mathcal{X}$ is finite, then $\#(\mathcal{X})$ will denote the number of its elements. For any premeasurable space $(\mathcal{X},\mho(\mathcal{X}))$ a ***probability simplex*** is defined as the set

$$L_1(\mathcal{X},\mho(\mathcal{X}))_1^+ := \{\tilde{\mu} \in \mathrm{Meas}^+(\mathcal{X},\mho(\mathcal{X})) \mid \tilde{\mu}(\mathcal{X}) = 1\}. \tag{148}$$

[16] Very similar perspective is implicitly contained in the work of Le Cam [435], and it was turned to an explicit category theoretic formulation by Huber (cf. [331]) around the same time as the Chencov–Morse–Sacksteder approach had appeared.

If $\mathcal{X}$ is a finite set with $\#(\mathcal{X}) =: n \in \mathbb{N}$ elements and $\mho(\mathcal{X}) \cong \wp(\mathcal{X})$ then the space $\mathrm{Meas}(\mathcal{X}, \mho(\mathcal{X}))$ is isomorphic to $\mathbb{R}^n$, and

$$L_1(\mathcal{X}, \wp(\mathcal{X}))^+_1 \cong \{(p_1, \ldots, p_n) \in [0,1]^n \mid \sum_{i=1}^n p_i = 1\}, \tag{149}$$

with $p_i := p(x_i)$ for $p \in L_1(\mathcal{X}, \wp(\mathcal{X}))^+_1$. Such probability simplex $L_1(\mathcal{X}, \mho(\mathcal{X}))^+_1$ will be called ***finite***. If $j \in \{1, 2\}$, $\mathcal{X}_j$ is a finite set, $\mho_j(\mathcal{X}_j) \cong \wp(\mathcal{X}_j)$, and $\#(\mathcal{X}_1) = \#(\mathcal{X}_2) =: n \in \mathbb{N}$, then a ***finite coarse graining*** is defined as such function

$$T_\star : L_1(\mathcal{X}_1, \mho_1(\mathcal{X}_1))^+_1 \ni p \mapsto T_\star(p) \in L_1(\mathcal{X}_2, \mho_2(\mathcal{X}_2))^+_1 \tag{150}$$

that can be represented as $n \times n$ matrix with entries in $\mathbb{R}$ (i.e., an element of $\mathrm{M}_n(\mathbb{R})$), which is a ***stochastic matrix***, defined by $(T_\star(p))_j = \sum_{i=1}^n (T_\star)_{ij} p_i$, $(T_\star)_{ij} \geq 0$ and $\sum_{j=1}^n (T_\star)_{ij} \mathbb{I}_j = \mathbb{I}_i$ for all $i, j \in \{1, \ldots, n\}$ (see e.g. [85]). It is the most general mapping between two finite probability simplexes. The category $\mathbf{ProbMod}_{\mathrm{fin}}$ consists of finite probability simplexes and finite coarse graining maps. If $\mathcal{X}$ is finite and $L_1(\mathcal{X}, \mho(\mathcal{X}))^+_1$ is a finite probability simplex, then the set of all finite coarse grainings from $L_1(\mathcal{X}, \mho(\mathcal{X}))^+_1$ into itself forms a semi-group with respect to composition. When considered together, they form a subcategory $\mathbf{ProbMod}_{\mathrm{fin}}(\mathcal{X})$ of $\mathbf{ProbMod}_{\mathrm{fin}}$ [148, 509] (compare with [432, 282]).

The assumptions of finite dimensionality and normalisation can be dropped. Let $(\mathcal{A}_1, \mu_1)$ and $(\mathcal{A}_2, \mu_2)$ be localisable measure algebras. A ***Markov map*** (named due to historical origins of this idea in [482]) is defined as a positive linear function

$$T : L_\infty(\mathcal{A}_1) \to L_\infty(\mathcal{A}_2) \text{ such that } T(\mathbb{I}) = \mathbb{I}. \tag{151}$$

A ***coarse graining*** is defined as a positive linear function

$$T_\star : L_1(\mathcal{A}_2, \mu_2) \to L_1(\mathcal{A}_1, \mu_1) \text{ such that } \|f\| = \|T_\star(f)\| \; \forall f \in L_1(\mathcal{A}_2, \mu_2)^+, \tag{152}$$

or, equivalently,

$$\int \mu_1 T_\star(f) = \int \mu_2 f \;\; \forall f \in L_1(\mathcal{A}_2, \mu_2). \tag{153}$$

Every positive linear function on a Banach lattice is norm continuous. Markov maps are dual to coarse grainings in terms of the Steinhaus–Nikodým duality $L_1(\mathcal{A}, \mu)^\star \cong L_\infty(\mathcal{A})$: for every $T_\star$ (or, respectively, T) there exists a unique T (or, respectively, $T_\star$) such that

$$\int \mu_2 T(p) f = \int \mu_1 p T_\star(f) \;\; \forall p \in L_\infty(\mathcal{A}_1) \; \forall f \in L_1(\mathcal{A}_2, \mu_2). \tag{154}$$

This allows us to define the category $\mathbf{ProbMod}$ of (measure algebraic representations of) probabilistic models $\mathcal{M}(\mathcal{A}, \mu) \subseteq L_1(\mathcal{A}, \mu)^+_1$ and coarse grainings, where $(\mathcal{A}, \mu)$ varies over all localisable measure algebras. Restriction to probabilistic models constructed over a fixed localisable measure algebra $(\mathcal{A}, \mu)$ defines a subcategory $\mathbf{ProbMod}(\mathcal{A}, \mu)$ of $\mathbf{ProbMod}$. In general commutative case, we define: a ***Markov map*** as a unit preserving positive ($x \geq 0 \Rightarrow T(x) \geq 0$) linear function between MI-spaces; a ***coarse graining*** as a positive linear function $T_\star : X_1 \to X_2$ between abstract L_1 spaces that satisfies $\|T_\star(x)\| = \|x\| \; \forall x \in X_1^+$ [435] (Le Cam calls it a 'transition'). The Banach space duality $L_1(\mathcal{A})^\star \cong L_\infty(\mathcal{A})$ and the Banach space duality between abstract L_1 spaces and proper abstract L_∞ spaces determines a duality between coarse grainings $T_\star : X_1 \to X_2$ and Markov maps $T : X_2^\star \to X_1^\star$ by means of

$$[\![T_\star(f), \phi]\!]_{X \times X^\star} = [\![f, T(\phi)]\!]_{X \times X^\star} \;\; \forall f \in X_1 \;\; \forall \phi \in X_2^\star. \tag{155}$$

The category of abstract L_1 spaces and coarse grainings is equivalent to a category of $L_1(\mathcal{A})$ spaces over mcb-algebras $\mathcal{A}$ and coarse grainings. As a result, the category $\mathbf{StatMod}$ of statistical models $\mathcal{M}(\mathcal{A}) \subseteq L_1(\mathcal{A})^+$ over all mcb-algebras $\mathcal{A}$ and coarse grainings is equivalent to a category $\mathbf{L_1^+cg}$ of *arbitrary* subsets of positive cones X^+ of abstract L_1 spaces X and coarse grainings. The latter embeds into the category $\mathbf{L}_1^+\mathbf{pos}$ of arbitrary subsets of positive cones X^+ of abstract L_1 spaces X and positive linear functions between them. Category $\mathbf{StatMod}(\mathcal{A})$ is defined as a subcategory of $\mathbf{StatMod}$ obtained by fixing the choice of a mcb-algebra $\mathcal{A}$.

Let $\mathcal{C},\mathcal{C}_1,\mathcal{C}_2$ be C^*-algebras and let $\mathrm{M}_n(\mathcal{C})$ denote the algebra $n\times n$ matrices with entries in $\mathcal{C}$. A function $T:\mathcal{C}_1\to\mathcal{C}_2$ is called: ***positive*** iff $T(\mathcal{C}_1^+)\subseteq\mathcal{C}_2^+$ (for $\mathcal{C}_1=\mathcal{C}_2=:\mathcal{C}$ this condition is sometimes strengthened to $T(\mathcal{C}^+)=\mathcal{C}^+$); n***-positive*** iff

$$T\otimes\mathrm{id}_{\mathrm{M}_n(\mathbb{C})}:\mathcal{C}_1\otimes\mathrm{M}_n(\mathbb{C})=\mathrm{M}_n(\mathcal{C}_1)\ni x\otimes y\mapsto T(x)\otimes y\in\mathrm{M}_n(\mathcal{C}_2)=\mathcal{C}_2\otimes\mathrm{M}_n(\mathbb{C}) \tag{156}$$

is positive for $n\in\mathbb{N}$; ***completely positive*** iff it is n-positive for all $n\in\mathbb{N}$ [690]. A set of all completely positive maps over $\mathcal{C}$ forms a semi-group with respect to a composition. For commutative $\mathcal{C}_1$ and $\mathcal{C}_2$ every positive linear function is also completely positive, and this is still true if just one of them is commutative [529, 694], but this is no longer true when both $\mathcal{C}_1$ and $\mathcal{C}_2$ are noncommutative. Every $*$-homomorphism of C^*-algebras is completely positive. Every conditional expectation $\mathcal{E}:\mathcal{N}_1\to\mathcal{N}_2$ of W^*-algebras $\mathcal{N}_2\subseteq\mathcal{N}_1$ is completely positive [717, 525]. If $\mathcal{N}_1$ and $\mathcal{N}_2$ are W^*-algebras then a function $f:(\mathcal{N}_2)_\star\to(\mathcal{N}_1)_\star$ is called ***positive*** iff $f((\mathcal{N}_2)_\star^+)\subseteq(\mathcal{N}_1)_\star^+$. If $T_\star:(\mathcal{N}_2)_\star\to(\mathcal{N}_1)_\star$ is a positive linear function then the adjoint function $T:\mathcal{N}_1\to\mathcal{N}_2$, defined by

$$[\![\phi_2,T(x)]\!]_{(\mathcal{N}_2)_\star\times\mathcal{N}_2}=[\![T_\star(\phi_2),x]\!]_{(\mathcal{N}_1)_\star\times\mathcal{N}_1}\quad\forall x\in\mathcal{N}_1\ \ \forall\phi_2\in(\mathcal{N}_2)_\star, \tag{157}$$

is positive and normal, where $[\![\cdot,\cdot]\!]_{\mathcal{N}_\star\times\mathcal{N}}$ is a Banach space duality between $\mathcal{N}_\star$ and $\mathcal{N}$. Moreover, every positive normal function $T:\mathcal{N}_1\to\mathcal{N}_2$ is an adjoint of a unique positive linear $T_\star$, and this holds also under restriction to $\mathcal{N}_i^{\mathrm{sa}}$ and $(\mathcal{N}_i)_\star^{\mathrm{sa}}$ for $i\in\{1,2\}$. A ***quantum Markov map*** is defined as a normal unital completely positive linear function $T:\mathcal{N}_1\to\mathcal{N}_2$. A ***quantum coarse graining*** is defined as such $T_\star:(\mathcal{N}_2)_\star\to(\mathcal{N}_1)_\star$ that (157) holds, where T is a quantum Markov map. Every quantum coarse graining is positive. If $\mathcal{N}_1=\mathfrak{B}(\mathcal{H}_1)$ and $\mathcal{N}_2=\mathfrak{B}(\mathcal{H}_2)$ for some Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, then every quantum coarse graining $T_\star:\mathfrak{G}_1(\mathcal{H}_2)\to\mathfrak{G}(\mathcal{H}_1)$ is completely positive. As a result, the category $\mathbf{QMod}^{\mathrm{M}}$ of quantum information models $\mathcal{M}(\mathcal{N})$ over all W^*-algebras $\mathcal{N}$ and their quantum coarse graining (where 'M' stands for 'Markov') is a subcategory of the category $\mathbf{QMod}^+$ of quantum information models $\mathcal{M}(\mathcal{N})$ over all W^*-algebras $\mathcal{N}$ and the positive linear functions between them. The restriction to a fixed W^*-algebra $\mathcal{N}$ defines the subcategories $\mathbf{QMod}^{\mathrm{M}}(\mathcal{N})$ and $\mathbf{QMod}^+(\mathcal{N})$. The set of all quantum Markov maps between W^*-algebras having $\mathcal{N}$ as a codomain will be denoted by $\mathrm{Mark}(\mathcal{N})$, while the set of all $T_\star\in\mathrm{Mor}(\mathbf{QMod}^{\mathrm{M}})$ such that $\mathrm{dom}(T_\star)=\mathcal{M}(\mathcal{N})$ will be denoted by $\mathrm{Mark}_\star(\mathcal{M}(\mathcal{N}))$. Under restriction to commutative W^*-algebra $\mathcal{N}\cong L_\infty(\mathcal{A})$, these sets will be denoted by $\mathrm{Mark}(\mathcal{A})$ and $\mathrm{Mark}_\star(\mathcal{M}(\mathcal{A}))$, respectively.

The categories introduced in this section form a commutative diagram

$$\begin{tikzcd}
 & & \mathbf{L}_1^+\mathbf{cg} \arrow[rr, tail] \arrow[d, shift left] \arrow[dr] & & \mathbf{L}_1^+\mathbf{pos} \arrow[d, tail] \\
\mathbf{ProbMod}_{\mathrm{fin}} \arrow[r, tail] & \mathbf{ProbMod} \arrow[r, tail] & \mathbf{StatMod} \arrow[r, tail] \arrow[u, tail, shift left] & \mathbf{QMod}^{\mathrm{M}} \arrow[r, tail] & \mathbf{QMod}^+ \\
\mathbf{ProbMod}_{\mathrm{fin}}(\mathcal{X}) \arrow[u, tail] & \mathbf{ProbMod}(\mathcal{A},\mu) \arrow[r, tail] \arrow[u, tail] & \mathbf{StatMod}(\mathcal{A}) \arrow[r, tail, "c"] \arrow[u, tail] & \mathbf{QMod}^{\mathrm{M}}(\mathcal{N}) \arrow[r] \arrow[u, tail] & \mathbf{QMod}^+(\mathcal{N}), \arrow[u, tail]
\end{tikzcd} \tag{158}$$

where the faithful functor c exists if $L_\infty(\mathcal{A})$ is a W^*-subalgebra of $\mathcal{N}$, and is full if $L_\infty(\mathcal{A})\cong\mathcal{N}$ (in such case it defines an equivalence of categories).

3 Information distances

Given any set X, a ***distance*** is defined as a map $D:X\times X\to[0,\infty]$ such that $D(x,y)=0\iff x=y$. A distance is called: ***bounded*** iff $\mathrm{ran}(D)=\mathbb{R}^+$; ***symmetric*** iff $D(x,y)=D(y,x)$; ***metrical*** (or ***Fréchet*** [256]) iff it is bounded, symmetric and satisfies ***triangle inequality***

$$D(x,z)\leq D(x,y)+D(y,z)\ \ \forall x,y,z\in X. \tag{159}$$

We will use the symbol d instead of D to denote metrical distances. We define a ***statistical distance*** as a distance on a statistical model $\mathcal{M}(\mathcal{A}) \subseteq L_1(\mathcal{A})^+$. A ***quantum distance*** is defined as a distance on a quantum model $\mathcal{M}(\mathcal{N}) \subseteq \mathcal{N}_\star^+$.[17] We will use the notion ***information distance*** to refer unspecifically to statistical and quantum distances. A ***relative entropy*** is defined as a map $\mathbf{S} : X \times X \to [-\infty, 0]$ such that $-\mathbf{S}$ is an information distance. This closely follows Wiener's idea that «amount of information is the negative of the quantity defined as entropy» [769]. First works in which information distances were discussed are [478, 74, 73, 75, 76, 601, 351, 284, 489, 303, 418]. See [312, 603] for a review of various quantifications of (dis)similarity in the statistical context, and [3, 1, 2, 182, 239, 186, 160] for review and characterisations of various information distances and other quantifications of (dis)similarity in the context of information theory. First known examples of metrical quantum distances on $\mathcal{N}_\star^+$ were: the Jauch–Misra–Gibson–Kronfli distance [343, 412, 300],

$$d_{L_1(\mathcal{N})}(\phi, \psi) := \frac{1}{2}\|\phi - \psi\|_{\mathcal{N}_\star}; \tag{160}$$

the Bures distance [127],

$$d_{\mathrm{Bures}}(\phi, \psi) := \inf_{(\mathcal{H},\pi)} \{\|\zeta_\pi(\phi) - \zeta_\pi(\psi)\|_{\mathcal{H}}\}, \tag{161}$$

where $\zeta_\pi(\omega) \in \mathcal{H}$ is defined by $\omega(x) = \langle\zeta_\pi(\omega), \pi(x)\zeta_\pi(\omega)\rangle_{\mathcal{H}}$ $\forall x \in \mathcal{N}$ for some representation $(\mathcal{H}, \pi)$ of $\mathcal{N}$ and the infimum varies over all possible representations; the Araki metrical distance [32, 35]; the Gudder distance [293]. For further examples of metrical quantum distances see [212, 301]. First nonmetrical quantum distance was introduced by Umegaki [735, 736] for W^*-algebras admittion faithful normal semi-finite traces. Its generalisation to arbitrary W^*-algebras was carried in [39, 40] (see also [726, 598] for equivalent constructions), and reads

$$D_1(\omega, \phi) = \begin{cases} (\phi - \omega)(\mathbb{I}) + \langle\xi_\pi(\omega), \log(\Delta_{\omega,\phi})\xi_\pi(\omega)\rangle_{\mathcal{H}} & : \ \omega \ll \phi \\ +\infty & : \ \text{otherwise.} \end{cases} \tag{162}$$

Quantum distances will serve us as a principal tool for quantification of information content of quantum states and quantum models. We will be interested in the families of quantum distances that usually will be nonmetrical and, moreover, nonsymmetric. In Section 3.1 we discuss the family of quantum distances that are nonincreasing under coarse grainings, called $\mathfrak{f}$-*distances*. This property can be understood as a requirement of compatibility of the quantum distance on a quantum model with the structure of the category $\mathbf{QMod}^{\mathrm{M}}$, expressing the requirement that "the coarse graining of information models should always be indicated by nonincreasing of the quantification of relative information content of information states". As a result, various information geometric structures on quantum models arising from this family of distances by means of Eguchi equations (416)-(418) are also compatible with the structure of $\mathbf{QMod}^{\mathrm{M}}$ (see Section 5.2). In Section 3.2 we consider a class $\widetilde{D}_\Psi$ of two-point functionals[18] on vector spaces, known as Brègman functionals, which have another remarkable property. They provide a generalisation of pythagorean theorem beyond the framework of Euclidean and Hilbert spaces, allowing (under some conditions) for an additive decomposition under *nonlinear* projection onto convex subset, where the projection is defined as a unique minimiser of this functional. While some of the Brègman functionals are also distances, which allows to consider them as information distances in the case of $L_1(\mathcal{A})$ or $L_1(\mathcal{N})$ vector spaces, this perspective is of limited applicability, especially when infinite dimensional (nonparametric) quantum models are considered.

[17] The functions that we call '(quantum/statistical) distances' are often called '(quantum) information divergences'. However, this causes very unfortunate collision of terms with well established notion of *divergence* used in differential calculus and differential geometry. Moreover, the term 'divergence' was introduced and used by Kullback and Leibler [418] in the context of relative entropy, but in order to refer to an example of what we call a *symmetric* distance. Rényi [614] proposed to use the term 'information gain'. Chencov [152] proposed to use the term 'deviation', but it seems for us to sound too awkward comparing with a generality and omnipresence of its designate. Eguchi [243] (following Pfanzagl [582, 583]) used the term 'contrast functional'. We think that it is more reasonable to extend the range of the meaning of term 'distance', which is also in agreement with some of the prominent works in the field of information theory, e.g. [144, 184, 557].

[18] As opposed to terminology of [404], in this paper we will use the term ***functional*** to refer to any $\mathbb{K}$-valued, but not necessarily $\mathbb{K}$-linear, function on the vector space over $\mathbb{K}$ or on a cartesian product of such vector spaces.

For this reason, in Section 3.3 we construct a family of dualistic Brègman distances. The key elements of this construction are the Young–Fenchel inequality, dual pairs of coordinate systems and a suitable generalisation of the *bijective* Legendre transform to the infinite dimensional case. This approach includes the large part of theory of Brègman (and Alber) functionals as a special case.

The families of $\mathfrak{f}$-distances and Brègman distances are definitely two most important classes of information distances (cf. e.g. [182, 185, 186]). This leads to ask about the class of quantum information distances that belong to both families. Amari has recently shown [19] that for the finite dimensional statistical models $L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+$ this intersection is characterised by the Liese–Vajda family of γ-distances. Following this result, in Section 3.4 we use the Falcone–Takesaki theory of noncommutative integration to construct the *canonical* noncommutative generalisation of the Liese–Vajda family, and show that the resulting family of quantum γ-distances belongs to an intersection of quantum $\mathfrak{f}$-distances $D_{\mathfrak{f}}$ and quantum Brègman distances D_Ψ. We conjecture that our family of quantum γ-distances on $L_1(\mathcal{N})^+\cong\mathcal{N}^+_\star$ is characterised by this property. Similarly to characterisation of quantum $\mathfrak{f}$-distances by the monotonicity under coarse grainings, and characterisation of Brègman distances by the generalised pythagorean equation, the proof of this conjecture remains an open problem.

Let D be an information distance on an information model $\mathcal{M}$. Given $\mathcal{Q}_1,\mathcal{Q}_2\subseteq\mathcal{M}$ and $\psi\in\mathcal{Q}_1$, we define a ***D-projection*** from $\mathcal{Q}_1$ to $\mathcal{Q}_2$ as a map

$$\mathfrak{P}^D_{\mathcal{Q}_2|\mathcal{Q}_1}:\psi\mapsto\arg\inf_{\phi\in\mathcal{Q}_2}\{D(\phi,\psi)\}, \tag{163}$$

whenever the right hand side is a singleton set. We will denote $\mathfrak{P}^D_{\mathcal{Q}_2}:=\mathfrak{P}^D_{\mathcal{Q}_2|\mathcal{M}}$. From definition of D it follows that $\mathfrak{P}^D_{\mathcal{Q}|\mathcal{Q}}(\psi)=\psi\;\forall\psi\in\mathcal{Q}\;\forall\mathcal{Q}\subseteq\mathcal{M}$, hence $\mathfrak{P}^D_{\mathcal{Q}|\mathcal{Q}}$ is an idempotent operation on an arbitrary information submodel $\mathcal{Q}$. A family of D-projections $\{\mathfrak{P}^D_{\mathcal{Q}_i|\mathcal{Q}_j}\mid i\in I,\;j\in J\}$, where I and J are arbitrary sets, and $\mathcal{Q}_i,\mathcal{Q}_j\subseteq\mathcal{M}\;\forall i\in I\;\forall j\in J$, will be called ***zone consistent*** iff $\mathfrak{P}^D_{\mathcal{Q}_i|\mathcal{Q}_j}=\mathfrak{P}^D_{\mathcal{Q}_i|\mathcal{Q}_k}\circ\mathfrak{P}^D_{\mathcal{Q}_k|\mathcal{Q}_j}\;\forall k\in I\cap J$. A category consisting of objects given by quantum models as objects and zone consistent D-projections as arrows will be denoted $\mathbf{QMod}^D$. A restriction of objects to subsets of $\mathcal{N}^+_\star$ for a given W^*-algebra $\mathcal{N}$ defines a category $\mathbf{QMod}^D(\mathcal{N})$. Note that $\mathbf{QMod}^D$ (respectively, $\mathbf{QMod}^D(\mathcal{N})$) is not a subcategory od $\mathbf{QMod}^+$ (respectively, $\mathbf{QMod}^+(\mathcal{N})$), because zone consistent D-projection $\mathfrak{P}^D_{\mathcal{Q}_2|\mathcal{Q}_1}$ may possess no extension to the full positive cone of $L_1(\mathcal{Q}_1)$.[19]

3.1 $\mathfrak{f}$-distances

A function $f:X\to[-\infty,+\infty]$ on a set X is called: ***proper*** iff it never takes the value $-\infty$ and its ***effective domain*** $\mathrm{efd}(f):=\{x\in X\mid f(x)\neq+\infty\}$ is nonempty. A proper function $f:X\to[-\infty,+\infty]$ is called ***convex*** iff [352, 353]

$$f(\lambda x+(1-\lambda)y)\leq\lambda f(x)+(1-\lambda)f(y)\;\;\forall x,y\in\mathrm{efd}(f)\;\forall\lambda\in\;[0,1]. \tag{164}$$

If this inequality is strict for all $\lambda\in]0,1[$ and all $x,y\in\mathrm{efd}(f)$ with $x\neq y$, then f is called ***strictly convex***. A function $f:X\to[-\infty,+\infty]$ is called: ***concave*** iff $-f$ is convex; ***strictly concave*** iff $-f$ is strictly convex. A proper function $f:X\to[-\infty,+\infty]$ is convex iff its ***epigraph***

$$\mathrm{epi}(f):=\{(x,t)\in X\times\mathbb{R}\mid f(x)\leq t\} \tag{165}$$

is convex. If f_1 is strictly convex and f_2 is convex, then f_1+f_2 is strictly convex whenever it is proper. A proper function $f:X\times X\to[-\infty,+\infty]$ is called ***jointly convex*** iff

$$f(\lambda x_1+(1-\lambda)x_2,\lambda x_3+(1-\lambda)x_4)\leq\lambda f(x_1,x_3)+(1-\lambda)f(x_2,x_4)\;\;\forall\lambda\in[0,1]\;\;\forall x_1,x_2,x_3,x_4\in\mathrm{efd}(f). \tag{166}$$

[19] For any quantum model $\mathcal{M}(\mathcal{N})$ we define $L_1(\mathcal{M}(\mathcal{N}))$ as a smallest space $L_1(\mathcal{C})$, by means of a partial order given by isometric embeddings, for some W^*-algebra $\mathcal{C}$ that contains the linear span of $\mathcal{M}(\mathcal{N})$. Note that the existence of infimum in such poset is not guaranteed a priori, and requires to be proven.

If X is a topological space, then a function $f : X \to [-\infty, +\infty]$ is called: ***closed*** iff its epigraph is a closed subset of $X \times \mathbb{R}$; ***lower semi-continuous at*** $x_0 \in X$ iff $\liminf_{x \to x_0} f(x) \geq f(x_0)$; ***lower semi-continuous*** iff it is lower semi-continuous at all $x \in X$. Every f is closed iff is it lower semi-continuous (with respect to the same topology) [48].

Consider a function $\mathfrak{f} : \mathbb{R}^+ \to \mathbb{R}$ that is convex on $]0, \infty[$, satisfies $\mathfrak{f}(1) = 0$, is strictly convex at 1, and $\mathfrak{f}(0) := \lim_{\lambda \to +0} \mathfrak{f}(\lambda) \in]-\infty, +\infty]$. If $\mathfrak{f}^{\mathbf{c}}(\lambda) := \lambda \mathfrak{f}(\frac{1}{\lambda})$ for $\lambda > 0$ and $\mathfrak{f}^{\mathbf{c}}(0) := \lim_{\lambda \to \infty} \frac{1}{\lambda} \mathfrak{f}(\lambda) \in]-\infty, +\infty]$, then $\mathfrak{f}^{\mathbf{c}} : \mathbb{R}^+ \to \mathbb{R}$ is convex on $]0, \infty[$, and $\mathfrak{f}^{\mathbf{cc}} = \mathfrak{f}$ [179]. The ***Csiszár–Morimoto*** $\mathfrak{f}$***-distance*** [176, 500, 11][20] on a statistical model $\mathcal{M}(\mathcal{A})$ is defined as a function $D_{\mathfrak{f}} : \mathcal{M}(\mathcal{A}) \times \mathcal{M}(\mathcal{A}) \to [0, \infty]$ such that

$$D_{\mathfrak{f}}(\omega, \phi) := \begin{cases} \int \nu_\phi \mathfrak{f}\left(\frac{\mu_\omega}{\nu_\phi}\right) & : \mu_\omega \ll \nu_\phi \\ +\infty & : \text{ otherwise,} \end{cases} \tag{167}$$

where μ_ω and ν_ϕ are measures on $\mathcal{A}$ determined by $\phi, \omega \in \mathcal{M}(\mathcal{A})$, and the conventions

$$\mathfrak{f}(0) \cdot 0 := 0 \text{ (even if } \mathfrak{f}(0) = \infty), \quad 0 \cdot \mathfrak{f}\left(\frac{0}{0}\right) := 0, \quad 0 \cdot \mathfrak{f}\left(\frac{\lambda}{0}\right) := \lambda \lim_{\lambda \to +0} \frac{\mathfrak{f}(\lambda)}{\lambda}, \quad \lambda > 0 \Rightarrow \lambda \mathfrak{f}\left(\frac{\lambda}{0}\right) := \lambda \mathfrak{f}^{\mathbf{c}}(0) \tag{168}$$

are used. Under the above assumptions, $\mathbb{R}^+ \times \mathbb{R}^+ \ni (\lambda_1, \lambda_2) \mapsto \lambda_2 \mathfrak{f}\left(\frac{\lambda_1}{\lambda_2}\right) \in]-\infty, +\infty]$ is jointly convex and lower semi-continuous. From this it follows that $D_{\mathfrak{f}}(\omega, \phi)$ is lower semi-continuous on $L_1(\mathcal{A})^+ \times L_1(\mathcal{A})_0^+$ endowed with norm topologies, and that it is jointly convex if $\mathfrak{f}(0) = 0$ [453]. The early examples of the Csiszár–Morimoto $\mathfrak{f}$-distance include: the ***Pearson–Kagan*** χ^2***-distance*** [558, 368],

$$\mathfrak{f}(\lambda) = (\lambda - 1)^2 \quad \Rightarrow \quad D_{\mathfrak{f}}(\omega, \phi) = \int \frac{(\mu_\omega - \nu_\phi)^2}{\nu_\phi} = \int \nu_\phi \left(\frac{\mu_\omega}{\nu_\phi} - 1\right)^2 =: \chi^2(\omega, \phi); \tag{169}$$

a ***total variation distance*** [646],

$$\mathfrak{f}(\lambda) = |\lambda - 1| \quad \Rightarrow \quad D_{\mathfrak{f}}(\omega, \phi) = \int |\mu_\omega - \nu_\phi|; \tag{170}$$

a *squared* ***Kakutani–Hellinger distance*** [321, 372],

$$\mathfrak{f}(\lambda) = (1 - \sqrt{\lambda})^2 \quad \Rightarrow \quad D_{\mathfrak{f}}(\omega, \phi) = \int (\sqrt{\mu_\omega} - \sqrt{\nu_\phi})^2; \tag{171}$$

the ***Onicescu distance*** ('information energy') [552, 562],

$$\mathfrak{f}(\lambda) = \lambda^2 \quad \lim \quad D_{\mathfrak{f}}(\omega, \phi) = \int \nu_\phi \left(\frac{\mu_\omega}{\nu_\phi}\right)^2; \tag{172}$$

see also [490, 187, 739]. The distance (170) is a unique, up to a positive scalar multiple, Csiszár–Morimoto $\mathfrak{f}$-distance that is also a metrical distance [388]. If there exists a function $\widetilde{\mathfrak{f}} : \mathbb{R}^+ \to \mathbb{R}$ that satisfies the same conditions as $\mathfrak{f}$ above and, moreover,

$$\mathfrak{f}(\lambda) = \widetilde{\mathfrak{f}}(\lambda) - \mathfrak{D}_+ \widetilde{\mathfrak{f}}(1), \tag{173}$$

where $\mathfrak{D}_+$ denotes the right derivative,

$$\mathfrak{D}_+ \widetilde{\mathfrak{f}}(\lambda) := \lim_{t \to +0} \frac{1}{t} \left(\widetilde{\mathfrak{f}}(\lambda + t) - \widetilde{\mathfrak{f}}(\lambda)\right), \tag{174}$$

and if some representation of $\mathcal{A}$ in terms of a localisable measurable space $(\mathcal{X}, \mho(\mathcal{X}), \mho^0(\mathcal{X}))$ is chosen, then the definition (167) can be extended to

$$D_{\mathfrak{f}}(\omega, \phi) := \int \tilde{\upsilon} \frac{\tilde{\nu}_\phi}{\tilde{\upsilon}} \mathfrak{f}\left(\frac{\tilde{\mu}_\omega}{\tilde{\upsilon}} \middle/ \frac{\tilde{\nu}_\phi}{\tilde{\upsilon}}\right), \tag{175}$$

[20] Somewhat similar functionals were considered earlier in [508] under the name "generalised Hellinger integrals", and with different assumptions on $\mathfrak{f}$ (it was considered to be a Young function). See also [724, 725].

where $\tilde{\mu}_\omega, \tilde{\nu}_\phi \in \mathrm{Meas}^+_{\mathrm{fin}}(\mathcal{X}, \mho(\mathcal{X}))$, while $\tilde{\upsilon} \in \mathrm{Meas}^+(\mathcal{X}, \mho(\mathcal{X}))$ is such that $\mho^0(\mathcal{X}) = \mho^{\tilde{\upsilon}}(\mathcal{X})$ (hence, $(\mathcal{X}, \mho(\mathcal{X}), \mho^{\tilde{\upsilon}}(\mathcal{X}))$ is localisable) and $\tilde{\nu}_\phi \ll \tilde{\upsilon}$ and $\tilde{\mu}_\omega \ll \tilde{\upsilon}$, but otherwise arbitrary. This extends the definition of (175) given in [706] to the case of not necessarily finite $\tilde{\upsilon}$ (see also [188, 454] for a formulation taking into account the singular parts of the measures and [294] for a generalisation of $D_{\mathfrak{f}}$ to a quantification of (dis)similarity of more than two arguments). The values of (175) are independent of the choice of $\tilde{\upsilon}$ [453]. It was introduced in [178] for probability models, see also [639, 453] for a discussion of relationship between (175) and (167) in that case.

Every Csiszár–Morimoto $\mathfrak{f}$-distance is a statistical distance (the assumption $\mathfrak{f}(1) = 0$ alone implies that $\omega = \phi \Rightarrow D_{\mathfrak{f}}(\omega, \phi) = 0$), is jointly convex in both variables,

$$D_{\mathfrak{f}}(\lambda\omega_1 + (1-\lambda)\omega_2, \lambda\phi_1 + (1-\lambda)\phi_2) \leq \lambda D_{\mathfrak{f}}(\omega_1, \phi_1) + (1-\lambda) D_{\mathfrak{f}}(\omega_2, \phi_2) \;\; \forall\lambda \in \; [0,1], \tag{176}$$

and satisfies [738]

$$D_{\mathfrak{f}}(\omega, \phi) = D_{\mathfrak{f}^{\mathrm{c}}}(\phi, \omega) \quad \Longleftrightarrow \quad \exists t \in \mathbb{R} \; \forall \lambda \in]0, \infty[\;\; \mathfrak{f}(\lambda) - \mathfrak{f}^{\mathbf{c}}(\lambda) = (\lambda - 1)t, \tag{177}$$

$$D_{\mathfrak{f}}(\omega, \phi) = D_{\mathfrak{f}}(\phi, \omega) \;\; \Leftarrow \;\; \mathfrak{f}(\lambda) = \mathfrak{f}^{\mathbf{c}}(\lambda) \;\; \forall \lambda \in]0, \infty[\,. \tag{178}$$

Moreover, if $\mathfrak{f}$ is twice differentiable at $\lambda = 1$ and $\mathfrak{f}''(1) > 0$, with $\mathfrak{f}''$ denoting the second derivative of f, then [194]

$$\lim_{\omega \to \phi} \frac{D_{\mathfrak{f}}(\omega, \phi)}{\chi^2(\omega, \phi)} = \frac{1}{2}\mathfrak{f}''(1). \tag{179}$$

The properties of $D_{\mathfrak{f}}$ are analysed in detail in [453, 55, 710, 454, 455, 706]. The family (167) was characterised by Csiszár [180], in the case then $\mathcal{M}(\mathcal{A})$ is a finite probability simplex $L_1(\mathcal{X}, \mho(\mathcal{X}))^+_1$, as a unique function $D : L_1(\mathcal{X}, \mho(\mathcal{X}))^+_1 \times L_1(\mathcal{X}, \mho(\mathcal{X}))^+_1 \to \mathbb{R}$ such that

1) $D(\omega, \phi) \geq D(T_\star(\omega), T_\star(\phi))$,

2) $D(\omega, \phi) = D(T_\star(\omega), T_\star(\phi)) \iff \frac{\mu_\omega}{\nu_\phi} = \frac{\mu_{T_\star(\omega)}}{\nu_{T_\star(\phi)}}$,

3) $D(\lambda\omega_1 + (1-\lambda)\omega_2, \lambda\phi_1 + (1-\lambda)\phi_2) \leq \lambda D(\phi_1, \omega_1) + (1-\lambda) D_{\mathfrak{f}}(\phi_2, \omega_2) \; \forall \lambda \in]0,1[$,

for all coarse grainings $T_\star : L_1(\mathcal{X}, \mho(\mathcal{X}))^+_1 \to L_1(\mathcal{X}, \mho(\mathcal{X}))^+_1$ satisfying $T_\star(\phi) = \sum_{x \in \mathcal{Y}_i} \phi(x)$, where $\mathcal{Y}_i \subseteq \mathcal{X}$ satisfy $\mathcal{Y}_i \cap \mathcal{Y}_j = \varnothing$ for $i \neq j$, $\bigcup_i \mathcal{Y}_i = \mathcal{X}$ and $i \in \{1, \ldots, m\}$ with $m \in \mathbb{N}$ arbitrary for each $T_\star$. Another characterisation of the Csiszár–Morimoto $\mathfrak{f}$-distances, involving a more general class of information models and coarse grainings, but also essentially dependent on the conditions 1) and 2) above, was given in [182].

In the most general representation independent formulation, the monotonicity inequality of $D_{\mathfrak{f}}$ (called also 'data processing theorem' [178, 172]) reads

$$D_{\mathfrak{f}}(\omega, \phi) \geq D_{\mathfrak{f}}(\omega \circ T, \phi \circ T) \;\; \forall \omega, \phi \in \mathcal{M}(\mathcal{A}) \;\; \forall T \in \mathrm{Mark}(L_\infty(\mathcal{A})), \tag{180}$$

or, equivalently,

$$D_{\mathfrak{f}}(\omega, \phi) \geq D_{\mathfrak{f}}(T_\star(\omega), T_\star(\phi)) \;\; \forall \omega, \phi \in \mathcal{M}(\mathcal{A}) \;\; \forall T_\star \in \mathrm{Mark}_\star(\mathcal{M}(\mathcal{A})). \tag{181}$$

For a detailed discussion of the case when the elements of $\mathcal{M}(\mathcal{A})$ are dominated by a finite countably additive measure (including the measure theoretic characterisation of equality in (181)), see [176, 178, 518, 453, 454].

A function $\mathfrak{f} : \mathbb{R}^+ \to \mathbb{R}$ is called: ***operator monotone increasing*** [470] iff

$$0 \leq x \leq y \quad \Rightarrow \quad \mathfrak{f}(x) \leq \mathfrak{f}(y) \;\; \forall x, y \in \mathfrak{B}(\mathcal{H}); \tag{182}$$

operator monotone decreasing iff $(-\mathfrak{f})$ is operator monotone increasing; ***operator convex*** [409] iff

$$\mathfrak{f}(\lambda x + (1-\lambda)y) \leq \lambda\mathfrak{f}(x) + (1-\lambda)\mathfrak{f}(y) \;\; \forall x, y \in \mathfrak{B}(\mathcal{H})^+ \;\; \forall \lambda \in \; [0,1]; \tag{183}$$

operator concave iff $(-\mathfrak{f})$ is operator convex [28]. Every operator monotone increasing function is operator concave, while an operator concave function is operator monotone increasing if it is bounded from below [307]. Every operator convex function on $\mathbb{R}^+$ is continuous on $]0,+\infty[$. See [66, 202, 234, 28, 72] for a discussion of operator convex and operator monotone increasing functions and [72, 326, 327] for their integral representations. If $\mathcal{C}_1$ and $\mathcal{C}_2$ are C^*-algebras, and $T:\mathcal{C}_1\to\mathcal{C}_2$ is a unital positive linear function, then the ***Schwarz–Kadison inequality*** $T(x^*x)\geq T(x^*)T(x)$ holds for all normal $x\in\mathcal{C}_1$ [367]. A linear function $T:\mathcal{C}_1\to\mathcal{C}_2$ will be called ***Schwarz***[21] iff

$$T(x^*x)\geq T(x^*)T(x)\;\;\forall x\in\mathcal{C}_1. \tag{184}$$

Every 2-positive linear function between C^*-algebras is Schwarz [157]. Hence, in particular, all completely positive linear maps, and all quantum Markov maps, are Schwarz.

If $\mathfrak{f}:\mathbb{R}^+\to\mathbb{R}$ is operator convex (hence, continuous on $]0,\infty[$) with $\mathfrak{f}(0)\leq 0$ and $\mathfrak{f}(1)=0$, and if $(\mathcal{H},\pi,J,\mathcal{H}^\natural)$ is standard representation of a W^*-algebra $\mathcal{N}$, then the ***Kosaki–Petz*** $\mathfrak{f}$***-distance*** [397, 565] is defined as a function $D_\mathfrak{f}:\mathcal{N}^+_\star\times\mathcal{N}^+_\star\to[0,\infty]$ such that

$$D_\mathfrak{f}(\omega,\phi):=\left\{\begin{array}{ll}\langle\xi_\pi(\phi),\mathfrak{f}(\Delta_{\omega,\phi})\xi_\pi(\phi)\rangle_\mathcal{H} & :\omega\ll\phi\\ +\infty & :\text{ otherwise,}\end{array}\right. \tag{185}$$

where $\xi_\pi(\phi)$ is standard vector representative of ϕ in $\mathcal{H}^\natural$. See [397, 400, 550] for some functional analytic properties of (185). It is a quantum distance. If $\mathcal{N}$ is a type I W^*-algebra, then (185) takes a form [567]

$$D_\mathfrak{f}(\omega,\phi):=\left\{\begin{array}{ll}\left\langle\mathbb{I},\mathfrak{f}(\mathfrak{L}_{\rho_\omega}\mathfrak{R}^{-1}_{\rho_\phi})\mathfrak{R}_{\rho_\phi}\mathbb{I}\right\rangle_{\mathfrak{G}_2(\mathcal{H})}=\left\langle\rho_\phi^{1/2},\mathfrak{f}(\mathfrak{L}_{\rho_\omega}\mathfrak{R}^{-1}_{\rho_\phi})\rho_\phi^{1/2}\right\rangle_{\mathfrak{G}_2(\mathcal{H})} & :\omega\ll\phi\\ +\infty & :\text{ otherwise.}\end{array}\right. \tag{186}$$

As an example of (186), the operator convex function $\mathfrak{f}(\lambda)=(\lambda-1)^2$ gives rise to a quantum analogue of the Pearson–Kagan distance (169), considered e.g. in [446],

$$D_\mathfrak{f}(\omega,\phi)=\mathrm{tr}(\rho_\phi-2\rho_\omega+\rho_\omega^2\rho_\phi^{-1})=\mathrm{tr}\left((\rho_\phi-\rho_\omega)\rho_\phi^{-1}(\rho_\phi-\rho_\omega)\right). \tag{187}$$

By Petz's theorem [565], if $\mathfrak{f}$ is bounded from above (hence, operator monotone decreasing), then $D_\mathfrak{f}$ given by (185) satisfies

$$D_\mathfrak{f}(\omega,\phi)\geq D_\mathfrak{f}(T_\star(\omega),T_\star(\phi))\;\;\forall\omega,\phi\in\mathcal{N}^+_\star \tag{188}$$

for any unital 2-positive function T such that $\mathrm{dom}(T_\star)=\mathcal{N}^+_\star$ (hence, in particular, for every quantum coarse graining $T_\star\in\mathrm{Mark}_\star(\mathcal{N}^+_\star)$), and the equality is attained iff $T_\star$ is an isomorphism (see [327] for a proof when $\mathcal{N}$ is of type I). In the case $\mathfrak{f}(\lambda)=\lambda\log\lambda$ this result is known as Uhlmann's monotonicity theorem [726]. If T is a $*$-homomorphism, then (188) holds without assuming that $\mathfrak{f}$ is bounded from above (which corresponds to weakening operator monotonicity to operator convexity) and without $\mathfrak{f}(0)\leq 0$. In [716] the inequality (188) has been shown to hold for any quantum coarse graining $T_\star$ such that $\mathrm{dom}(T_\star)=\mathcal{N}^+_\star$ and for any Kosaki–Petz $\mathfrak{f}$-distance (without assuming that $\mathfrak{f}$ is bounded from above), which provides a direct generalisation of (181) to

$$D_\mathfrak{f}(\omega,\phi)\geq D_\mathfrak{f}(T_\star(\omega),T_\star(\phi))\;\;\forall\omega,\phi\in\mathcal{M}(\mathcal{N})\;\forall T_\star\in\mathrm{Mark}_\star(\mathcal{M}(\mathcal{N})). \tag{189}$$

In [327] the inequality (188) has been shown to hold for all unital Schwarz functions T such that $\mathrm{dom}(T_\star)=\mathcal{N}^+_\star$ and for any Kosaki–Petz distance. These two results were proven only for type I W^*-algebras, but it seems that they hold for all W^*-algebras. It is not known [577] whether (188) or (189) characterise Kosaki–Petz $\mathfrak{f}$-distances in a way similar to Csiszár's characterisation of the Csiszár–Morimoto $\mathfrak{f}$-distances. Every Kosaki–Petz $\mathfrak{f}$-distance $(\omega,\phi)\mapsto D_\mathfrak{f}(\omega,\phi)$ is lower semi-continuous on $\mathcal{N}^+_\star\times\mathcal{N}^+_{\star0}$ endowed with the product of norm topologies [565]. If $\mathfrak{f}(0)=0$ then $D_\mathfrak{f}(\omega,\phi)$ is jointly

[21] In [327] such functions are called 'Schwarz contractions', their preduals are called 'substochastic maps', while the preduals of unital 'Schwarz contractions' are called 'stochastic maps'.

convex in ω and ϕ (in the sense of (176)) [397, 565]. If $\mathcal{N}$ is commutative and $\mathcal{N} = L_\infty(\mathcal{A})$ for some mcb-algebra $\mathcal{A}$, then every Kosaki–Petz $\mathfrak{f}$-distance on $\mathcal{N}_\star^+$ is a Csiszár–Morimoto $\mathfrak{f}$-distance. Like its commutative counterpart, the Kosaki–Petz $\mathfrak{f}$-distance satisfies also (178).

The early results on existence and characterisation of the projections determined by constrained minimisation of the Csiszár–Morimoto $\mathfrak{f}$-distances [178, 452, 639, 640, 453, 182, 712, 184] were developed as an extension and generalisation of the early results on existence and characterisation of the projections determined by minimisation of the WGKL distance (see Section 3.4) under linear constraints [416, 417, 561, 150, 179, 181] (see [189] for a recent account of this topic). Further development of the general theory of constrained minimisation of the Csiszár–Morimoto $\mathfrak{f}$-distances was strongly influenced by the results of Borwein and Lewis [98, 99, 100, 101, 102, 103, 104, 97, 105, 106, 107, 110, 449, 109, 111, 108], who applied Rockafellar's approach to minimisation problems based on convex duality [627, 628, 631, 630, 632] (see also [12, 86] for more recent discussion), and by the partially related development of the 'maximum entropy on the mean' approach [530, 531, 268, 196, 269, 208, 433, 69, 270, 434, 188]. This has led to establishing three general approaches to *constrained* variational extremisation of functions of type (167), developed by Broniatowski & Keziou [119], Csiszár & Matúš [189, 190, 191, 193, 192], and Léonard [441, 440, 442, 443, 444, 445], respectively. These approaches are not equivalent, and each has its own merits. In particular, [119] covers the problems of minimisation of Kakutani–Hellinger and Burg[22] distances (cf. e.g. [128, 68] for their relevance), [192] establishes the pythagorean theorem, while [443] obtains existence and uniqueness results for a very general family of constraints. In all these approaches the minimisation problem is considered over the spaces of nonnormalised measures. Unfortunately, none of these approaches has been generalised to noncommutative case yet, and also no other approaches dealing with the constrained variational extremisation of functionals (185) are known. We consider Léonard's approach as the most promising candidate for a generalisation to the case of W^*-algebras and their preduals, due to its geometric character (it relies almost exclusively on convex and Banach space geometry and duality), as opposed to measure theoretic and topological character of two alternative approaches.

3.2 Brègman functionals

At least six different *inequivalent* general notions of Brègman functional are present in the literature, each one having its own virtues and flaws (we review them below, to a reasonable extent determined by our later applications). The substantial part of the theory of Brègman functionals is developed for the reflexive Banach spaces. However, this excludes the discussion of the most interesting case of L_1 spaces, which are naturally related with the WGKL and the Umegaki–Araki distances. For that case, there are at least three approaches possible: the general approach based on one-sided Gâteaux derivatives on arbitrary Banach spaces, the measure theoretic approach based on integrals over premeasurable spaces and pointwise composition of gradients over $\mathbb{R}^n$ with $\mathbb{R}^n$-valued measure functions, and the intermediate approach, which can be applied to arbitrary Banach space, but requires its Fréchet differentiability.

The results on Brègman functionals and Brègman functional projections are scattered through the literature, and the main role of this section is to collect them together in a systematic way. Our main references for infinite dimensional convex variational analysis are [52, 584, 633, 162, 132, 108, 793, 116, 115, 421, 497, 113]. Standard expositions of a finite dimensional theory are [629, 632, 246, 161].

A ***dual pair*** is defined [214, 215, 476] as a triple $(X, X^{\mathbf{d}}, [\![\cdot,\cdot]\!]_{X\times X^{\mathbf{d}}})$, where X and $X^{\mathbf{d}}$ are vector spaces over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, equipped with a bilinear duality pairing $[\![\cdot,\cdot]\!]_{X\times X^{\mathbf{d}}} : X \times X^{\mathbf{d}} \to \mathbb{K}$ satisfying [23]

$$[\![x, y]\!]_{X\times X^{\mathbf{d}}} = 0 \;\forall x \in X \;\Rightarrow\; y = 0, \tag{190}$$

$$[\![x, y]\!]_{X\times X^{\mathbf{d}}} = 0 \;\forall y \in X^{\mathbf{d}} \;\Rightarrow\; x = 0. \tag{191}$$

[22] Defined as a Csiszár–Morimoto $\mathfrak{f}$-distance with $\mathfrak{f}(\lambda) = -\log\lambda + \lambda - 1$. It will be identified in Section 3.4 with a distance $D_{\gamma=0}$ in the Liese–Vajda family of γ-distances.

[23] We use here the general setting of dual vector spaces, and do not restrict our considerations to locally convex topological vector spaces, because we have in mind the possible future use of convenient vector spaces [266, 411].

An example of a dual pair is given by a Banach space X, $X^{\mathbf{d}} = X^\star$, and the dual pairing given by the Banach space duality. The ***Fenchel subdifferential*** [251, 498, 121] of a proper $\Psi : X \to [-\infty, +\infty]$ at $x \in \mathrm{efd}(\Psi)$ is a set

$$\partial\Psi(x) := \{\hat{y} \in X^{\mathbf{d}} \mid \Psi(z) - \Psi(x) \geq \mathrm{re}\, [\![z - x, \hat{y}]\!]_{X \times X^{\mathbf{d}}}\, \forall z \in X\}. \tag{192}$$

For $x \in X \setminus \mathrm{efd}(\Psi)$ one defines $\partial\Psi(x) := \varnothing$. The elements of $\partial\Psi(x)$ are called ***Fenchel subgradients*** at x. The ***Fenchel dual*** of Ψ is defined as $\Psi^{\mathbf{L}} : X^{\mathbf{d}} \to [-\infty, +\infty]$ such that [80, 481, 251]

$$\Psi^{\mathbf{L}}(\hat{y}) := \sup_{x \in X}\{\mathrm{re}\, [\![x, \hat{y}]\!]_{X \times X^{\mathbf{d}}} - \Psi(x)\} \;\; \forall \hat{y} \in X^{\mathbf{d}}. \tag{193}$$

Given $X^{\mathbf{dd}}$ such that $(X^{\mathbf{d}}, X^{\mathbf{dd}}, [\![\cdot,\cdot]\!]_{X^{\mathbf{d}} \times X^{\mathbf{dd}}})$ is a dual pair and $X \subseteq X^{\mathbf{dd}}$, one defines $\Psi^{\mathbf{LL}} : X \to [-\infty, +\infty]$ by $\Psi^{\mathbf{LL}} := (\Psi^{\mathbf{L}})^{\mathbf{L}}$. The functions $\Psi^{\mathbf{L}}$ and $\Psi^{\mathbf{LL}}$ are convex for any Ψ, and $\Psi^{\mathbf{LL}}|_X \leq \Psi$. If $\mathrm{efd}(\Psi) \neq \varnothing$, then $\Psi^{\mathbf{L}}(x) > -\infty \; \forall x \in X^{\mathbf{d}}$. If $(X, X^{\mathbf{t}})$ is a dual pair of locally convex topological vector spaces, equipped with weak-$\star$ and weak topologies, respectively, and Ψ is proper, then $\Psi^{\mathbf{L}}$ is weakly-$\star$ lower semi-continuous, $\Psi^{\mathbf{LL}}$ is weakly lower semi-continuous, and ($\Psi^{\mathbf{LL}}|_X = \Psi$ holds iff Ψ is weakly lower semi-continuous and convex) [330, 120]. A lower semi-continuous convex Ψ on X is proper iff $\Psi^{\mathbf{L}}$ on $X^{\mathbf{t}}$ is proper. If X is a Banach space and $\Psi : X \to [-\infty, +\infty]$ is proper, convex, then it is lower semi-continuous in norm topology of X iff it is lower semi-continuous in weak topology on X. In what follows, we will always assume $\mathrm{efd}(\Psi) \neq \varnothing$. If $\Psi : X \to \mathbb{R} \cup \{+\infty\}$ is convex and $\hat{y} \in X^{\mathbf{d}}$, then the ***Young–Fenchel inequality*** [779, 251]

$$\Psi(x) + \Psi^{\mathbf{L}}(\hat{y}) - \mathrm{re}\, [\![x, \hat{y}]\!]_{X \times X^{\mathbf{d}}} \geq 0 \tag{194}$$

holds, with equality iff $\hat{y} \in \partial\Psi(x)$. If $(X, X^{\mathbf{t}})$ is a dual pair of locally convex topological vector spaces, and Ψ is proper, convex, and lower semi-continuous, then equality in (194) holds iff $x \in \partial\Psi^{\mathbf{L}}(\hat{y})$. There exist various criteria for nonemptiness of Fenchel subdifferential. In particular, if $(X, X^{\mathbf{d}})$ are Banach spaces, $[\![\cdot,\cdot]\!]_{X \times X^{\mathbf{d}}}$ is a Banach space duality, and Ψ is proper, convex and lower semi-continuous in norm topology of X, then the Fenchel–Rockafellar theorem [629, 115] states that $\partial\Psi(x) \neq \varnothing \; \forall x \in \mathrm{core}(\mathrm{efd}(\Psi))$, where

$$\mathrm{core}(Y) := \{x \in Y \mid \forall h \in \{z \in X \mid \|z\| = 1\} \; \exists \varepsilon > 0 \;\; \forall t \in [0, \varepsilon] \;\; x + th \in Y, \;\; \varnothing \neq Y \subseteq X\}, \tag{195}$$

and $\mathrm{int}(Y) \subseteq \mathrm{core}(Y)$. If X is a Banach space and $\Psi : X \to [-\infty, +\infty]$ is proper, convex, and lower semi-continuous in norm topology of X, then equivalent are: $x \in \mathrm{int}(\mathrm{efd}(\Psi))$; $x \in \mathrm{core}(\mathrm{efd}(\Psi))$; Ψ is continuous at x. The key role of Fenchel subdifferential $\partial\Psi(x)$ is to characterise minimisers of Ψ at x. In particular, if X is a Banach space, $x \in X$, and $\Psi : X \to [-\infty, +\infty]$ is proper and convex, then

$$x_0 \in \arg\inf_{x \in X}\{\Psi(x)\} \quad \Longleftrightarrow \quad 0 \in \partial\Psi(x_0). \tag{196}$$

If Ψ is also lower semi-continuous with respect to norm topology on X, then the conditions in (196) are equivalent to $\partial\Psi^{\mathbf{L}}(0) \cap X^\star \neq \varnothing$, where $\Psi^{\mathbf{L}}$ is a Fenchel dual with respect to the Banach duality of X and $X^\star$.

If $(X, X^{\mathbf{d}}, [\![\cdot,\cdot]\!]_{X \times X^{\mathbf{d}}})$ is a dual pair and $\Psi : X \to]-\infty, +\infty]$ is proper, then:

$$\mathrm{efd}(\partial\Psi) := \{x \in \mathrm{efd}(\Psi) \mid \partial\Psi(x) \neq \varnothing\}, \tag{197}$$

$$\mathrm{efc}(\partial\Psi) := \{\hat{y} \in X^{\mathbf{d}} \mid \hat{y} \in \partial\Psi(x), \; x \in \mathrm{efd}(\partial\Psi)\}, \tag{198}$$

$$(\partial\Psi)^{-1} : X^{\mathbf{d}} \ni \hat{y} \mapsto (\partial\Psi)^{-1}(\hat{y}) := \{x \in X \mid \hat{y} \in \partial\Psi(x)\} \in \wp(X), \tag{199}$$

and Ψ is called ***cofinite*** iff $\Psi^{\mathbf{L}}$ is everywhere finite. If $(X^{\mathbf{d}}, X^{\mathbf{dd}}, [\![\cdot,\cdot]\!]_{X^{\mathbf{d}} \times X^{\mathbf{dd}}})$ is a dual pair and $X \subseteq X^{\mathbf{dd}}$, then Ψ is called ***adequate*** [745] iff $\mathrm{efd}((\partial\Psi)^{-1}) = \mathrm{efd}(\partial\Psi^{\mathbf{L}}) \neq \varnothing$ and $(\partial\Psi)^{-1}(\hat{y}) = \{*\}$ $\forall \hat{y} \in \mathrm{efd}((\partial\Psi)^{-1})$. If X is a Banach space, $X^{\mathbf{d}} = X^\star$, and Ψ is proper, convex, and lower semi-continuous in norm topology on X, then $\mathrm{int}(\mathrm{efd}(\Psi)) \subseteq \mathrm{efd}(\partial\Psi)$, and $\mathrm{efd}(\partial\Psi)$ is dense in $\mathrm{efd}(\Psi)$.

If X is a vector space over $\mathbb{K}$, $t \in \mathbb{R}$, and $\Psi : X \to [-\infty, +\infty]$ is proper then the ***right Gâteaux derivative*** of Ψ at $x \in X$ in the direction $h \in X$ reads

$$X \times X \ni (x, h) \mapsto \mathfrak{D}^{\mathrm{G}}_{+}\Psi(x; h) := \lim_{t \to^{+}0} \frac{\Psi(x + th) - \Psi(x)}{t} \in [0, +\infty]. \tag{200}$$

If x is fixed and (200) exists for all $h \in X$, then Ψ is called ***Gâteaux differentiable at*** x. If $\Psi : X \to]-\infty, +\infty]$ is convex and Gâteaux differentiable at x, then $\mathfrak{D}^{\mathrm{G}}_{+}\Psi(x; \cdot) \in \partial\Psi(x)$. If $\Psi : X \to]-\infty, +\infty]$ is convex and continuous at x, then $\partial\Psi(x) = \{*\}$ iff Ψ is Gâteaux differentiable at x. If $\Psi : X \to]-\infty, +\infty]$ is convex, lower semi-continuous, and Gâteaux differentiable at x, then it is continuous at x. If X is a Banach space and Ψ is convex and lower semi-continuous, then $\mathfrak{D}^{\mathrm{G}}_{+}\Psi(x; \cdot)$ is convex on X, and continuous on $\mathrm{int}(\mathrm{efd}(\Psi))$, while $\mathfrak{D}^{\mathrm{G}}_{+}\Psi(\cdot, \cdot)$ is finite and upper semi-continuous on $\mathrm{int}(\mathrm{efd}(\Psi)) \times X$. If $x \in \mathrm{efd}(\Psi)$ and $\mathfrak{D}^{\mathrm{G}}_{+}\Psi(x; \cdot)$ is continuous at some $h \in X$, then $\partial\Psi(x) \neq \varnothing$. If X is a Banach space and Ψ is Gâteaux differentiable at $x \in X$, then

$$\mathfrak{D}^{\mathrm{G}}_{+}\Psi(x; y) =: [\![y, \mathfrak{D}^{\mathrm{G}}_{x}\Psi]\!]_{X \times X^\star} \;\; \forall y \in X \tag{201}$$

defines the ***Gâteaux derivative*** [271, 272, 273] $(\mathfrak{D}^{\mathrm{G}}\Psi)(x) \equiv \mathfrak{D}^{\mathrm{G}}_{x}\Psi \in X^\star$ of Ψ at x. A function Ψ is called ***Gâteaux differentiable*** iff $\mathrm{int}(\mathrm{efd}(\Psi)) \neq \varnothing$ and Ψ is Gâteaux differentiable for all $x \in \mathrm{int}(\mathrm{efd}(\Psi))$. If X is a Banach space, $\Psi : X \to [-\infty, +\infty]$ is proper, convex, and lower semi-continuous in norm topology, then: (i) if $\Psi^{\mathbf{L}}$ (with respect to Banach space duality) is strictly convex at all elements of $\mathrm{efd}(\Psi^{\mathbf{L}})$, then Ψ is Gâteaux differentiable; (ii) if $\Psi^{\mathbf{L}}$ is Gâteaux differentiable at all $x \in X^\star$, then Ψ is strictly convex at all elements of $\mathrm{int}(\mathrm{efd}(\Psi))$.

The above definitions of Gâteaux differentiablity can be provided alternatively for $[-\infty, +\infty]$ replaced by some locally convex vector space Y and $\Psi : X \to [-\infty, +\infty]$ replaced by $f : U \to Y$, where U is an open subset of a locally convex vector space X, but we are more concerned here with the way in which the infinite values of Ψ are handled. If X and Y are normed vector spaces, then the ***Fréchet derivative*** [257, 259, 260] of a function $f : X \to Y$ at $x \in X$ in the direction $h \in X$ is defined as a function $\mathfrak{D}^{\mathrm{F}} f : X \times X \to Y$ such that $\mathfrak{D}^{\mathrm{F}} f(x, \cdot)$ is linear and bounded and

$$\lim_{\|h\|_X \to 0} \frac{\left\| f(x + h) - f(x) - \mathfrak{D}^{\mathrm{F}} f(x, h) \right\|_Y}{\|h\|_X} = 0. \tag{202}$$

If such $\mathfrak{D}^{\mathrm{F}} f$ exists, then it is unique and $\mathfrak{D}^{\mathrm{F}} f(x, \cdot) \equiv \mathfrak{D}^{\mathrm{F}}_{x} f$. If X is a vector space over $\mathbb{K}$ and $Y = \mathbb{K}$, then $\mathfrak{D}^{\mathrm{F}}_{x} f \in X^\star$. If $f : X \to \mathbb{R}$ above is replaced by $\Psi : X \to [-\infty, +\infty]$ then the extension of the Fréchet derivative to this case is done under the same conditions as for the Gâteaux derivative. A function $\Psi : X \to [-\infty, +\infty]$ on a normed space X is called: ***Fréchet differentiable at*** x iff $x \in \mathrm{efd}(\Psi)$ and

$$\exists \hat{y} \in X^\star \;\; \forall \epsilon_1 > 0 \;\; \exists \epsilon_2 > 0 \;\; \forall h \in X \;\; \|h\|_X < \epsilon_2 \;\; \Rightarrow \;\; |\Psi(x + h) - \Psi(x) - \hat{y}(h)| \leq \epsilon_1 \|h\|_X; \tag{203}$$

Fréchet differentiable iff $\mathrm{int}(\mathrm{efd}(\Psi)) \neq \varnothing$ and Ψ is Fréchet differentiable at all $x \in \mathrm{int}(\mathrm{efd}(\Psi))$. If Ψ is Fréchet differentiable, then it is also norm continuous and Gâteaux differentiable. For $\dim X < \infty$ these two notions of derivative coincide.

A Banach space X is called: ***strictly convex*** [258, 163] iff

$$\forall x, y \in X \;\; \|x + y\| = \|x\| + \|y\|, \;\; x \neq 0 \neq y \;\; \Rightarrow \;\; \exists \lambda > 0 \;\; y = \lambda x; \tag{204}$$

Gâteaux differentiable [44, 492] iff $\|\cdot\|$ is Gâteaux differentiable at every $x \in X \setminus \{0\}$; ***uniformly convex*** [163] iff

$$\forall \epsilon_1 > 0 \;\; \exists \epsilon_2 > 0 \;\; \forall x, y \in X \;\; \|x\| = \|y\| = 1, \;\; \|x - y\| \geq \epsilon_1 \;\; \Rightarrow \;\; \left\| \frac{x + y}{2} \right\| \leq 1 - \epsilon_2; \tag{205}$$

uniformly Fréchet differentiable [681] iff

$$\forall \epsilon_1 > 0 \;\; \exists \epsilon_2 > 0 \;\; \forall x, y \in X \;\; \|x\| = 1, \;\; \|y\| \leq \epsilon_2 \;\; \Rightarrow \;\; \|x + y\| + \|x - y\| \leq 2 + \epsilon_1 \|y\|; \tag{206}$$

reflexive [302] iff the map $j : X \to X^{\star\star}$, defined by $j(x)(\hat{y}) := \hat{y}(x)\ \forall x \in X\ \forall \hat{y} \in X^\star$ is an isometric isomorphism. If X (resp. $X^\star$) is Gâteaux differentiable, then $X^\star$ (resp. X) is strictly convex [680, 392]. A Banach space X is uniformly convex (rep. uniformly Fréchet differentiable) iff $X^\star$ is uniformly Fréchet differentiable (resp. uniformly convex) [205]. If X is uniformly convex (resp. uniformly Fréchet differentiable), then it is also strictly convex (resp. Gâteaux differentiable). If X is uniformly convex or uniformly Fréchet differentiable, then it is reflexive [496, 369, 563, 681]. If X is Gâteaux differentiable, then there exists a norm-to-weak-$\star$ continous map $\hat{}\, : \{x \in X \mid \|x\|_X = 1\} \to \{x \in X^\star \mid \|x\|_{X^\star} = 1\}$ that is uniquely determined by a condition $[\![x, \hat{x}]\!]_{X\times X^\star} = 1$ [680].

Let X be a Banach space with a norm $\|\cdot\|$. In what follows, we will refer to Banach spaces assuming implicitly that they are over $\mathbb{R}$. For Banach spaces over $\mathbb{C}$ all definitions and results require to replace $[\![\cdot,\cdot]\!]_{X\times X^\star}$ by re $[\![\cdot,\cdot]\!]_{X\times X^\star}$. A function $T : X \to \wp(X^\star)$ is called ***locally bounded*** at $x \in X$ iff [684]

$$\exists\epsilon > 0\quad \sup\left\{\|T(x+\epsilon y)\| \mid y \in X,\ \|y\| \leq 1\right\} < +\infty. \tag{207}$$

If $\Psi : X \to]-\infty, +\infty]$ is proper, then

$$(\partial\Psi)^{-1}(\hat{y}) = \arg\min_{x\in X}\left\{\Psi(x) - [\![x, \hat{y}]\!]_{X\times X^\star}\right\}. \tag{208}$$

A function $\Psi : X \to]-\infty, +\infty]$ is called ***coercive*** iff $\lim_{\|x\|\to+\infty}\Psi(x) = +\infty$. A Banach space X is reflexive iff every proper, convex, coercive function that is lower semi-continuous in norm topology attains its minimum on X. If $\Psi : X \to [-\infty, +\infty]$ is proper, convex, lower semi-continuous and $\Psi^{\mathbf{L}}$ denotes its Fenchel dual with respect to the Banach space duality of X and $X^\star$, then Ψ is called [629, 61, 112, 114]:

- ***essentially Gâteaux differentiable*** iff ($\partial\Psi$ is locally bounded on $\mathrm{efd}(\partial\Psi)$ or $\mathrm{int}(\mathrm{efd}(\Psi)) \neq \varnothing$) and $\partial\Psi(x) = \{*\}\ \forall x \in \mathrm{efd}(\partial\Psi)$;
- ***essentially strictly convex*** iff $(\partial\Psi)^{-1}$ is locally bounded on $\mathrm{efd}((\partial\Psi)^{-1})$ and Ψ is strictly convex on every convex subset of $\mathrm{efd}(\partial\Psi)$;
- ***Legendre*** iff Ψ is essentially Gâteaux differentiable and essentially strictly convex;
- ***essentially Fréchet differentiable*** iff it is essentially Gâteaux differentiable and Fréchet differentiable for all $x \in \mathrm{int}(\mathrm{efd}(\Psi))$;
- ***Fréchet–Legendre*** iff Ψ and $\Psi^{\mathbf{L}}$ are essentially Fréchet differentiable.

If Ψ is continuous and is Gâteaux differentiable at all $x \in X$ then it is essentially Gâteaux differentiable. If Ψ is essentially Gâteaux differentiable then $\mathrm{int}(\mathrm{efd}(\Psi)) \neq \varnothing$ and Ψ is Gâteaux differentiable on $\mathrm{int}(\mathrm{efd}(\Psi))$ [61]. If X is reflexive, then Ψ is essentially Gâteaux differentiable (resp. Legendre, Fréchet–Legendre) iff $\Psi^{\mathbf{L}}$ is essentially strictly convex (resp. Legendre, Fréchet–Legendre). If X is reflexive and Ψ is Legendre, then

$$\mathfrak{D}^{\mathrm{G}}\Psi : \mathrm{int}(\mathrm{efd}(\Psi)) \to \mathrm{int}(\mathrm{efd}(\Psi^{\mathbf{L}})) \tag{209}$$

is bijective, $(\mathfrak{D}^{\mathrm{G}}\Psi)^{-1} = \mathfrak{D}^{\mathrm{G}}(\Psi^{\mathbf{L}})$, and both $\mathfrak{D}^{\mathrm{G}}\Psi$ and $\mathfrak{D}^{\mathrm{G}}(\Psi^{\mathbf{L}})$ are norm-to-weak continuous and locally bounded on their respective domains [61]. If X is an arbitrary Banach space, $\Psi : X \to]-\infty, +\infty]$ is proper and weakly lower semi-continuous, and $\mathrm{efd}((\partial\Psi)^{-1})$ is open, then [745]

1) if Ψ is essentially Gâteaux differentiable, then Ψ is adequate,
2) if X is reflexive, then $\Psi^{\mathbf{L}}$ is essentially Gâteaux differentiable iff Ψ is adequate.

Let X be a Banach space, and let $\Psi : X \to]-\infty, +\infty]$ be proper. Then the ***Brègman functional*** $\widetilde{D}_\Psi : X \times X \to [0, +\infty]$ can be defined in any of the following *inequivalent* ways (see also [130]):

(B_1) for Ψ convex, with $\mathrm{efd}(\Psi) \neq \varnothing$ [389, 390, 391, 131, 133]:

$$\widetilde{D}_\Psi : X \times X \ni (x, y) \mapsto \begin{cases} \Psi(x) - \Psi(y) - \mathfrak{D}^{\mathrm{G}}_{+}\Psi(y; x-y) & : y \in \mathrm{efd}(\Psi) \\ +\infty : \mathrm{otherwise}; \end{cases} \tag{210}$$

(B$_2$) for Ψ convex and lower semi-continuous, with $\mathrm{int}(\mathrm{efd}(\Psi)) \neq \varnothing$ [61]:

$$\widetilde{D}_\Psi : X \times X \ni (x,y) \mapsto \begin{cases} \Psi(x) - \Psi(y) - \mathfrak{D}^{\mathrm{G}}_{+}\Psi(y; x-y) & : y \in \mathrm{int}(\mathrm{efd}(\Psi)) \\ +\infty : \text{otherwise}; \end{cases} \tag{211}$$

(B$_3$) for Ψ convex, lower semi-continuous, and Gâteaux differentiable on $\mathrm{int}(\mathrm{efd}(\Psi)) \neq \varnothing$ [9]:

$$\widetilde{D}_\Psi : X \times X \ni (x,y) \mapsto \begin{cases} \Psi(x) - \Psi(y) - \left[\!\left[x-y, \mathfrak{D}^{\mathrm{G}}_{y}\Psi \right]\!\right]_{X\times X^\star} & : y \in \mathrm{int}(\mathrm{efd}(\Psi)) \\ +\infty : \text{otherwise}; \end{cases} \tag{212}$$

(B$_4$) for Ψ convex, lower semi-continuous, and Fréchet differentiable on $\mathrm{int}(\mathrm{efd}(\Psi)) \neq \varnothing$ [265, 264][24]:

$$\widetilde{D}_\Psi : X \times X \ni (x,y) \mapsto \begin{cases} \Psi(x) - \Psi(y) - \left[\!\left[x-y, \mathfrak{D}^{\mathrm{F}}_{y}\Psi \right]\!\right]_{X\times X^\star} & : y \in \mathrm{int}(\mathrm{efd}(\Psi)) \\ +\infty : \text{otherwise}; \end{cases} \tag{213}$$

(B$_5$) for $\mathrm{MeFun}(\mathcal{X}, \mho(\mathcal{X}); \mathbb{R}^+)$ denoting the space of $\mho(\mathcal{X})$-measurable functions $h : \mathcal{X} \to \mathbb{R}^+$, $\tilde{\mu}$ denoting a countably additive finite measure on $\mho(\mathcal{X})$, $\check{\Psi} : \mathbb{R} \to]-\infty, +\infty]$ proper, strictly convex, and differentiable on $]0, +\infty[$ with $\check{\Psi}(0) = \lim_{t\to+0}\check{\Psi}(t)$ and $t < 0 \Rightarrow \check{\Psi}(t) = +\infty$, X given by a suitable Banach space of some elements of $\mathrm{MeFun}(\mathcal{X}, \mho(\mathcal{X}); \mathbb{R}^+)$, $\Psi(x) := \int \tilde{\mu}(\chi)\check{\Psi}(x(\chi))$ [366, 182, 183, 184, 190]:

$$\widetilde{D}_\Psi : X \times X \to [0, +\infty], \tag{214}$$
$$\widetilde{D}_\Psi : (x,y) \mapsto \int_{\mathcal{X}} \tilde{\mu}(\chi) \left(\check{\Psi}(x(\chi)) - \check{\Psi}(y(\chi)) - \left((\mathrm{grad}\check{\Psi})(y(\chi)) \right) (x(\chi) - y(\chi)) \right); \tag{215}$$

(B$_6$) for $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ as above, $n \in \mathbb{N}$, X given by a suitable Banach space in $\mathrm{MeFun}(\mathcal{X}, \mho(\mathcal{X}); \mathbb{R}^n) \cap \mathcal{L}_0(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}; \mathbb{R}^n)$, $\check{\Psi} : \mathbb{R}^n \to]-\infty, +\infty]$ proper, convex, closed, and $\Psi(x) := \int \tilde{\mu}(\chi)\check{\Psi}(x(\chi))$ [191]:

$$\widetilde{D}_\Psi : X \times X \ni (x,y) \mapsto \int_{\mathcal{X}} \tilde{\mu}(\chi) \check{D}_{\check{\Psi}}(x(\chi), y(\chi)) \in [0, +\infty], \tag{216}$$

where

$$\check{D}_{\check{\Psi}}(\lambda_1, \lambda_2) := \begin{cases} \check{\Psi}(\lambda_1) - \check{\Psi}(\lambda_2) - \sup_{\lambda_3 \in \partial\check{\Psi}(\lambda_2)} \{ [\![\lambda_1 - \lambda_2, \lambda_3]\!]_{\mathbb{R}^n \times \mathbb{R}^n} \} & : \lambda_1 \in \mathrm{efd}(\check{\Psi}), \lambda_2 \in \mathrm{efd}(\check{\Psi}) \cap \mathrm{efd}(\partial\check{\Psi}) \\ 0 & : \lambda_2 = \lambda_1 \in \mathrm{efd}(\check{\Psi}), \lambda_2 \notin \mathrm{efd}(\partial\check{\Psi}) \\ +\infty & : \text{otherwise}. \end{cases} \tag{217}$$

Some of these definitions are special cases of others, which can be written symbolically as:

$$(\mathrm{B}_1) \supseteq (\mathrm{B}_2) \supseteq (\mathrm{B}_3) \supseteq (\mathrm{B}_4) \supseteq (\mathrm{B}_5) \subseteq (\mathrm{B}_6). \tag{218}$$

The definitions (B$_1$), (B$_2$) and (B$_6$) are intended to deal with nondifferentiable functions Ψ. The definition (B$_6$) (and (B$_5$), if equipped with some additional condition) can be also considered as a definition of the *Brègman distance* (see next section). In all cases, (B$_1$)-(B$_6$), the convexity of Ψ implies $\widetilde{D}_\Psi(x,y) \geq 0$. If Ψ is strictly convex, (B$_1$) is used, and any of the following inequivalent conditions holds,

$$\mathfrak{D}^{\mathrm{G}}_{+}\Psi(y; x-y) = \sup_{\hat{z} \in \partial\Psi(y)} \{ [\![x-y, \hat{z}]\!]_{X\times X^\star} \}, \tag{219}$$
$$\mathfrak{D}^{\mathrm{G}}_{+}\Psi(y; x-y) = -\sup_{\hat{z} \in \partial\Psi(y)} \{ [\![y-x, \hat{z}]\!]_{X\times X^\star} \}, \tag{220}$$

[24]Here we generalise the definition given in cited papers.

then [389]

$$\widetilde{D}_{\Psi}(x,y)=0 \iff x=y \;\; \forall x,y\in \mathrm{efd}(\Psi). \tag{221}$$

The equation (221) holds also for (B_2)-(B_4) under the same conditions as above, if $\forall x,y\in\mathrm{efd}(\Psi)$ is replaced by $\forall x,y\in\int(\mathrm{efd}(\Psi))$. If Ψ is strictly convex and (B_6) is used, then [191]

$$\widetilde{D}_{\Psi}(x,y)=0 \iff x=_{\tilde{\mu}} y \;\; \forall x,y\in \mathrm{efd}(\Psi). \tag{222}$$

For (B_3) the strict convexity of Ψ on $\mathrm{efd}(\Psi)$ implies that $D_{\Psi}(\cdot,y)$ is strictly convex on $\mathrm{efd}(\Psi)$ [9]. If X is reflexive and (B_2) is used, then for $(x,y)\in\mathrm{int}(\mathrm{efd}(\Psi))$ [61]:

1) $\widetilde{D}_{\Psi}(\cdot,y)$ is proper, convex, lower semi-continuous, with $\mathrm{efd}(\widetilde{D}_{\Psi}(\cdot,y))=\mathrm{efd}(\Psi)$;

2) $\widetilde{D}_{\Psi}(x,y)=\Psi(x)-\Psi(y)+\max_{\hat{z}\in\partial\Psi(y)}\left\{[\![y-x,\hat{z}]\!]_{X\times X^\star}\right\}$;

3) $\widetilde{D}_{\Psi}(x,y)=\Psi(x)+\Psi^{\mathbf{L}}(\hat{z})-[\![x,\hat{z}]\!]_{X\times X^\star}$ for all $\hat{z}\in\partial\Psi(y)$ such that

$$[\![y-x,\hat{z}]\!]_{X\times X^\star}=\max_{\hat{w}\in\partial\Psi(y)}\left\{[\![y-x,\hat{w}]\!]_{X\times X^\star}\right\}; \tag{223}$$

4) if Ψ is Gâteaux differentiable at y, then

$$\widetilde{D}_{\Psi}(x,y)=\Psi(x)-\Psi(y)-\left[\left[x-y,\mathfrak{D}^{\mathrm{G}}_{y}\Psi\right]\right]_{X\times X^\star}=\Psi(x)+\Psi^{\mathbf{L}}(\mathfrak{D}^{\mathrm{G}}_{y}\Psi)-\left[\left[x,\mathfrak{D}^{\mathrm{G}}_{y}\Psi\right]\right]_{X\times X^\star}; \tag{224}$$

5) if Ψ is essentially strictly convex, then

$$\widetilde{D}_{\Psi}(x,y)=0 \iff x=y; \tag{225}$$

6) if Ψ is Gâteaux differentiable at $\mathrm{int}(\mathrm{efd}(\Psi))$ and essentially strictly convex, then

$$\widetilde{D}_{\Psi}(x,y)=\widetilde{D}_{\Psi^{\mathbf{L}}}(\mathfrak{D}^{\mathrm{G}}_{y}\Psi,\mathfrak{D}^{\mathrm{G}}_{x}\Psi) \;\; \forall x\in\mathrm{int}(\mathrm{efd}(\Psi)). \tag{226}$$

We can conclude that the Brègman functional can be considered a distance if (Ψ is strictly convex, one of the conditions (223) holds, and (B_1) is used) or (Ψ is essentially strictly convex, X is reflexive, and (B_2) is used) or (Ψ is strictly convex and (B_6) is used).

If X is a Banach space and $\Psi: X\to]-\infty,+\infty]$ is proper, then an ***Alber functional*** on X is defined as [6, 7, 8]

$$W_{\Psi}: X\times X^\star\ni(x,\hat{y})\mapsto\Psi(x)+\Psi^{\mathbf{L}}(\hat{y})-[\![x,\hat{y}]\!]_{X\times X^\star}\in[0,+\infty]. \tag{227}$$

The condition (Ψ is Gâteaux differentiable at x and $\hat{y}=\mathfrak{D}^{\mathrm{G}}_{x}\Psi$) is equivalent to $W_{\Psi}(x,\hat{y})=0$. If Ψ is also convex, lower semi-continuous, and Gâteaux differentiable on $\mathrm{int}(\mathrm{efd}(\Psi))\neq\varnothing$, and X is reflexive, then the Young–Fenchel inequality gives

$$\Psi(x)+\Psi^{\mathbf{L}}(\mathfrak{D}^{\mathrm{G}}_{x}\Psi)-\left[\left[x,\mathfrak{D}^{\mathrm{G}}_{x}\Psi\right]\right]_{X\times X^\star}=0 \;\; \forall x\in\mathrm{int}(\mathrm{efd}(\Psi)) \tag{228}$$

and

$$\begin{aligned} W_{\Psi}(x,\mathfrak{D}^{\mathrm{G}}_{y}\Psi) &= \Psi(x)+\Psi^{\mathbf{L}}(\mathfrak{D}^{\mathrm{G}}_{y}\Psi)-\left[\left[x,\mathfrak{D}^{\mathrm{G}}_{y}\Psi\right]\right]_{X\times X^\star} \\ &= \Psi(x)-\Psi(y)-\left[\left[x-y,\mathfrak{D}^{\mathrm{G}}_{y}\Psi\right]\right]_{X\times X^\star} \\ &= \widetilde{D}_{\Psi}(x,y) \end{aligned} \tag{229}$$

for all $x,y\in\mathrm{int}(\mathrm{efd}(\Psi))$, with $\widetilde{D}_{\Psi}$ given by (B_3). These equations are special cases of (224).

If X is a Banach space and (B_3) is used, then for every $x, y \in X$ and $z, w \in \mathrm{int}(\mathrm{efd}(\Psi))$ [146, 64, 62]

$$\widetilde{D}_\Psi(z,w) + \widetilde{D}_\Psi(w,z) = \left[\!\left[z - w, \mathfrak{D}^{\mathrm{G}}_z\Psi - \mathfrak{D}^{\mathrm{G}}_w\Psi\right]\!\right]_{X\times X^\star}, \tag{230}$$

$$\widetilde{D}_\Psi(x,w) + \widetilde{D}_\Psi(w,z) = \widetilde{D}_\Psi(x,z) + \left[\!\left[x - w, \mathfrak{D}^{\mathrm{G}}_z\Psi - \mathfrak{D}^{\mathrm{G}}_w\Psi\right]\!\right]_{X\times X^\star}, \tag{231}$$

$$\widetilde{D}_\Psi(x,w) + \widetilde{D}_\Psi(y,z) - \widetilde{D}_\Psi(x,z) - \widetilde{D}_\Psi(y,w) = \left[\!\left[x - y, \mathfrak{D}^{\mathrm{G}}_z\Psi - \mathfrak{D}^{\mathrm{G}}_w\Psi\right]\!\right]_{X\times X^\star}. \tag{232}$$

The equation (231) is an instance of a ***generalised cosine equation***, while the equation (232) is an instance of a ***quadrilateral equation***.

A ***Brègman functional projection*** [144, 59, 61, 62] from a set $C_1 \subseteq X$ onto a set $C_2 \subseteq X$ is the function

$$\widetilde{\mathfrak{P}}^\Psi_{C_2|C_1} : C_1 \ni y \mapsto \left\{x \in C_2 \cap \mathrm{efd}(\Psi) \mid \widetilde{D}_\Psi(x,y) = \inf_{z\in C_2}\left\{\widetilde{D}_\Psi(z,y)\right\} < +\infty\right\} \in \wp(C_2). \tag{233}$$

For $C_1 = X$ we denote $\widetilde{\mathfrak{P}}^\Psi_{C_2} := \widetilde{\mathfrak{P}}^\Psi_{C_2|X}$. If $\widetilde{\mathfrak{P}}^\Psi_C(y) = \{x\}$, then we will use the notation $\widetilde{\mathfrak{P}}^\Psi_C(y) = x$. The main problems considered in the context of Brègman functional projections are their existence, uniqueness, characterisation, and stability (which means the behaviour of sequences converging to the unique solution of the minimisation problem). Various results, depending on different sets of assumptions, are present in the literature. Here we will present the main existence, uniqueness and characterisation results obtained for the Banach space setting and the measure theoretic setting (which generalise earlier results of [144, 207, 145, 711, 240, 59], obtained for $\mathbb{R}^n$).

(P_1) [9, 8]. If (B_3) is used, Ψ is strictly convex on $\mathrm{efd}(\Psi)$, $C \subseteq X$ is convex, and $C \cap \mathrm{efd}(\Psi) \neq \varnothing$, then $\widetilde{\mathfrak{P}}^\Psi_C(y)$ contains *at most* one element. If, in addition, X is reflexive and C is nonempty and weakly closed[25], then $\widetilde{\mathfrak{P}}^\Psi_C(y) = \{*\}$ $\forall y \in \mathrm{int}(\mathrm{efd}(\Psi))$ whenever ($C \cap \mathrm{efd}(\Psi)$ is norm bounded or $\lim_{\|x\|\to+\infty} \frac{\Psi(x)}{\|x\|} \to +\infty$ $\forall x \in C \cap \mathrm{efd}(\Psi)$). Moreover, if X is an arbitrary Banach space, (B_3) is used, Ψ is strictly convex, $C \subseteq X$ is nonempty and convex, $y \in X$, $x \in C$, then equivalent are:

$$\widetilde{D}_\Psi(z,x) + \widetilde{D}_\Psi(x,y) \leq \widetilde{D}_\Psi(z,y) \;\; \forall z \in C, \tag{234}$$

$$\left[\!\left[z - x, \mathfrak{D}^{\mathrm{G}}_y\Psi - \mathfrak{D}^{\mathrm{G}}_x\Psi\right]\!\right]_{X\times X^\star} \leq 0 \;\; \forall z \in C, \tag{235}$$

$$x = \widetilde{\mathfrak{P}}^\Psi_C(y). \tag{236}$$

If X and Ψ are as above, K is a vector subspace of X, $C \subseteq K$ is a nonempty closed convex set, then

$$\widetilde{D}_\Psi(x,y) = \widetilde{D}_\Psi(x, \widetilde{\mathfrak{P}}^\Psi_K(x)) + \widetilde{D}_\Psi(\widetilde{\mathfrak{P}}^\Psi_K(x), y) \quad \forall (x,y) \in K \times X. \tag{237}$$

(P_2) [61, 62, 63]. If (B_2) is used, Ψ is Legendre (or if Ψ is strictly convex, essentially strictly convex, and Gâteaux differentiable at y), $y \in \mathrm{int}(\mathrm{efd}(\Psi))$, X is reflexive, and $C \subseteq X$ is nonempty convex closed, $C \cap \mathrm{int}(\mathrm{efd}(\Psi)) \neq \varnothing$, then $\widetilde{\mathfrak{P}}^\Psi_C(y) = \{*\}$ and

$$x = \widetilde{\mathfrak{P}}^\Psi_C(y) \iff \left(x \in C \;\text{ and }\; C \subseteq \left\{z \in X \mid \left[\!\left[z - x, \mathfrak{D}^{\mathrm{G}}_y\Psi - \mathfrak{D}^{\mathrm{G}}_x\Psi\right]\!\right]_{X\times X^\star} \leq 0\right\}\right) \tag{238}$$

$\forall x, y \in \mathrm{int}(\mathrm{efd}(\Psi))$, which is equivalent to characterisation of $\widetilde{\mathfrak{P}}^\Psi_C(y)$ as a unique $x \in C \cap \mathrm{int}(\mathrm{efd}(\Psi))$ that satisfies

$$\widetilde{D}_\Psi(z,x) + \widetilde{D}_\Psi(x,y) \leq \widetilde{D}_\Psi(z,y) \;\; \forall z \in C. \tag{239}$$

(P_3) [132, 615, 134, 8]. A function $\Psi : X \to]-\infty, +\infty]$ is called ***totally convex*** [131] at $y \in \mathrm{efd}(\Psi)$ iff

$$\inf\{\widetilde{D}_\Psi(x,y) \mid x \in \mathrm{efd}(\Psi),\; \|x - y\| = t\} > 0 \;\; \forall t > 0, \tag{240}$$

[25] Note that, by the Hahn–Banach theorem, each norm closed convex set in a reflexive Banach space is weakly closed.

where (B_1) is used. For $C\subseteq X$ and $\hat{y}\in X^\star$ its ***Alber projection*** reads

$$\hat{\mathfrak{P}}^\Psi_C(\hat{y}) := \arg\inf_{x\in C}\left\{W_\Psi(x,\hat{y})\right\}. \tag{241}$$

If X is reflexive, Ψ is strictly convex and lower semi-continuous on $\mathrm{efd}(\Psi)$, $\hat{y}\in\mathrm{efd}(\Psi^{\mathbf{L}})$, $C\subseteq\mathrm{efd}(\Psi)$ is nonempty, convex, and closed, and the set $\{z\in\mathrm{efd}(\Psi)\mid W_\Psi(z,\hat{y})\leq\lambda\}$ is bounded for any $\lambda\in\mathbb{R}^+$ (which holds if $\lim_{\|x\|\to+\infty}\frac{\Psi(x)}{\|x\|}=+\infty$, or if ($\hat{y}\in\mathrm{efc}(\partial\Psi)$ and Ψ is totally convex at each $x\in\mathrm{efd}(\Psi)$)), then $\hat{\mathfrak{P}}^\Psi_C(\hat{y})=\{*\}$. If, in addition, Ψ is Gâteaux differentiable on $\mathrm{int}(\mathrm{efd}(\Psi))\neq\varnothing$, then $\hat{\mathfrak{P}}^\Psi_C(\hat{y})$ is characterised as such $x\in C$ that is a unique solution in C of

$$\left[\!\left[x-z,\hat{y}-\mathfrak{D}^{\mathrm{G}}_x\Psi\right]\!\right]_{X\times X^\star}\leq 0\ \ \forall z\in C, \tag{242}$$

or, equivalently, of

$$W_\Psi(x,\hat{y})+W_\Psi(z,\mathfrak{D}^{\mathrm{G}}_x\Psi)\leq W_\Psi(z,\hat{y})\ \ \forall z\in C. \tag{243}$$

From $\mathfrak{D}^{\mathrm{G}}_x\Psi\in\mathrm{int}(\mathrm{efd}(\Psi^{\mathbf{L}}))$ $\forall x\in\mathrm{int}(\mathrm{efd}(\Psi))$ it follows that

$$\widetilde{\mathfrak{P}}^\Psi_C(x)=\hat{\mathfrak{P}}^\Psi_C(\mathfrak{D}^{\mathrm{G}}_x\Psi)=\left(\hat{\mathfrak{P}}^\Psi_C\circ\mathfrak{D}^{\mathrm{G}}\Psi\right)(x). \tag{244}$$

This implies existence and uniqueness of the Brègman functional projection for (B_3) under the above conditions, with $x=\widetilde{\mathfrak{P}}^\Psi_C(y)$ for $y\in\mathrm{int}(\mathrm{efd}(\Psi))$ characterised as a unique solution in C of

$$\left[\!\left[x-z,\mathfrak{D}^{\mathrm{G}}_y\Psi-\mathfrak{D}^{\mathrm{G}}_x\Psi\right]\!\right]_{X\times X^\star}\geq 0\ \ \forall z\in C, \tag{245}$$

or, equivalently, of

$$\widetilde{D}_\Psi(z,x)+\widetilde{D}_\Psi(x,y)\leq\widetilde{D}_\Psi(z,y)\ \ \forall z\in C. \tag{246}$$

(P_4) [745]. If (B_3) is used, $C\subseteq X$, $C\cap\mathrm{efd}(\Psi)\neq\varnothing$, $\mathfrak{D}^{\mathrm{G}}\Psi(\mathrm{int}(\mathrm{efd}(\Psi)))=X^\star$, and $y\in\mathrm{int}(\mathrm{efd}(\Psi))$, then

$$(\widetilde{\mathfrak{P}}^\Psi_C(y)=\{*\})\iff(\Psi+\imath_C\text{ is adequate}) \tag{247}$$

$$\overset{(\mathrm{P}^1_4)}{\iff}(\Psi+\imath_C\text{ is essentially strictly convex}) \tag{248}$$

$$\overset{(\mathrm{P}^2_4)}{\iff}(C\text{ is convex}), \tag{249}$$

where

$$\imath_C: X\ni x\mapsto\begin{cases}0: & x\in C\\ +\infty & : x\notin C\end{cases} \tag{250}$$

is called an ***indicator function***, (P^1_4) denotes an additional assumption that X is reflexive, while (P^2_4) denotes an additional assumption that X reflexive, Ψ is Legendre and cofinite, and C is weakly closed set with $C\subseteq\mathrm{int}(\mathrm{efd}(\Psi))$.

Equation (237) is an instance of a ***generalised pythagorean equation***. In [210, 58] an instance of (237) has been established for $\widetilde{\mathfrak{P}}^\Psi_C(y)$ with $\widetilde{D}_\Psi$ defined by (B_2), $X=\mathbb{R}^n$, Ψ Legendre but not necessarily lower semi-continuous, and (P_2) with $C=K+x_0$, where $K\subseteq X$ is a closed vector subspace and $x_0\in\mathrm{int}(\mathrm{efd}(\Psi))$. Another instance of a generalised pythagorean equation, independent of (237), was established in a measure theoretic setting of (B_5) in [190].

The Brègman functional projection from X onto $C_1\subseteq X$ is called ***zone consistent*** iff it is a singleton and it can be subjected to another Brègman functional projection with the same Ψ, with X replaced by C_1, and with C_1 replaced by C_2, which is of the same type as C_1 [144, 145, 59].[26]

[26]More precisely, and in a strict agreement with the definition at the beginning of the Section 3, a family of Brègman functional projections $\{\widetilde{\mathfrak{P}}^\Psi_{C_i|C_j}\mid(i,j)\in I\times J\}$, for some sets I and J, is called ***zone consistent*** iff $\widetilde{\mathfrak{P}}^\Psi_{C_i|C_j}(\psi)=\{*\}$ $\forall(i,j)\in I\times J$ $\forall\psi\in C_j$ and $\widetilde{\mathfrak{P}}^\Psi_{C_i|C_j}(\psi)=\widetilde{\mathfrak{P}}^\Psi_{C_i|C_k}\circ\widetilde{\mathfrak{P}}^\Psi_{C_k|C_j}(\psi)$ $\forall\psi\in C_j$ $\forall k\in I\cap J$. In the given case, the zone consistency is obtained by requiring Ψ to be a Legendre function on a reflexive Banach space X, together with the condition (P_2), which imposes the restriction of all C_i to nonempty convex closed subsets of $\mathrm{int}(\mathrm{efd}(\Psi))$ as well as the restriction of all C_j to nonempty subsets of $\mathrm{int}(\mathrm{efd}(\Psi))$.

According to [61], if the conditions (P_2) for $\widetilde{\mathfrak{P}}^{\Psi}_{C}(y) = \{*\}$ are used with Ψ Legendre, then $\widetilde{\mathfrak{P}}^{\Psi}_{C}(y)$ is zone consistent (meaning: $\widetilde{\mathfrak{P}}^{\Psi}_{C}(y) \in \mathrm{int}(\mathrm{efd}(\Psi))$) and $\widetilde{\mathfrak{P}}^{\Psi}_{C}(\widetilde{\mathfrak{P}}^{\Psi}_{C}(y)) = \widetilde{\mathfrak{P}}^{\Psi}_{C}(y)$. According to [8], if the conditions (P_1) are used, then $\widetilde{\mathfrak{P}}^{\Psi}_{C}(y) = \widetilde{\mathfrak{P}}^{\Psi}_{C}(\widetilde{\mathfrak{P}}^{\Psi}_{K}(x)) = \widetilde{\mathfrak{P}}^{\Psi}_{K}(\widetilde{\mathfrak{P}}^{\Psi}_{C}(x))$ for C as in (P_3) and K a vector subspace of X with $C \subseteq K$.

The Brègman functional (B_5) has been characterised in [366] by means of a generalised pythagorean equation. The Brègman functional (B_3) has been characterised in finite dimensional case of $X = \mathbb{R}^n$ (for which it coincides with (B_4)) in [182] by a set of conditions which have geometric character, and in [50] by the condition that

$$\arg\inf_{y\in X}\left\{\int_{\mathcal{X}} \tilde{\mu}(\chi)\widetilde{D}_{\Psi}(x(\chi), y)\right\} = \int_{\mathcal{X}} \tilde{\mu}(\chi)x(\chi) \tag{251}$$

for some measure space $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ and $\tilde{\mu}$-integrable function $x : \mathcal{X} \to X$. Generalisation of equation (251) (but not of the associated characterisation) to (B_4) in arbitrary dimension, under some additional conditions, was provided in [264, 253]. The equality (251) was proved for the family of Liese–Vajda γ-distances (283) in [792].

3.3 Brègman distances

Our main objects of interest are not Brègman *functionals*, but Brègman *distances*, considered over information models. While most of research deals only with Brègman functionals on vector spaces as presented in the previous section, we will follow here the idea represented in [23, 784, 360, 20, 19, 403], according to which Brègman distances shall be defined in terms of Brègman functionals on vector spaces composed with (nonlinear) embeddings of statistical or quantum models. Apart from requirement $D_{\Psi}(\psi, \phi) = 0 \iff \phi = \psi$, this approach stresses that a Brègman distance is an information distance defined by means of some choice of representation of information model in a linear space, which forms a domain for corresponding Brègman functional. This formulation amounts to expose the dualistic properties of Brègman distance that are responsible for generalised pythagorean theorem. The novel aspect of our work is a systematic treatment of an extension of this approach to infinite dimensional case. The main idea is to introduce a generalisation of a Brègman *functional* using the Young–Fenchel inequality (194), and to subsequently define a Brègman distance over an arbitrary set Z, using this functional together with a *pair* of (not necessarily linear) embeddings $(\ell, \ell^{@}) : Z \times Z \to X \times X^{\mathbf{d}}$ into a dual pair of vector spaces. Finally, we define quantum Brègman distance as a dualistic Brègman distance with $Z = \mathcal{N}^{+}_{\star}$.

The current stage of development of this approach does not lead to any strong theorems. Nevertheless, it introduces a valuable structural clarification, and we consider it an important tool that might help unify various results in the theory of Brègman distances. In particular, we will use it in Section 3.4 to analyse the properties of a family of quantum γ-distances (299).

Given a dual pair $(X, X^{\mathbf{d}}, [\![\cdot,\cdot]\!]_{X\times X^{\mathbf{d}}})$ over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ and a convex proper $\Psi : X \to \mathbb{R} \cup \{+\infty\}$, let us define a ***generalised Alber functional*** as a map

$$X \times X^{\mathbf{d}} \ni (x, \hat{y}) \mapsto W_{\Psi}(x, \hat{y}) := \Psi(x) + \Psi^{\mathbf{L}}(\hat{y}) - \mathrm{re}\,[\![x, \hat{y}]\!]_{X\times X^{\mathbf{d}}} \in [0, \infty]. \tag{252}$$

By definition and (194),

i) $W_{\Psi}(x, \hat{y})$ is convex in each variable separately,

ii) $W_{\Psi}(x, \hat{y}) \geq 0\ \forall (x, \hat{y}) \in X \times X^{\mathbf{d}}$,

iii) $W_{\Psi}(x, \hat{y}) = 0 \iff (\hat{y} \in \partial\Psi(x)$ and $x \in \mathrm{efd}(\partial\Psi))$.

If X is a Banach space and $X^{\mathbf{d}} = X^{\star}$ with duality given by Banach space duality, then a generalised Alber functional (252) coincides with an Alber functional (227).[27]

[27] We have proposed the definition (252) in [403], while being unaware of Alber's work (which is summarised in Section 3.2).

For a given dual pair $(X, X^{\mathbf{d}}, [\![\cdot,\cdot]\!]_{X\times X^{\mathbf{d}}})$ a ***dual coordinate system*** on a set Z is defined as a map

$$(\ell, \ell^{@}) : Z \times Z \ni (\omega, \phi) \mapsto (\ell(\omega), \ell^{@}(\phi)) \in X \times X^{\mathbf{d}}. \tag{253}$$

If $W_\Psi : X \times X^{\mathbf{d}} \to [0, \infty]$ is a generalised Alber functional and $(\ell_\Psi, \ell^{@}_\Psi) : Z \times Z \to X \times X^{\mathbf{d}}$ is a dual coordinate system such that

$$\begin{cases} \partial\Psi(x) \neq \varnothing \ \ \forall x \in \mathrm{efd}(\partial\Psi) \cap \mathrm{cod}(\ell_\Psi) \\ \ell^{@}_\Psi(\omega) \in \partial\Psi(\ell_\Psi(\omega)) \ \ \forall \omega \in Z, \end{cases} \tag{254}$$

then a ***Brègman pre-distance*** is defined as a function

$$D_\Psi : Z \times Z \ni (\omega, \phi) \mapsto D_\Psi(\omega, \phi) := W_\Psi(\ell_\Psi(\omega), \ell^{@}_\Psi(\phi)) \in [0, \infty]. \tag{255}$$

The conditions (254) can be understood either as constraints on allowed dual coordinate systems if Ψ is given, or as constraints on Ψ if $(\ell_\Psi, \ell^{@}_\Psi)$ is given. By definition, $D_\Psi(\omega, \phi)$ is convex in each variable separately, $D_\Psi(\omega, \phi) \geq 0 \ \forall \omega, \phi \in Z$, and $\omega = \phi \Rightarrow D_\Psi(\omega, \phi) = 0 \ \forall \omega \in Z$. This weakening of the usual property of distance ($\omega = \phi \iff D(\omega, \phi) = 0$) is caused by restriction of domain of W_Ψ to $\mathrm{cod}(\ell_\Psi) \times \mathrm{cod}(\ell^{@}_\Psi)$. In order to impose an implication in the opposite direction, one would have to impose additional conditions that are not natural at this level of generality (they will be discussed below).

Definition (255) exposes the dualistic and variational structures underlying Brègman distances. However, the standard definition of Brègman distance uses only a single coordinate system instead of a dual pair, exposing geometric properties of Brègman distance and imposing $D_\Psi(x, y) = 0 \iff x = y$ at the price of nontrivial restrictions on the domain of duality and convexity. Usually these restrictions are introduced in order to adapt to *presupposed* topological and differential framework (e.g. of a reflexive Banach space), which imposes some specific restrictions on Brègman distance (as exemplified by various definitions of Brègman functional in previous section), and requires one to prove that such Brègman distance encodes the Legendre case of the Fenchel duality with the dual variable $y \in X^{\mathbf{d}}$ given by some suitably defined notion of derivative (e.g. Fréchet, Gâteaux, right Gâteaux), see e.g. [144, 59, 134, 113] for standard examples in commutative case, [576] for an example in the finite dimensional noncommutative case, and [360] for an example in the infinite dimensional noncommutative case.

Our approach is different, because we do not assume any fixed framework for continuity or smoothness, so we can consider general properties of the relationship between explicitly *dualistic* Brègman distance and its *standard* (hence, restricted) version, which has both arguments represented on the same space. The transition between these two formulations in the real finite dimensional case is provided by means of bijective ***Legendre transformation*** $\mathbf{L}_\Psi : \Theta \to \Xi$, which acts between suitable open subsets $\Theta \subset \mathbb{R}^n$ and $\Xi \subset \mathbb{R}^n$, and is given by the gradient,

$$\mathbf{L}_\Psi : \Theta \ni \theta \mapsto \eta := \mathrm{grad}\Psi(\theta) \in \Xi. \tag{256}$$

In the coordinate-dependent form this reads

$$\eta_i = (\mathbf{L}_\Psi(\theta))_i := \frac{\partial\Psi(\theta)}{\partial\theta^i}, \tag{257}$$

$$\theta^i = (\mathbf{L}^{-1}_\Psi(\eta))^i := \frac{\partial\Psi^{\mathbf{L}}(\eta)}{\partial\eta_i}, \tag{258}$$

whenever the duality pairing is given by

$$[\![\cdot,\cdot]\!]_{\mathbb{R}^n\times\mathbb{R}^n} : \mathbb{R}^n \times \mathbb{R}^n \ni (\theta, \eta) \mapsto \theta \cdot \eta^\top := \sum_{i=1}^n \theta^i \eta_i \in \mathbb{R}. \tag{259}$$

We will now construct a general framework for conversion between these two forms of the Brègman distance, which is independent of any particular assumptions about continuity or differentiability. The key element in this setting is the (generally, nonlinear) *dualiser* function. It will provide also an infinite dimensional generalisation of the bijective transformation between the dual coordinate systems that strengthens (254).

The relationship between dual coordinate systems is in the infinite dimensional case is more complicated than just replacing gradient by the Gâteaux derivative. It involves characterisation in terms of subdifferential, and depends on the function Ψ *and* on the specific structure of the dual pair $(X, X^{\mathbf{d}}, [\![\cdot,\cdot]\!]_{X\times X^{\mathbf{d}}})$ of vector spaces. In [403] we have proposed the following generalisation of the Legendre transformation to the case of arbitrary dual pair of vector spaces of arbitrary dimension, which preserves its bijective character without any fixed choice of topological background. The generalisation of (256) is provided by the ***dualiser***, defined as a map $\mathbf{L}_\Psi : X \to X^{\mathbf{d}}$ associated with a convex proper function $\Psi : X \to \mathbb{R} \cup \{+\infty\}$ such that there exists a nonempty set $\Theta_\Psi \subseteq \mathrm{efd}(\Psi)$ satisfying:

(i) $\mathbf{L}_\Psi$ is a bijection on Θ_Ψ,

(ii) $\Psi^{\mathbf{L}}(\mathbf{L}_\Psi(y)) - \Psi(y) = \mathrm{re}\, [\![y, \mathbf{L}_\Psi(y)]\!]_{X\times X^{\mathbf{d}}}\ \forall y \in \Theta_\Psi$,

(iii) $\mathbf{L}_\Psi(y) \in \partial\Psi(x) \iff x = y\ \forall x, y \in \mathrm{efd}(\partial\Psi)$.

If such $\mathbf{L}_\Psi$ exists, then Θ_Ψ will be called an ***admissible domain*** of $\mathbf{L}_\Psi$ and denoted $\mathrm{add}(\mathbf{L}_\Psi)$, while $\mathrm{adc}(\mathbf{L}_\Psi) \equiv \Xi_\Psi := \mathbf{L}_\Psi(\Theta_\Psi)$ will be called its ***admissible codomain***. The function Ψ will be called ***dualisable*** with respect to $(X, X^{\mathbf{d}}, [\![\cdot,\cdot]\!]_{X\times X^{\mathbf{d}}})$ iff there exists at least one dualiser $\mathbf{L}_\Psi$. Each triple $(\Theta_\Psi, \Xi_\Psi, \mathbf{L}_\Psi)$ will be called a ***generalised Legendre transformation***. A bijection

$$\mathbf{L}_\Psi : X \supseteq \Theta_\Psi \mapsto \Xi_\Psi \subseteq X^{\mathbf{d}}, \tag{260}$$

is a generalisation of (256). A change of domain X or a change of duality structure $[\![\cdot,\cdot]\!]_{X\times X^{\mathbf{d}}}$ on X changes the available dualisers. Also, there might be several different dualisers for a given quadruple $((X, X^{\mathbf{d}}, [\![\cdot,\cdot]\!]_{X\times X^{\mathbf{d}}}), \Psi)$. The existence of different dualisers is equivalent to $\partial\Psi$ being a nonsingleton, nonempty, set-valued function.

Given a generalised Legendre transformation $(\Theta_\Psi, \Xi_\Psi, \mathbf{L}_\Psi)$, we can define the ***Brègman functional*** $\bar{D}_\Psi : X \times X \to [0, +\infty]$ associated to a generalised Alber functional W_Ψ [403],

$$\bar{D}_\Psi(x, y) := \begin{cases} W_\Psi(x, \mathbf{L}_\Psi(y)) = \Psi(x) - \Psi(y) - \mathrm{re}\, [\![x - y, \mathbf{L}_\Psi(x)]\!]_{X\times X^{\mathbf{d}}} & : y \in \Theta_\Psi \\ +\infty & : \text{otherwise.} \end{cases} \tag{261}$$

The equality above follows from the property (ii) of $\mathbf{L}_\Psi$. The bounded version of this functional is given by restriction of the domain of (261) to $\bar{D}_\Psi : \mathrm{efd}(\Psi) \times \Theta_\Psi \to [0, \infty[$. From the property (iii) of $\mathbf{L}_\Psi$ it follows that $\bar{D}_\Psi$ satisfies

$$\bar{D}_\Psi(x, y) = 0 \iff x = y \quad \forall (x, y) \in X \times X, \tag{262}$$

or for all $(x, y) \in \mathrm{efd}(\Psi) \times \Theta_\Psi$ whenever $\bar{D}_\Psi$ is bounded. The equivalence appears here at the price of loss of convexity of $\bar{D}_\Psi$ in the second variable (it is a common problem in standard treatments, see e.g. [60]). This is because using the inverse of a dualiser $\mathbf{L}_\Psi$ may not preserve the convexity properties. From $\Theta \subseteq \mathrm{efd}(\Psi)$ is follows that the definition (261) is a generalisation of (B_3). We will call this definition ($\mathrm{B}_{\bar{D}}$), and consider it as an alternative to (B_2), aimed at preservation of convex and dualistic properties without reducing them to the setting of topological differentiability. From the results discussed in the previous section it follows that (B_2) with reflexive X and Legendre Ψ is a special case of ($\mathrm{B}_{\bar{D}}$). More precisely, if X is a reflexive Banach space, $X^{\mathbf{d}} = X^\star$, Ψ is convex, proper, lower semi-continuous, and Legendre, then $(\Theta_\Psi, \Xi_\Psi, \mathbf{L}_\Psi)$ is given by $(\mathrm{int}(\mathrm{efd}(\Psi)), \mathrm{int}(\mathrm{efd}(\Psi^{\mathbf{L}})), \mathfrak{D}^{\mathrm{G}})$ due to (209), and in such case (261) reduces to (212). Properties (261) and (262) follow then from (229), and property 5) in Section 3.2, respectively.

Let $(X, X, [\![\cdot,\cdot]\!]_{X\times X^{\mathbf{d}}})$ be a dual pair, let $\Psi : X \to \mathbb{R} \cup \{+\infty\}$ be a convex proper function, let $(\Theta_\Psi, \Xi_\Psi, \mathbf{L}_\Psi)$ be a generalised Legendre transformation, let Z be a set, and let $(\ell_\Psi, \ell_\Psi^{@}) : Z \times Z \to$

$X\times X^{\mathbf{d}}$ be a dual coordinate system such that $\mathrm{cod}(\ell^{@}_{\Psi})\subseteq\Xi_{\Psi}$. Then we define the ***dualistic Brègman distance*** on Z as a function $D_{\Psi}:Z\times Z\to[0,\infty]$ such that

$$\begin{aligned}D_{\Psi}(\omega,\phi)&:=W_{\Psi}(\ell_{\Psi}(\omega),\ell^{@}_{\Psi}(\phi))\\&=\bar{D}_{\Psi}(\ell_{\Psi}(\omega),\mathbf{L}^{-1}_{\Psi}\circ\ell^{@}_{\Psi}(\phi))\\&=\Psi(\ell_{\Psi}(\omega))-\Psi(\mathbf{L}^{-1}_{\Psi}\circ\ell^{@}_{\Psi}(\phi))-\mathrm{re}\,\left[\!\!\left[\ell_{\Psi}(\omega)-\mathbf{L}^{-1}_{\Psi}\circ\ell^{@}_{\Psi}(\phi),\ell^{@}_{\Psi}(\phi)\right]\!\!\right]_{X\times X^{\mathbf{d}}}.\end{aligned}\tag{263}$$

Note that it is possible to weaken the above definition by weakening the condition (iii) of definition of $\mathbf{L}_{\Psi}$ by replacing $\mathrm{efd}(\partial\Psi)$ and $\mathbf{L}_{\Psi}(y)$ by $\mathrm{efd}(\partial\Psi)\cap\mathrm{cod}(\ell_{\Psi})$ and $\mathbf{L}_{\Psi}(y)\cap\mathrm{cod}(\ell^{@}_{\Psi})$ respectively. Both definitions imply

$$D_{\Psi}(\omega,\phi)=0\iff\omega=\phi\quad\forall\omega,\phi\in Z.\tag{264}$$

It follows that a single Brègman pre-distance (255) may have several different representations in terms of dualistic Brègman distances, depending on the choice of the dualiser $\mathbf{L}_{\Psi}$ (263), corresponding to the choice of the generalised Legendre transformation $(\Theta_{\Psi},\Xi_{\Psi},\mathbf{L}_{\Psi})$. If $\bar{D}_{\Psi,\mathbf{L}_1}$ and $\bar{D}_{\Psi,\mathbf{L}_2}$ are two Brègman functionals defined from a single generalised Alber functional W_{Ψ} by two dualisers $\mathbf{L}_1$ and $\mathbf{L}_2$ of Ψ, then they are equal to each other on $V\subseteq\mathrm{add}(\mathbf{L}_1)\cap\mathrm{add}(\mathbf{L}_2)$ iff there exists a dualiser $\mathbf{L}_3$ of Ψ such that $\mathrm{add}(\mathbf{L}_3)=V$. Every choice of a triple $(\Theta_{\Psi},\Xi_{\Psi},\mathbf{L}_{\Psi})$ that turns Brègman pre-distance to a dualistic Brègman distance can be considered as a *localisation* of the former.

Especially interesting case of the dualistic Brègman distance (263) is when the equality

$$\ell^{@}_{\Psi}=\mathbf{L}_{\Psi}\circ\ell_{\Psi}\tag{265}$$

holds for all elements of Z. Relation (265) is a special case of (254) and allows to rewrite (263) as

$$D_{\Psi}(\omega,\phi)=\bar{D}_{\Psi}(\ell_{\Psi}(\omega),\ell_{\Psi}(\phi))=\Psi(\ell_{\Psi}(\omega))-\Psi(\ell_{\Psi}(\phi))-\mathrm{re}\,[\![\ell_{\Psi}(\omega)-\ell_{\Psi}(\phi),\mathbf{L}_{\Psi}\circ\ell_{\Psi}(\phi)]\!]_{X\times X^{\mathbf{d}}},\tag{266}$$

which does not depend on $\ell^{@}_{\Psi}$. Functional of the form (266) will be called a ***standard Brègman distance***. In particular, if $X=X^{\mathbf{d}}=\mathbb{R}^n$ with duality given by (259), $\Psi:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ is convex and proper, $\mathbf{L}_{\Psi}$ is given by the Legendre transformation (256), $\bar{D}_{\Psi}$ is given by a functional introduced originally by Brègman in [118],

$$\bar{D}_{\Psi}(x,y)=\Psi(x)-\Psi(y)-[\![x-y,\mathrm{grad}\Psi(y)]\!]_{\mathbb{R}^n\times\mathbb{R}^n},\tag{267}$$

$Z=\mathcal{M}(\mathcal{A})\subseteq L_1(\mathcal{A})^+$ for some mcb-algebra $\mathcal{A}$ or $Z=\mathcal{M}(\mathcal{N})\subseteq L_1(\mathcal{N})^+$ for some W^*-algebra $\mathcal{N}$, $\dim Z=:n<\infty$, while $(\ell_{\Psi},\ell^{@}_{\Psi})$ satisfies (265) by means of

$$\ell^{@}_{\Psi}=\mathrm{grad}\Psi(\ell_{\Psi}(\cdot)),\tag{268}$$

so the generalised Legendre transformation is determined by such $(\Theta_{\Psi},\Xi_{\Psi})$ that $\mathrm{cod}(\ell_{\Psi})\subseteq\Xi_{\Psi}$, then the associated standard Brègman distance reads

$$D_{\Psi}(\omega,\phi)=\Psi(\ell_{\Psi}(\omega))-\Psi(\ell_{\Psi}(\phi))-\sum_{i=1}^{n}\left(\ell_{\Psi}(\omega)-\ell_{\Psi}(\phi)\right)^i\left(\mathrm{grad}\Psi(\ell_{\Psi}(\phi))\right)_i.\tag{269}$$

If $\mathcal{A}$ is represented in terms of a measureable space $(\mathcal{X},\mho(\mathcal{X}),\mho^0(\mathcal{X}))$ and if ϕ_1 and ϕ_2 and densities in $\mathrm{MeFun}(\mathcal{X},\mho(\mathcal{X});\mathbb{R}^n)$ with respect to a fixed measure $\tilde{\mu}$ on $(\mathcal{X},\mho(\mathcal{X}))$ such that $\mho^{\tilde{\mu}}(\mathcal{X})=\mho^0(\mathcal{X})$, so that they can be identified with the elements of $L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu};\mathbb{R}^n)$, and if

$$\Psi(\ell_{\Psi}(\phi_i))=\int_{\mathcal{X}}\tilde{\mu}(x)\check{\Psi}(\phi_i(x)),\tag{270}$$

then (269) takes the form (B$_5$), with domain of $\check{\Psi}$ generalised from $\mathbb{R}^+$ to $(\mathbb{R}^+)^n$. If $\mathcal{N}=\mathfrak{B}(\mathcal{H})$ with $\dim\mathcal{H}<\infty$, and $\phi_1,\phi_2\in\mathfrak{G}_1(\mathcal{H})^+_0$, then a condition analogous to (270) reads (cf. [313, 211])

$$\Psi(\ell_{\Psi}(\phi_i))=\mathrm{tr}(\check{\Psi}(\phi_i)),\tag{271}$$

where $\check{\Psi} : \mathbb{R} \to]-\infty, +\infty]$ is proper, operator strictly convex function, differentiable on $]0, +\infty[$ with $\check{\Psi}(0) = \lim_{t\to+0} \check{\Psi}(t)$ and $t < 0 \Rightarrow \check{\Psi}(t) = +\infty$, and it is applied to density operator ϕ_i in terms of functional calculus on its spectrum.

Note that the relations (268), (265), and (254) quite specifically correspond to three sectors of the information geometry theory: finite dimensional, infinite dimensional with good duality properties, and generally infinite dimensional.

From the definitions (255) and (252) it follows that every dualistic Brègman distance D_Ψ with its corresponding dual coordinate system $(\ell_\Psi, \ell^{@}_\Psi)$ satisfies the ***quadrilateral equation***

$$D_\Psi(z_1, z_2) + D_\Psi(z_4, z_3) - D_\Psi(z_1, z_3) - D_\Psi(z_4, z_2) = \mathrm{re}\ \left[\!\!\left[\ell_\Psi(z_1) - \ell_\Psi(z_4), \ell^{@}_\Psi(z_3) - \ell^{@}_\Psi(z_2)\right]\!\!\right]_{X\times X^{\mathbf{d}}}, \tag{272}$$

and the ***generalised cosine equation***

$$D_\Psi(z_1, z_2) + D_\Psi(z_2, z_3) - D_\Psi(z_1, z_3) = \mathrm{re}\ \left[\!\!\left[\ell_\Psi(z_1) - \ell_\Psi(z_2), \ell^{@}_\Psi(z_3) - \ell^{@}_\Psi(z_2)\right]\!\!\right]_{X\times X^{\mathbf{d}}}, \tag{273}$$

for all $z_1, z_2, z_3, z_4 \in Z$ (cf. [785]). From the definition (261) of bounded Brègman functional $\bar{D}_\Psi$ it follows that $\bar{D}_\Psi$ satisfies the generalised cosine equation that generalises (231),

$$\bar{D}_\Psi(x_1, x_2) + \bar{D}_\Psi(x_2, x_3) - \bar{D}_\Psi(x_1, x_3) = \mathrm{re}\ [\![x_1 - x_2, \mathbf{L}_\Psi(x_3) - \mathbf{L}_\Psi(x_2)]\!]_{X\times X^{\mathbf{d}}} \tag{274}$$

$\forall x_1, x_2, x_3 \in \mathrm{add}(\mathbf{L}_\Psi)\cap\mathrm{efd}(\Psi)$, and it also satisfies the corresponding generalisation of the quadrilateral relation (232). From (274) it follows that for any given $x, y, \bar{y} \in \mathrm{add}(\mathbf{L}_\Psi) \cap \mathrm{efd}(\Psi)$, the ***generalised orthogonal decomposition***

$$\bar{D}_\Psi(x, \bar{y}) + \bar{D}_\Psi(\bar{y}, y) = \bar{D}_\Psi(x, y) \quad \forall x \in \mathrm{add}(\mathbf{L}_\Psi) \cap \mathrm{efd}(\Psi) \tag{275}$$

is equivalent with the ***orthogonality condition***,

$$\mathrm{re}\ [\![x - \bar{y}, \mathbf{L}_\Psi(y) - \mathbf{L}_\Psi(\bar{y})]\!]_{X\times X^{\mathbf{d}}} = 0. \tag{276}$$

Moreover, the equivalence holds also if $=$ is replaced by $\geq$ in (275) and $=$ is replaced by $\leq$ in (276). The generalised orthogonal decomposition (518) is a special case of (275). As we will see below, under suitable assumptions that guarantee the existence and uniqueness of solution of the corresponding variational problem, the generalised orthogonal decomposition can be turned into a theorem stating the existence and uniqueness of generalised additive decomposition of information distance under projection onto subspace (submodel), known as *generalised pythagorean theorem* (or *equation*).

Let $y \in \mathrm{add}(\mathbf{L}_\Psi)\cap\mathrm{efd}(\Psi)$, let $C \subseteq \mathrm{add}(\mathbf{L}_\Psi)\cap\mathrm{efd}(\Psi)$ be nonempty, convex, and containing at least one element z such that $\bar{D}_\Psi(z, y) < \infty$, let $x \in C$. In such case the Brègman functional projection (233) of y using $\bar{D}_\Psi$ will be denoted

$$\bar{y} \in \bar{\mathfrak{P}}^\Psi_C(y) = \arg\inf_{x\in C} \left\{\bar{D}_\Psi(x, y)\right\}. \tag{277}$$

The main problem with this definition is that in general case $\bar{\mathfrak{P}}^\Psi_C(y)$ might not exist or might be nonunique. The existence and uniqueness can follow from various assumptions. In particular, if X is a locally convex space, C is weakly compact, and $\bar{D}_\Psi$ is weakly lower semi-continuous, then the existence can be guaranteed by means of Bauer's theorem [56]. On the other hand, if X is a reflexive Banach space, C is closed, $\bar{D}_\Psi$ is lower semi-continuous, strictly convex, and Gâteaux differentiable at y, with $\mathrm{int}(\mathrm{efd}(\bar{D}_\Psi)) \neq \varnothing$, $C \cap \mathrm{efd}(\bar{D}_\Psi) \neq \varnothing$ and $y \in \mathrm{int}(\mathrm{efd}(\bar{D}_\Psi))$, then $\bar{\mathfrak{P}}^\Psi_C(y)$ is at most a singleton [113]. The conjunction of these two conditions is sufficient to guarantee the existence and uniqueness of $\bar{\mathfrak{P}}^\Psi_C(y)$. Unfortunately, we know neither the sufficient conditions for existence that would not require lower semi-continuity, nor the sufficient conditions for uniqueness that would not require Gâteaux differentiability.

If there exists a unique Brègman functional projection $\bar{y} = \bar{\mathfrak{P}}^\Psi_C(y)$ for $y \in \mathrm{add}(\mathbf{L}_\Psi) \cap \mathrm{efd}(\Psi)$, such that $(y, \bar{y})$ satisfies the orthogonality condition (276), then $\bar{y} = \bar{\mathfrak{P}}^\Psi_C(y)$ is called ***orthogonal***. Property

(275) generalises in such case the additive decompositions of norm under linear projections on closed convex subsets in the Hilbert space to the class of nonlinear $\bar{\mathfrak{P}}^{\Psi}_{C}$ projections onto convex subsets C in the linear space X. Note that the 'orthogonality' of projection is understood in the sense of the bilinear duality pairing $[\![\cdot,\cdot]\!]_{X\times X^{\mathbf{d}}}$, while the nonlinearity of projection $\bar{\mathfrak{P}}^{\Psi}_{C}$ corresponds to the nonlinear dualiser $\mathbf{L}_{\Psi}$. In particular, if $\bar{D}_{\Psi}$ is given by (B_3), then condition (276) turns to equality in (245), so the orthogonality condition (276) satisfied by $\bar{y} = \bar{\mathfrak{P}}^{\Psi}_{C}(y)$ turns to generalised pythagorean equation (237).

Given a dualistic Brègman distance D_{Ψ} on Z and $K_1, K_2 \subseteq Z$, we define a ***dualistic Brègman projection*** as a map

$$\mathfrak{P}^{D_{\Psi}}_{K_2|K_1} : K_1 \ni \phi \mapsto \arg\inf_{\omega\in K_2}\{D_{\Psi}(\omega,\phi)\} \subseteq \wp(K_2), \tag{278}$$

with $\mathfrak{P}^{D_{\Psi}}_{K_2} := \mathfrak{P}^{D_{\Psi}}_{K_2|Z}$. If $\ell_{\Psi}\times\ell^{@}_{\Psi}$ is bijective on $K_2\times K_1$, then the existence (resp., uniqueness) of $\mathfrak{P}^{D_{\Psi}}_{K_2|K_1}(\phi)$ follows from the existence(resp., uniqueness) of $\bar{\mathfrak{P}}^{D_{\Psi}}_{\ell_{\Psi}(K_2)|\mathbf{L}_{\Psi}^{-1}\circ\ell^{@}_{\Psi}(K_1)}\left(\ell^{@}_{\Psi}(\phi)\right)$. The generalised cosine equation (274) and the above discussion leads us to call a dualistic Brègman projection $\mathfrak{P}^{D_{\Psi}}_{K}(\psi)$ ***orthogonal*** iff it is a singleton and satisfies

$$\mathrm{re}\left[\!\!\left[\ell_{\Psi}(\phi)-\ell_{\Psi}(\mathfrak{P}^{D_{\Psi}}_{K}(\psi)),\ell^{@}_{\Psi}(\psi)-\ell^{@}_{\Psi}(\mathfrak{P}^{D_{\Psi}}_{K}(\psi))\right]\!\!\right]_{X\times X^{\mathbf{d}}} = 0 \;\;\forall\phi\in K, \tag{279}$$

which is equivalent the ***generalised pythagorean equation***

$$D_{\Psi}(\phi,\mathfrak{P}^{D_{\Psi}}_{K}(\psi)) + D_{\Psi}(\mathfrak{P}^{D_{\Psi}}_{K}(\psi),\psi) = D_{\Psi}(\phi,\psi) \;\;\forall\phi\in K. \tag{280}$$

The problem of characterisation of orthogonal $\mathfrak{P}^{D_{\Psi}}_{K}$ for a given D_{Ψ} and K remains open.

Let us summarise the insights gained in last two sections. There are few different candidates for the general notion of a ***Brègman distance*** on a general Banach space:

(BD_1) the Brègman functional $\widetilde{D}_{\Psi}$ defined by (B_1) under additional assumptions that Ψ is strictly convex on $\mathrm{efd}(\Psi)$ and that one of the equations (220) holds;

(BD_2) the Brègman functional $\bar{D}_{\Psi}$ defined by ($\mathrm{B}_{\bar{D}}$), with duality given by Banach space duality;

(BD_3) the dualistic Brègman distance (263), which is defined as a special case of ($\mathrm{B}_{\bar{D}}$), but its domain is shifted to the space Z, which in turn can be an arbitrary subset of a Banach space;

(BD_4) the Brègman functional $\widetilde{D}_{\Psi}$ defined by (B_2) for reflexive X and Ψ essentially strictly convex on $\mathrm{int}(\mathrm{efd}(\Psi)) \neq \varnothing$;

(BD_5) defined as (BD_4), but with an additional assumption of essential Gâteaux differentiability on $\mathrm{int}(\mathrm{efd}(\Psi))$. This is a special case of both (BD_1) and (BD_2).

In principle, there are three main properties that one would expect from a general notion of the Brègman distance:

- it should be a distance;
- it should possess well defined existence and uniqueness properties for the Brègman projections onto a well defined class of subsets;
- it should allow for generalised pythagorean and cosine theorems;
- it should be zone consistent.

All above candidates satisfy the first condition. The second and fourth conditions can be guaranteed at the level of (BD_5). The third condition requires either to strenghten (B_2) in (BD_5) to (B_3), in order to use (P_1), or to use (BD_3) with an additional orthogonality condition (279). However, the condition (279) is abstract and we do not know what are necessary and sufficient conditions for it to hold. On the other hand, using (BD_5) as a Brègman distance restricts the underlying Banach space to be reflexive. This is unacceptable restriction, because our main objective is to consider distances on subsets $\mathcal{M}(\mathcal{N}) \subseteq \mathcal{N}_\star^+$ of nonreflexive noncommutative $L_1(\mathcal{N}) \cong \mathcal{N}_\star^+$ spaces. This problem can be solved in a twofold way: either by using definition (BD_1) with $X = \mathcal{N}_\star$ and proving suitable existence, uniqueness, cosine, pythagorean, and zone consistency theorems by imposing additional conditions in a case-by-case mode, or by using definition (BD_3) with $Z = \mathcal{N}_\star^+$ and embeddings $(\ell_\Psi, \ell_\Psi^@) : \mathcal{N}_\star^+ \times \mathcal{N}_\star^+ \to X \times X^\star$ into a suitable Banach dual pair of reflexive Banach spaces, for which the existence, uniqueness, and pythagorean theorems based on (BD_5), with (B_3) instead of (B_2) inside, can be applied. These two distances will be called, respectively, ***weak quantum Brègman distance*** and ***dualistic quantum Brègman distance***. For $\mathcal{N} = L_\infty(\mathcal{A})$ an adjective 'quantum' will be replaced by 'statistical'. Some properties of (B_1) applied to $X = \mathfrak{G}_1(\mathcal{H}) = \mathfrak{B}(\mathcal{H})_\star$ for $\dim \mathcal{H} < \infty$ were analysed by Petz in [576] (see also [211]).

Note that ideally we would like to define *quantum Brègman distance* as such distance $D_\Psi : \mathcal{N}_\star^+ \times \mathcal{N}_\star^+ \to [0, \infty]$ that depends explicitly on Ψ and satisfies

(i) $\exists K_1 \subseteq \mathcal{N}_\star^+ \; \forall C \subseteq \mathcal{N}_\star^+ \; \forall (\phi, \psi) \in C \times K_1$

$$\mathfrak{P}^{D_\Psi}_{C|K_1}(\psi) = \{*\} \;\; \Rightarrow \;\; D_\Psi(\phi, \psi) = D_\Psi(\phi, \mathfrak{P}^{D_\Psi}_{C|K_1}(\psi)) + D_\Psi(\mathfrak{P}^{D_\Psi}_{C|K_1}(\psi), \psi); \tag{281}$$

(ii) $\exists K_2 \subseteq \mathcal{N}_\star^+ \; \forall C_1, C_2, C_3 \subseteq K_2$

$$\left(\mathfrak{P}^{D_\Psi}_{C_2|C_1}(\psi) = \{*\} \; \forall \psi \in C_1, \;\; \mathfrak{P}^{D_\Psi}_{C_3|C_2}(\phi) = \{*\} \; \forall \phi \in C_2\right) \; \Rightarrow \; \mathfrak{P}^{D_\Psi}_{C_3|C_1}(\psi) = \{*\} \; \forall \psi \in C_1. \tag{282}$$

However, as indicated by the gap between the weak and dualistic definitions above, it is still unclear how to maintain these two properties in a general nonreflexive setting of $\mathcal{N}_\star^+$ while: (1) keeping the explicit dependence on Ψ, (2) keeping the explicit dependence on Banach dual spaces, (3) allowing Araki distance D_1 (see next section) to be a special case of this definition (it can be defined in terms of (BD_1) by right Gâteaux derivative, but not in terms of (BD_3) by duality based on reflexive Banach spaces).

3.4 γ-distances

By imposing the condition of monotonicity under coarse graining on the dualistic Brègman distances (or on the corresponding standard Brègman distances), one obtains a strong restriction on the allowed forms of the 'generating' function Ψ and the corresponding dual coordinate systems $(\ell_\Psi, \ell_\Psi^@)$. Such families of information distances are of special interest, because they satisfy two main information theoretic constraints: existence of orthogonal decomposition under nonlinear projection 'onto submodel' and nonincreasing under 'information loss'.

Given $\gamma \in [0, 1]$ and an mcb-algebra $\mathcal{A}$, consider a family of the ***Liese–Vajda γ-distances*** [453, 790, 791, 792] on $L_1(\mathcal{A})^+$,

$$D_\gamma(\omega, \phi) := \begin{cases} \int \frac{1}{\gamma(1-\gamma)} \left(\gamma \mu_\omega + (1-\gamma)\nu_\phi - \nu_\phi \left(\frac{\mu_\omega}{\nu_\phi}\right)^\gamma\right) & : \; \gamma \in]0,1[, \; \mu_\omega \ll \nu_\phi \\ \int \lim_{\tilde{\gamma} \to^\pm \gamma} \frac{1}{\tilde{\gamma}(1-\tilde{\gamma})} \left(\tilde{\gamma}\mu_\omega + (1-\tilde{\gamma})\nu_\phi - \nu_\phi \left(\frac{\mu_\omega}{\nu_\phi}\right)^{\tilde{\gamma}}\right) & : \; \gamma \in \{0, 1\}, \; \mu_\omega \ll \nu_\phi \\ +\infty & : \; \text{otherwise}, \end{cases} \tag{283}$$

where the right limit, $\tilde{\gamma} \to^+ \gamma$, is considered for $\gamma = 0$, while the left limit, $\tilde{\gamma} \to^- \gamma$, is considered for $\gamma = 1$. Here μ_ω and ν_ϕ are finite positive measures corresponding to the positive integrals ω and ϕ, while $(\frac{\mu_\omega}{\nu_\phi})^\gamma$ denotes the γ-th power of the Radon–Nikodým quotient $\frac{\mu_\omega}{\nu_\phi}$, see Section 2.1. The boundary cases take the form

$$D_1(\omega, \phi) = \begin{cases} \int \left(\nu_\phi - \mu_\omega + \mu_\omega \log \frac{\mu_\omega}{\nu_\phi}\right) & : \; \mu_\omega \ll \nu_\phi \\ +\infty & : \; \text{otherwise}, \end{cases} \tag{284}$$

and

$$D_0(\omega,\phi)=\left\{\begin{array}{ll}\int\left(\mu_\omega-\nu_\phi-\nu_\phi\log\frac{\mu_\omega}{\nu_\phi}\right) & :\ \mu_\omega\ll\nu_\phi\\ +\infty & :\ \text{otherwise.}\end{array}\right. \tag{285}$$

It follows directly that D_γ satisfies

i) $\nu\ll\mu\ll\nu\Rightarrow D_\gamma(\mu,\nu)=D_{1-\gamma}(\mu,\nu)\ \forall\gamma\in[0,1]$,

ii) $D_\gamma(\lambda\mu,\lambda\nu)=\lambda D_\gamma(\mu,\nu)\ \forall\lambda\in]0,\infty[$.

A direct calculation shows that D_γ is a Csiszár–Morimoto $\mathfrak{f}$-distance with

$$\mathfrak{f}_\gamma(t)=\left\{\begin{array}{ll}\frac{1}{\gamma}+\frac{1}{1-\gamma}t-\frac{1}{\gamma(1-\gamma)}t^\gamma & :\gamma\in]0,1[\\ t\log t-(t-1) & :\gamma=1\\ -\log t+(t-1) & :\gamma=0,\end{array}\right. \tag{286}$$

which corresponds to

$$\mathfrak{f}^{\mathbf{c}}_\gamma(t)=\left\{\begin{array}{ll}\frac{1}{\gamma(1-\gamma)}(1-t^{1-\gamma})+\frac{1}{\gamma}(t-1) & :\gamma\in]0,1[\\ t\log t-(t-1) & :\gamma=0\\ -\log t+(t-1) & :\gamma=1.\end{array}\right. \tag{287}$$

These functions satisfy

$$\mathfrak{f}_0(t)=\lim_{\gamma\to^+0}\mathfrak{f}_\gamma(t)=\mathfrak{f}^{\mathbf{c}}_1(t), \tag{288}$$

$$\mathfrak{f}_1(t)=\lim_{\gamma\to^-1}\mathfrak{f}_\gamma(t)=\mathfrak{f}^{\mathbf{c}}_0(t). \tag{289}$$

Under restriction to $L_1(\mathcal{A})^+_1$, $D_1(\omega,\phi)$ becomes the ***Wald–Good–Kullback–Leibler distance*** [755, 284, 418, 416] (cf. [285, 51])

$$D_1|_{L_1(\mathcal{A})^+_1}(\omega,\phi)=\left\{\begin{array}{ll}\int\mu_\omega\log\frac{\mu_\omega}{\nu_\phi} & :\mu_\omega\ll\nu_\phi\\ +\infty & :\text{otherwise,}\end{array}\right. \tag{290}$$

which is a Csiszár–Morimoto $\mathfrak{f}$-distance with $\mathfrak{f}(\lambda)=\lambda\log(\lambda)$. More generally, (287) turns at $L_1(\mathcal{A})^+_1$ to

$$\mathfrak{f}^{\mathbf{c}}_\gamma|_{L_1(\mathcal{A})^+_1}(t)=\left\{\begin{array}{ll}\frac{1}{\gamma(1-\gamma)}(1-t^{1-\gamma}) & :\gamma\in]0,1[\\ t\log t & :\gamma=0\\ -\log t & :\gamma=1.\end{array}\right. \tag{291}$$

All above properties hold for the domain of γ extended from $[0,1]$ to $\mathbb{R}$ with the conditions satisfied for $\gamma\in]0,1[$ extending to $\gamma\in\mathbb{R}\setminus\{0,1\}$. Nevertheless, we will consider this extension separately.

The Liese–Vajda γ-distances are generalised Brègman distances for $\gamma\in]0,1[$ (see below), while for $\gamma\in\{0,1\}$ and $\dim(L_1(\mathcal{A}))=:n<\infty$ they are standard Brègman distances in the sense of (B_5) and (269) with $X=\mathbb{R}^n$ and $\Psi_{\gamma=1}(x)=\sum_{i=1}^n(x_i\log(x_i)-x_i+1)$.

Amari [19] has shown that the Liese–Vajda γ-distances with $\gamma\in\mathbb{R}$ can be characterised in the finite dimensional case as a unique class of standard Brègman distances that are monotone under coarse grainings.[28] Csiszár [182] (see also [557, 308]) has shown that under restriction to $L_1(\mathcal{A})^+_1$ the uniqueness result is stronger, characterising the pair $\{D_1|_{L_1(\mathcal{A})^+_1},D_0|_{L_1(\mathcal{A})^+_1}\}$. So far no corresponding characterisation results for the noncommutative case are known.[29]

[28] The assumption of *decomposability* used in Amari's proof is a discrete version of (270), so, together with (268), it amounts to a choice of a specific dual coordinate system.

[29] However, one should note Donald's [229] characterisation of Donald's distance, which coincides with $D_1|_{\mathcal{N}^+_{\star1}}$ at least for injective W^*-algebras, as well as Petz's [574] characterisation of $D_1|_{\mathcal{N}^+_{\star1}}$ (discussed below), which holds for injective W^*-algebras too. See also [571, 550] for other related results.

Consider the γ-***embedding*** functions on $\mathcal{N}_\star^+$ valued in $L_{1/\gamma}(\mathcal{N})^+$ spaces:

$$\ell_\gamma : \mathcal{N}_\star^+ \ni \omega \mapsto \ell_\gamma(\omega) := \frac{\omega^\gamma}{\gamma} \in L_{1/\gamma}(\mathcal{N}), \tag{292}$$

with $\gamma \in]0,1]$. These functions arise as restrictions of

$$\tilde{\ell}_\gamma : \mathcal{N}_\star \ni \omega \mapsto \tilde{\ell}_\gamma(\omega) := \frac{u|\omega|^\gamma}{\gamma} \in L_{1/\gamma}(\mathcal{N}), \tag{293}$$

which are bijections due to uniqueness of the polar decomposition $\omega = |\omega|(\cdot\, u)$. In particular, $\tilde{\ell}_{1/2}$ maps bijectively $\mathcal{N}_\star$ onto Hilbert space $L_2(\mathcal{N})$. The special case of the function (292) was introduced by Nagaoka and Amari [519, 524] in commutative finite dimensional setting,

$$\ell_\gamma : \mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \ni p(\chi) \mapsto \ell_\gamma(p(\chi)) := \begin{cases} \frac{1}{\gamma} p(\chi)^\gamma & : \gamma \in]0,1] \\ \log p(\chi) & : \gamma = 0 \end{cases} \in L_{1/\gamma}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+. \tag{294}$$

Since then it became a standard tool of information geometry theory. However, the Nagaoka–Amari formulation (294), as well as its noncommutative generalisations [313, 23],

$$\ell_\gamma : \mathfrak{G}_1(\mathcal{H})^+ \cong \mathfrak{B}(\mathcal{H})_\star^+ \ni \rho \mapsto \ell_\gamma(\rho) := \begin{cases} \frac{1}{\gamma}\rho^\gamma & : \gamma \neq 0 \\ \log \rho & : \gamma = 0 \end{cases} \in L_{1/\gamma}(\mathfrak{B}(\mathcal{H}), \mathrm{tr})^+ = \mathfrak{G}_{1/\gamma}(\mathcal{H})^+, \tag{295}$$

and [360]

$$\ell_\gamma^\psi : \mathcal{N}_\star^+ \ni \omega \mapsto \ell_\gamma^\psi(\omega) := \frac{\Delta_{\omega,\psi}^\gamma}{\gamma} \in L_{1/\gamma}(\mathcal{N}, \psi) \;\text{ for } \gamma \in]0,1[, \tag{296}$$

use γ-powers of densities (Radon–Nikodým quotients) with respect to a *fixed* reference measure $\tilde{\mu}$, trace tr, or weight $\psi \in \mathcal{W}_0(\mathcal{N})$, respectively. This restricts the generality of formulation. An important attempt to solve this problem in the commutative case was made by Zhu [788, 792], who considered the spaces of measures constructed through an equivalence relation based on γ-powers of Radon–Nikodým quotients, but without fixing any particular reference measure (hence, without passing to densities). However, his work remained unfinished and widely unknown, and it covered only the commutative case. The embeddings (292) solve these problems in the noncommutative case, while our construction of canonical $L_p(\mathcal{A})$ spaces over mcb-algebras $\mathcal{A}$, provided in Section 2.1, solves this problem in the commutative case (our definition of (283) already incorporates this solution).

The most general quantum distance that that has been known so far to be a standard Brègman distance that is monotone under coarse grainings is the ***Jenčová–Ojima*** γ-***distance*** [357, 360, 551]

$$D_\gamma(\omega, \phi) := \begin{cases} \frac{\gamma\omega(\mathbb{I}) + (1-\gamma)\phi(\mathbb{I}) - [\![\Delta_{\omega,\psi}^\gamma, \Delta_{\phi,\psi}^{1-\gamma}]\!]_\psi}{\gamma(1-\gamma)} & : \omega \ll \phi \\ +\infty & : \text{otherwise}, \end{cases} \tag{297}$$

where $\gamma \in]0,1[$, $\psi \in \mathcal{W}_0(\mathcal{N})$ is an arbitrary reference functional, $[\![\cdot,\cdot]\!]_\psi$ is the Banach space duality pairing between the Araki–Masuda noncommutative $L_{1/\gamma}(\mathcal{N}, \psi)$ and $L_{1/(1-\gamma)}(\mathcal{N}, \psi)$ spaces (see [42] or [404]). However, the Jenčová–Ojima distance is not a canonical noncommutative generalisation of the Liese–Vajda distance. The construction of the former is dependent on the choice of fixed reference weight ψ, while the latter does not depend on any additional measure. (Nevertheless, the *values* taken by the Jenčová–Ojima distance are independent of the choice of ψ.) Using the Falcone–Takesaki theory (see Section 2.2) we can make the reference-independent approach valid in all cases, including the infinite dimensional noncommutative one [402, 403].

Definition 3.1. *Given* $\gamma \in [0,1]$*, a* ***quantum*** γ-***distance*** *is a map*

$$D_\gamma : \mathcal{N}_\star^+ \times \mathcal{N}_\star^+ \ni (\omega, \phi) \mapsto D_\gamma(\omega, \phi) \in [0, \infty] \tag{298}$$

such that

$$D_\gamma(\omega,\phi) := \left\{ \begin{array}{ll} \int \frac{1}{\gamma(1-\gamma)} \left(\gamma\omega + (1-\gamma)\phi - \omega^\gamma\phi^{1-\gamma}\right) & : \gamma \in]0,1[,\ \omega \ll \phi \\ \int \lim_{\tilde{\gamma}\to^{\pm}\gamma} \frac{1}{\tilde{\gamma}(1-\tilde{\gamma})} \left(\tilde{\gamma}\omega + (1-\tilde{\gamma})\phi - \omega^{\tilde{\gamma}}\phi^{1-\tilde{\gamma}}\right) & : \gamma \in \{0,1\},\ \omega \ll \phi \\ +\infty & : \textit{otherwise}, \end{array} \right. \tag{299}$$

where the right limit, $\tilde{\gamma} \to^{+} \gamma$, is considered for $\gamma = 0$, and the left limit, $\tilde{\gamma} \to^{-} \gamma$, is considered for $\gamma = 1$.

Remark 3.2. The mathematical form of quantum γ-distance (299) exhibits strong formal similarity with its commutative special case (283) due to the canonical character of the Falcone–Takesaki construction, as well as the canonical character of the construction of $L_p(\mathcal{A})$ spaces provided in Section 2.1. Whenever required, the family (299) can be extended to the range $\gamma \in \mathbb{R}$ with the condition $\gamma \in]0,1[$ replaced by $\gamma \in \mathbb{R}\setminus\{0,1\}$, using the fact that (133) is well defined for any $\gamma > 0$, and *defining* $D_\gamma(\phi,\omega)$ for $\gamma < 0$ as $D_{1-\gamma}(\omega,\phi)$.

Proposition 3.3. *A quantum γ-distance* (299) *for $\gamma \in [0,1]$ is a Kosaki–Petz $\mathfrak{f}$-distance on $\mathcal{N}_\star^+$ with $\mathfrak{f}$ given by* (286)*.*

Proof. Applying (286) for $\gamma \in]0,1[$ to (185) for $\omega \ll \phi$ and using identity (133), we obtain

$$\begin{aligned} D_{\mathfrak{f}_\gamma}(\omega,\phi) &= \left\langle \xi_\pi(\phi), \left(\frac{1}{\gamma} + \frac{1}{1-\gamma}\Delta_{\omega,\phi} - \frac{1}{\gamma(1-\gamma)}\Delta^\gamma_{\omega,\phi}\right)\xi_\pi(\phi)\right\rangle_{\mathcal{H}} \\ &= \frac{1}{\gamma}\phi(\mathbb{I}) + \frac{1}{1-\gamma}\omega(\mathbb{I}) - \frac{1}{\gamma(1-\gamma)}\int \omega^\gamma\phi^{1-\gamma} \\ &= D_\gamma(\omega,\phi). \end{aligned} \tag{300}$$

We have also used the identity $\Delta^{1/2}_{\omega,\phi}\xi_\pi(\phi) = \mathrm{supp}(\phi)\xi_\pi(\phi)$, which holds for any $\phi,\omega \in \mathcal{N}_\star^+$. Using (288)-(289), we obtain $D_{\mathfrak{f}_\gamma}(\omega,\phi) = D_\gamma(\omega,\phi)$ also for $\gamma \in \{0,1\}$. □

Corollary 3.4. *From the above proof it follows that, for $\gamma \in \{0,1\}$,* (299) *can be written explicitly as*

$$\begin{aligned} D_0(\omega,\phi) &= \langle \xi_\pi(\phi), (-\log(\Delta_{\omega,\phi}) + \Delta_{\omega,\phi} - \mathbb{I})\rangle_{\mathcal{H}} \\ &= (\omega-\phi)(\mathbb{I}) - \langle \xi_\pi(\phi), \log(\Delta_{\omega,\phi})\xi_\pi(\phi)\rangle_{\mathcal{H}} \end{aligned} \tag{301}$$

and

$$\begin{aligned} D_1(\omega,\phi) &= \langle \xi_\pi(\phi), (\Delta_{\omega,\phi}\log(\Delta_{\omega,\phi}) - \Delta_{\omega,\phi} + \mathbb{I})\,\xi_\pi(\phi)\rangle_{\mathcal{H}} \\ &= (\phi-\omega)(\mathbb{I}) + \langle \xi_\pi(\phi), (\Delta_{\omega,\phi}\log(\Delta_{\omega,\phi}))\,\xi_\pi(\phi)\rangle_{\mathcal{H}} \\ &= (\phi-\omega)(\mathbb{I}) + \langle \xi_\pi(\omega), \log(\Delta_{\omega,\phi})\xi_\pi(\omega)\rangle_{\mathcal{H}}. \end{aligned} \tag{302}$$

Hence,

$$\phi \ll \omega \ll \phi \;\;\Rightarrow\;\; D_\gamma(\omega,\phi) = D_{1-\gamma}(\phi,\omega) \quad \forall\gamma \in [0,1], \tag{303}$$

$$D_\gamma(\omega,\phi) = D_\gamma(\phi,\omega) \iff \gamma = \frac{1}{2}. \tag{304}$$

Remark 3.5. The special cases of the distance (299) are:

- the Jenčová–Ojima γ-distance (297) for $\gamma \in]0,1[$, and for any choice of a reference weight $\psi \in \mathcal{W}_0(\mathcal{N})$, which determines the representation of the Falcone–Takesaki $L_p(\mathcal{N})$ space for every $p \in [1,\infty]$ provided by means of an isometric isomorphism with the Araki–Masuda $L_p(\mathcal{N},\psi)$ space.

- the ***Hasegawa* γ*-distance*** [313][30]

$$D_\gamma|_{\mathcal{N}^+_{\star 1}}(\omega,\phi)=\frac{\tau(\rho_\omega-\rho^\gamma_\omega\rho^{1-\gamma}_\phi)}{\gamma(1-\gamma)}=\frac{\tau(\rho_\omega-\Delta^\gamma_{\omega,\phi}\rho_\phi)}{\gamma(1-\gamma)}=\frac{1}{\gamma(1-\gamma)}-\tau(\ell_\gamma(\rho_\omega)\ell_{1-\gamma}(\rho_\phi)), \tag{305}$$

for $\gamma\in]0,1[$, $\omega\ll\phi$, semi-finite $\mathcal{N}$, and $\rho_\phi,\rho_\omega\in L_1(\mathcal{N},\tau)^+$ defined as the Dye–Segal densities (101) of $\omega,\phi\in\mathcal{N}^+_{\star 1}$ with respect to a faithful normal semi-finite trace τ on $\mathcal{N}$, i.e. $\phi(\cdot)=\tau(\rho_\phi\,\cdot)$ and $\omega(\cdot)=\tau(\rho_\omega\,\cdot)$. The map $\ell_\gamma:L_1(\mathcal{N},\tau)^+\ni\phi\mapsto\gamma^{-1}\rho^\gamma_\phi\in L_{1/\gamma}(\mathcal{N},\tau)$ is a straightforward generalisation of (295) and a special case of (292). If $\mathcal{N}$ is a type I factor, then the standard representation of $\mathcal{N}$ on $\mathcal{H}$ is isomorphic to $\mathfrak{B}(\mathcal{H})$ as a von Neumann algebra and $\tau(\cdot)=\mathrm{tr}(\cdot)$, where tr is a standard normalised ($\tau(\mathbb{I})=1$) trace on $\mathfrak{B}(\mathcal{H})$.

- the ***Araki distance*** [36, 39, 40]

$$D_1|_{\mathcal{N}^+_{\star 1}}(\omega,\phi)=\begin{cases}-\langle\xi_\pi(\omega),\log(\Delta_{\phi,\omega})\xi_\pi(\omega)\rangle_{\mathcal{H}} & :\omega\ll\phi\\ +\infty & :\text{otherwise},\end{cases} \tag{306}$$

which is a Kosaki–Petz $\mathfrak{f}$-distance with an operator convex function $\mathfrak{f}(\lambda)=-\log\lambda$. The alternative definitions generalising WGKL distance to $\mathcal{N}^+_{\star 1}$, given by Uhlmann [726] and by Pusz and Woronowicz [598], were shown to be equal to (306) in [329] and [229], respectively. If $D_1|_{\mathcal{N}^+_{\star 1}}(\omega,\phi)<\infty$, then (306) takes the form [564, 566]

$$D_1|_{\mathcal{N}^+_{\star 1}}(\omega,\phi)=\begin{cases}\mathrm{i}\lim_{t\to+0}\frac{\omega}{t}\left([\phi:\omega]_t-\mathbb{I}\right) & :\ \omega\ll\phi\\ +\infty & :\ \text{otherwise}.\end{cases} \tag{307}$$

For a semi-finite $\mathcal{N}$, normal faithful semi-finite trace τ on $\mathcal{N}$ and ρ_ϕ and ρ_ω defined as in (305), the Araki distance (306) turns to the ***Umegaki distance*** [735, 736] (cf. also [36, 37])

$$\begin{aligned}D_1|_{\mathcal{N}^+_{\star 1}}(\omega,\phi)&=\tau(\rho_\omega(\log\rho_\omega-\log\rho_\phi))=\tau\left(\rho^{1/2}_\omega(\log\Delta_{\omega,\phi})\rho^{1/2}_\omega\right)\\&=\int_0^1\mathrm{d}\lambda\tau\left(\rho_\omega\frac{1}{\rho_\phi+\lambda\mathbb{I}}(\rho_\omega-\rho_\phi)\frac{1}{\rho_\omega+\lambda\mathbb{I}}\right)\end{aligned} \tag{308}$$

if $\omega\ll\phi$, and $D_1|_{\mathcal{N}^+_{\star 1}}(\omega,\phi)=+\infty$ otherwise.

- the Liese–Vajda γ-distance (283) for $\gamma\in[0,1]$, and commutative $\mathcal{N}$, such that $\mathcal{N}=L_\infty(\mathcal{A})$.[31]

- the ***Amari–Cressie–Read* γ*-distance***[32] [16, 174, 17, 453, 606]

$$D_\gamma|_{L_1(\mathcal{A})^+_1}(\omega,\phi)=\begin{cases}\frac{1}{\gamma(1-\gamma)}\int\left(\mu_\omega-\nu_\phi\left(\frac{\mu_\omega}{\nu_\phi}\right)^\gamma\right) & :\mu_\omega\ll\nu_\phi\\ +\infty & :\ \text{otherwise}\end{cases} \tag{310}$$

for $\gamma\in]0,1[$, commutative $\mathcal{N}$, and normalised measures ν_ϕ and μ_ω ($\int\mu_\omega=1=\int\nu_\phi$) on the mcb-algebra $\mathcal{A}$ associated with $\mathcal{N}$ by means of $L_\infty(\mathcal{A})=\mathcal{N}$. Consider the ***Kakutani–Hellinger***

[30] Here we have generalised the original definition given by Hasegawa in a way analogous to Umegaki's definition of $D_1|_{\mathcal{N}^+_{\star 1}}$ given in [735] and (308).

[31] Under extension of the domain of γ to $\mathbb{R}$, the $D_{\gamma=2}$ distance is a Csiszár–Morimoto $\mathfrak{f}$-distance with $\mathfrak{f}(\lambda)=\frac{1}{2}(\lambda-1)^2$, and coincides, up to multiplication by 2, with the χ^2 distance (see Section 3.1),

$$D_{\gamma=2}(\omega,\phi)=2\chi^2(\omega,\phi). \tag{309}$$

[32] The family (310) can be considered for $\gamma\in]0,1[$ replaced by $\gamma\in\mathbb{R}\setminus\{0,1\}$. This family corresponds bijectively, but is not equal, to the γ-distance families of Chernoff [156], Kraft [407], Rényi [612, 613], Pérez [561], Havrda–Chárvat [320], Linhard–Nielsen [461], and Tsallis [722] (for a review with calculations, see e.g. [175]). One should note that, in particular, the Bhattacharyya, Chernoff, and Rényi distances *do not* belong to the class of the Csiszár–Morimoto $\mathfrak{f}$-distances [453].

distance [372, 321, 490, 491][33], defined as a square root of the Csiszár–Morimoto $\mathfrak{f}$-distance with $\mathfrak{f}(\lambda) = (1 - \sqrt{\lambda})^2$,

$$D_{\mathrm{KH}}(\omega, \phi) := \sqrt{\int \left(\sqrt{\nu_\phi} - \sqrt{\mu_\omega}\right)^2}. \tag{311}$$

The case $\gamma = 1/2$ satisfies

$$D_{\mathrm{KH}}|_{L_1(\mathcal{A})_1^+}(\omega, \phi) = \sqrt{\frac{1}{4} D_{1/2}|_{L_1(\mathcal{A})_1^+}(\omega, \phi)} \tag{312}$$

and allows to define the ***Bhattacharyya distance*** [74, 73, 75, 76]

$$D_{\mathrm{B}}|_{L_1(\mathcal{A})_1^+}(\omega, \phi) := 4 - 4 D_{1/2}|_{L_1(\mathcal{A})_1^+}(\omega, \phi). \tag{313}$$

If a representation of $\mathcal{A}$ in terms of some $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\upsilon})$ is given, with $\tilde{\mu}_\omega \ll \tilde{\upsilon}$ and $\tilde{\nu}_\phi \ll \tilde{\upsilon}$ such that $p_\omega := \tilde{\mu}_\omega / \tilde{\upsilon}$ and $q_\omega := \tilde{\nu}_\omega / \tilde{\upsilon}$, then (312) and (313) imply

$$\begin{aligned} D_{KH}|_{L_1(\mathcal{A})_1^+}(\omega, \phi) &= \sqrt{\frac{1}{2} \int \tilde{\upsilon} (\sqrt{p_\omega} - \sqrt{q_\phi})^2} = \sqrt{1 - \int \tilde{\upsilon} \sqrt{p_\omega q_\phi}} \\ &= \frac{1}{\sqrt{2}} \|\sqrt{p_\omega} - \sqrt{q_\phi}\|_{L_2(\mathcal{X}, \mho(\mathcal{X}), \tilde{\upsilon})} = \sqrt{1 - D_B|_{L_1(\mathcal{A})_1^+}(\omega, \phi)}. \end{aligned} \tag{314}$$

- the WGKL distance (290), for commutative $\mathcal{N} = L_\infty(\mathcal{A})$, and $\int \mu_\omega = 1 = \int \nu_\phi$. (For an explicit derivation of the WGKL distance from the Araki distance for $\mathcal{N} = L_\infty(\mathcal{X}, \mho(\mathcal{X}), \tilde{\upsilon})$ see [329].)

Proposition 3.6. *If $\gamma \in]0, 1[$, then quantum γ-distance (299) is both a dualistic Brègman distance (263) and a standard Brègman distance (266) on $\mathcal{N}_\star^+$, with a dual coordinate system $(\ell_\gamma, \ell_{1-\gamma})$ given by (292), with convex proper function*

$$\Psi_\gamma : L_{1/\gamma}(\mathcal{N}) \ni x \mapsto \Psi_\gamma(x) := \frac{1}{1-\gamma} \int (\gamma x)^{1/\gamma} = \frac{1}{1-\gamma} \|\gamma x\|_{1/\gamma}^{1/\gamma} \in [0, +\infty[, \tag{315}$$

with a dualiser

$$\mathbf{L}_{\Psi_\gamma} := \tilde{\ell}_{1-\gamma} \circ \tilde{\ell}_\gamma^{-1} : L_{1/\gamma}(\mathcal{N}) \ni \frac{1}{\gamma} u |\phi|^\gamma \mapsto \frac{1}{1-\gamma} u |\phi|^{1-\gamma} \in L_{1/(1-\gamma)}(\mathcal{N}) \tag{316}$$

and with a Brègman functional, in the sense of ($B_{\bar{D}}$) and (B_4),

$$L_{1/\gamma}(\mathcal{N}) \times L_{1/\gamma}(\mathcal{N}) \ni (x, y) \mapsto \bar{D}_{\Psi_\gamma}(x, y) = \Psi_\gamma(x) + \Psi_{1-\gamma}(\mathbf{L}_{\Psi_\gamma}(y)) - \mathrm{re}\, [\![x, \mathbf{L}_{\Psi_\gamma}(y)]\!]_{L_{1/\gamma}(\mathcal{N}) \times L_{1/(1-\gamma)}(\mathcal{N})}. \tag{317}$$

Proof. Our method of proof will be based on the approach of [360] (which in turn used some of the ideas introduced in [276]).

The embeddings ℓ_γ defined by (292) allow to construct the real valued functional on $\mathcal{N}_\star^+$ using the duality (119),

$$\mathcal{N}_\star^+ \times \mathcal{N}_\star^+ \ni (\omega, \phi) \mapsto \int \ell_\gamma(\omega) \ell_{1-\gamma}(\phi) = [\![\ell_\gamma(\omega), \ell_{1-\gamma}(\phi)]\!]_{L_{1/\gamma}(\mathcal{N}) \times L_{1/(1-\gamma)}} \in \mathbb{R}. \tag{318}$$

In these terms, D_γ defined in (299) for $\gamma \in]0, 1[$ is equal to

$$D_\gamma(\omega, \phi) = \int \left(\frac{\omega}{1-\gamma} + \frac{\phi}{\gamma} - \ell_\gamma(\omega) \ell_{1-\gamma}(\phi) \right) = \frac{\omega(\mathbb{I})}{1-\gamma} + \frac{\phi(\mathbb{I})}{\gamma} - [\![\ell_\gamma(\omega), \ell_{1-\gamma}(\phi)]\!]_\gamma, \tag{319}$$

[33] As pointed in [438], the reference to [321] is traditional, but quite irrelevant. The referenced paper contains only the integrals of the form $\int \frac{\tilde{\mu}_1 \tilde{\mu}_2}{\tilde{\upsilon}}$ for $\tilde{\mu}_1 \ll \tilde{\upsilon}$ and $\tilde{\mu}_2 \ll \tilde{\upsilon}$.

where we have simplified the notation by $[\![\cdot,\cdot]\!]_\gamma := [\![\cdot,\cdot]\!]_{L_{1/\gamma}(\mathcal{N})\times L_{1/(1-\gamma)}(\mathcal{N})}$.

We begin by proving that that a function $\mathbf{L}_{\Psi_\gamma}$ is a homeomorphisms in the corresponding norm topologies. Its bijectivity follows from the bijectivity of $\tilde{\ell}_\gamma$. For $\phi \in \mathcal{N}_\star$ denote its unique polar decomposition as $|\phi|(\,\cdot\, u)$. From (118) it follows that

$$\|u|\phi|^\gamma\|_{1/\gamma} = (|\phi|(\mathbb{I}))^\gamma\,, \tag{320}$$

so

$$\begin{aligned}\|\gamma x\|_{1/\gamma}^{1/\gamma} &:= \left\|(1-\gamma)\mathbf{L}_{\Psi_\gamma}(x)\right\|_{1/(1-\gamma)}^{1/(1-\gamma)} = \int |\phi|\mathbb{I} = \int |\phi|^\gamma|\phi|^{1-\gamma}\mathrm{supp}(\phi)\\ &= \int u|\phi|^\gamma|\phi|^{1-\gamma}u^* = \int u|\phi|^\gamma\left(u|\phi|^{1-\gamma}\right)^*\\ &= \gamma(1-\gamma)\left[\!\left[x,(\mathbf{L}_{\Psi_\gamma}(x))^*\right]\!\right]_\gamma.\end{aligned} \tag{321}$$

For a Banach space X let $v_{x/\|x\|}$ denote a unique point on a unit sphere in $X^\star$ such that

$$\left[\!\left[x, v_{x/\|x\|}\right]\!\right]_{X\times X^\star} = \|x\|_{X^\star}. \tag{322}$$

According to [195], if X is uniformly convex and $\|\cdot\|_X$ is Fréchet differentiable, then a map

$$F_v : \begin{cases} X\setminus\{0\}\ni x & \mapsto \|x\|_X v_{x/\|x\|}\in X^\star\setminus\{0\}\\ X\ni 0 & \mapsto 0\in X^\star\end{cases} \tag{323}$$

is a homeomorphism in the norm topologies of X and $X^\star$. The function

$$v_\gamma(x) := \|\gamma x\|_{1/\gamma}^{1-1/\gamma}(1-\gamma)(\mathbf{L}_{\Psi_\gamma}(x))^* \tag{324}$$

satisfies

$$\begin{aligned}[\![x, v_\gamma(x)]\!]_\gamma &= \|\gamma x\|_{1/\gamma}^{1-1/\gamma}(1-\gamma)\left[\!\left[x,(\mathbf{L}_{\Psi_\gamma}(x))^*\right]\!\right]_\gamma\\ &= \|\gamma x\|_{1/\gamma}^{1-1/\gamma}(1-\gamma)\|\gamma x\|_{1/\gamma}^{1/\gamma}\gamma^{-1}(1-\gamma)^{-1}\\ &= \|x\|_{1/\gamma},\end{aligned} \tag{325}$$

hence $v_\gamma(x) = v_{x/\|x\|}$ for $X = L_{1/\gamma}(\mathcal{N})$. From (321) it follows that $\mathbf{L}_{\Psi_\gamma}(x)$ is continuous at 0. From uniform convexity and uniform Fréchet differentiability of $L_{1/\gamma}(\mathcal{N})$ for $\gamma\in]0,1[$ it follows that for $x\in L_{1/\gamma}(\mathcal{N})\setminus\{0\}$ the function F_{v_γ} reads

$$F_{v_\gamma}(x) = \|x\|_{1/\gamma}v_\gamma(x) = (1-\gamma)\gamma^{1-1/\gamma}\|x\|_{1/\gamma}^{2-1/\gamma}(\mathbf{L}_{\Psi_\gamma}(x))^*, \tag{326}$$

which implies that $\mathbf{L}_{\Psi_\gamma}$ is also a homeomorphism.

Next, we will prove that Ψ_γ is Fréchet differentiable, with

$$(\mathfrak{D}^{\mathrm{F}}_x\Psi_\gamma)(y) = \mathrm{re}\,\left[\!\left[y,\mathbf{L}_{\Psi_\gamma}(x)\right]\!\right]_\gamma\quad\forall x\in L_{1/\gamma}(\mathcal{N}) \tag{327}$$

and

$$\Psi_\gamma(x) + \Psi_{1-\gamma}(\mathbf{L}_{\Psi_\gamma}(x)) - \mathrm{re}\,\left[\!\left[x,\mathbf{L}_{\Psi_\gamma}(x)\right]\!\right]_\gamma = 0\quad\forall x\in L_{1/\gamma}(\mathcal{N}). \tag{328}$$

If a Banach space X is Gâteaux differentiable except $0\in X$, then

$$\left[\!\left[y,\mathfrak{D}^{\mathrm{G}}_x\|\cdot\|\right]\!\right]_{X\times X^\star} = \mathrm{re}\,\left[\!\left[y, v_{x/\|x\|}\right]\!\right]_{X\times X^\star}. \tag{329}$$

From the uniform Fréchet differentiability of $L_{1/\gamma}(\mathcal{N})$ it follows that $\|\cdot\|_{1/\gamma}$ is Fréchet differentiable at any $x\in L_{1/\gamma}(\mathcal{N})\setminus\{0\}$, and

$$(\mathfrak{D}^{\mathrm{F}}_x\|\cdot\|_{1/\gamma})(y) = \mathrm{re}\,[\![y, v_\gamma]\!]_\gamma\quad\forall y\in L_{1/\gamma}(\mathcal{N}), \tag{330}$$

so

$$(\mathfrak{D}^{\mathrm{F}}_x\Psi_\gamma)(y) = \left(\mathfrak{D}^{\mathrm{F}}\left(\frac{1}{1-\gamma}\|\gamma x\|^{1/\gamma}_{1/\gamma}\right)\right)(y) = \left(\frac{1}{1-\gamma}\|\gamma x\|^{1/\gamma-1}_{1/\gamma}\mathfrak{D}^{\mathrm{F}}\|x\|_{1/\gamma}\right)(y)$$
$$= \mathrm{re}\,\left[\!\left[y, \frac{1}{1-\gamma}\|\gamma x\|^{1/\gamma-1}_{1/\gamma}\|\gamma x\|^{1-1/\gamma}_{1/\gamma}(1-\gamma)(\mathbf{L}_{\Psi_\gamma}(x))^*\right]\!\right]_\gamma$$
$$= \mathrm{re}\,\left[\!\left[y, \mathbf{L}_{\Psi_\gamma}(x)\right]\!\right]_\gamma. \tag{331}$$

The function $\|\gamma x\|^{1/\gamma}_{1/\gamma}$ is also Fréchet differentiable at $x=0$, which implies

$$(\mathfrak{D}^{\mathrm{F}}_0\Psi_\gamma)(y) = 0 = \mathrm{re}\,\left[\!\left[y, \mathbf{L}_{\Psi_\gamma}(0)\right]\!\right]_\gamma. \tag{332}$$

This gives (327). The equation (328) follows as straightforward calculation. Note that (328) is just $\bar{D}_{\Psi_\gamma}(x,x)=0$ for $\bar{D}_{\Psi_\gamma}$ given by (317). From the fact that (317) satisfies (B$_4$), it follows that $\bar{D}_{\Psi_\gamma}(x,y)\geq 0$. Moreover, from Fréchet differentiability and continuity of $\Psi_{1-\gamma}$ on all $L_{1/(1-\gamma)}(\mathcal{N})$ and reflexivity of $L_{1/\gamma}(\mathcal{N})$ spaces it follows that Ψ_γ is essentially strictly convex, hence, due to (225), $\bar{D}_{\Psi_\gamma}(x,y)=0 \iff x=y$. This implies that the equation (328) is a unique solution of the variational problem

$$\Psi_{1-\gamma}(\mathbf{L}_{\Psi_\gamma}(x)) = \sup_{y\in L_{1/\gamma}(\mathcal{N})}\left\{\mathrm{re}\,\left[\!\left[y, \mathbf{L}_{\Psi_\gamma}(x)\right]\!\right]_\gamma - \Psi_\gamma(y)\right\}, \tag{333}$$

because

$$y\neq x \;\;\Rightarrow\;\; \Psi_\gamma(y) + \Psi_{1-\gamma}(\mathbf{L}_{\Psi_\gamma}(x)) - \mathrm{re}\,\left[\!\left[y, \mathbf{L}_{\Psi_\gamma}(x)\right]\!\right]_\gamma > 0, \tag{334}$$
$$\Psi_{1-\gamma}(\mathbf{L}_{\Psi_\gamma}(x)) > \mathrm{re}\,\left[\!\left[y, \mathbf{L}_{\Psi_\gamma}(x)\right]\!\right]_\gamma - \Psi_\gamma(y). \tag{335}$$

Comparing (333) with (193), we see that

$$\Psi_{1-\gamma} = \Psi^{\mathbf{L}}_\gamma, \tag{336}$$

with respect to the duality $[\![\cdot,\cdot]\!]_\gamma$.

If X is a Banach space and $f: X\to\mathbb{R}$ is norm continuous and convex function, then f is Gâteaux differentiable iff $\partial f(x)=\{*\}\;\forall x\in X$. The norm continuity and Fréchet differentiability of Ψ_γ on $L_{1/\gamma}(\mathcal{N})$ implies that

$$\partial\Psi_\gamma(x) = \{*\} = \mathfrak{D}^{\mathrm{F}}_x\Psi_\gamma, \tag{337}$$

so

$$\mathbf{L}_{\Psi_\gamma}(y)\in\partial\Psi_\gamma(x) \quad\iff\quad x=y \;\;\forall x,y\in\mathrm{efd}(\partial\Psi_\gamma). \tag{338}$$

Hence, $(L_{1/\gamma}(\mathcal{N}), L_{1/(1-\gamma)}(\mathcal{N}), \mathbf{L}_{\Psi_\gamma})$ is a generalised Legendre transform, and $D_\gamma(\omega,\phi)$ is a dualistic Brègman distance of the form

$$D_{\Psi_\gamma}(\omega,\phi) = \Psi_\gamma(\ell_\gamma(\omega)) + \Psi_{1-\gamma}(\ell_{1-\gamma}(\phi)) - [\![\ell_\gamma(\omega),\ell_{1-\gamma}(\phi)]\!]_{L_{1/\gamma}(\mathcal{N})\times L_{1/(1-\gamma)(\mathcal{N})}} \tag{339}$$

with $\Psi_\gamma(\ell_\gamma(\omega)) = \frac{1}{1-\gamma}(\mathbb{I})$. □

Proposition 3.7. *If $\gamma\in\,]0,1[$, then $D_\gamma(\omega,\phi)$ satisfies the generalised cosine equation*

$$D_\gamma(\omega,\phi) + D_\gamma(\phi,\psi) = D_\gamma(\omega,\psi) + \int(\ell_\gamma(\omega)-\ell_\gamma(\phi))\,(\ell_{1-\gamma}(\psi)-\ell_{1-\gamma}(\phi)). \tag{340}$$

In finite dimensional setting (340) holds also for $\gamma\in\{0,1\}$, with ℓ_γ given by (295).

Proof. Straightforward calculation based on equations (319) and (273). □

Corollary 3.8. *The equation (340) is an instance of the 'standard' generalised cosine equation (274) applied to $\bar{D}_{\Psi_\gamma}$ given by (317), while the equation (303) follows from the 'representation-index duality' equation*[34]

$$\bar{D}_{\Psi_\gamma}(x,y) = \bar{D}_{\Psi_{1-\gamma}}(\mathbf{L}_{\Psi_\gamma}(y), \mathbf{L}_{\Psi_\gamma}(x)), \tag{341}$$

where $x, y \in L_{1/\gamma}(\mathcal{N})$. For $\gamma = 1/2$ the $L_{1/\gamma}(\mathcal{N})$ space becomes a Hilbert space $\mathcal{H}$ (see Section 2.2), the generalised Brègman functional $\bar{D}_{\Psi_\gamma}$ becomes the norm distance on it,

$$\bar{D}_{\Psi_{1/2}}(x,y) = \frac{1}{2}\|x-y\|^2_{\mathcal{H}}, \tag{342}$$

so the generalised cosine equation for $\bar{D}_{\Psi_\gamma}$ turns to the cosine equation in Hilbert space $\mathcal{H}$,

$$\|x-y\|^2_{\mathcal{H}} + \|y-z\|^2_{\mathcal{H}} = \|x-z\|^2_{\mathcal{H}} + 2\left\langle x-y, z-y\right\rangle_{\mathcal{H}}. \tag{343}$$

Remark 3.9. From the fact that (299) is a Kosaki–Petz $\mathfrak{f}$-distance it follows that it has the following properties [397, 565, 550, 360]:

1) $D_\gamma(\omega,\phi) \geq D_\gamma(T_\star(\omega), T_\star(\phi))$,
2) D_γ is jointly convex on $\mathcal{N}^+_\star \times \mathcal{N}^+_\star$,
3) for $\gamma \in [0,1]$, D_γ is lower semi-continuous on $\mathcal{N}^+_\star \times \mathcal{N}^+_{\star 0}$ endowed with the product of norm topologies, while for $\gamma \in \{0,1\}$ it is also lower semi-continuous on $\mathcal{N}^+_\star \times \mathcal{N}^+_\star$ endowed with the product of weak-$\star$ topologies.

The joint convexity of the Umegaki distance $D_1|_{\mathcal{N}^+_{\star 1}}$ was proved by Lindblad [456] (it can be derived from Lieb's theorem [450], cf. [72], and the converse is also true [721], see also [29]), while the generalisation of this proof to the Araki distance $D_1|_{\mathcal{N}^+_{\star 1}}$ over preduals of arbitrary W^*-algebras was provided in [40, 598, 397]. For $D_1|_{\mathcal{N}^+_{\star 1}}$, the markovian monotonicity (property 1) was shown for type I factors by Lindblad [456, 457] (using the subadditivity of $D_1|_{\mathcal{N}^+_{\star 1}}$, proved in [451]), was extended to some other W^*-algebras by Araki [40] and was proved in general case by Uhlmann [726] (see also [598]), while the weak-$\star$ lower semi-continuity (property 3) was proved for type I factors by Wehrl [758], for some other W^*-algebras by Araki [40], and the complete proof was given independently by Kosaki [400] and Donald [229]. Moreover, $D_1|_{\mathcal{N}^+_{\star 1}}(\omega,\phi)$ satisfies also

(i) $\lambda_1, \lambda_2 > 0 \Rightarrow D_1|_{\mathcal{N}^+_{\star 1}}(\lambda_1\omega, \lambda_2\phi) = \lambda_1 D_1|_{\mathcal{N}^+_{\star 1}}(\omega,\phi) + \lambda_1\omega(\mathbb{I})(\log\lambda_1 - \log\lambda_2)\;\; \forall\lambda_1,\lambda_2 > 0$, [40]

(ii) $D_1|_{\mathcal{N}^+_{\star 1}}(\omega,\phi) \geq \omega(\mathbb{I})(\log\omega(\mathbb{I}) - \log\phi(\mathbb{I}))$ [598],

(iii) $\phi_1 \leq \phi_2 \Rightarrow D_1|_{\mathcal{N}^+_{\star 1}}(\omega,\phi_1) \geq D_1|_{\mathcal{N}^+_{\star 1}}(\omega,\phi_2)$ [40],

(iv) $\mathrm{supp}(\omega_1)\mathrm{supp}(\omega_2) = 0 \Rightarrow D_1|_{\mathcal{N}^+_{\star 1}}(\omega_1,\phi) + D_1|_{\mathcal{N}^+_{\star 1}}(\omega_2,\phi) = D_1|_{\mathcal{N}^+_{\star 1}}(\omega_1+\omega_2,\phi)$, [40]

(v) $\omega = \sum_{i=1}^n \omega_i \Rightarrow D_1|_{\mathcal{N}^+_{\star 1}}(\omega,\phi) + \sum_{i=1}^n D_1|_{\mathcal{N}^+_{\star 1}}(\omega_i,\omega) = \sum_{i=1}^n D_1|_{\mathcal{N}^+_{\star 1}}(\omega_i,\phi)\;\forall\phi,\omega_1,\ldots,\omega_n \in \mathcal{N}^+_\star$, [231, 41, 232, 570]

(vi) if $\mathcal{E} : \mathcal{N} \to \mathcal{N}_0 \subseteq \mathcal{N}$ is a conditional expectation of norm 1 that is ω-stable[35], then [328, 566, 570]

$$D_1|_{\mathcal{N}^+_{\star 1}}(\omega,\phi) = D_1|_{\mathcal{N}^+_{\star 1}}(\omega|_{\mathcal{N}_0}, \phi|_{\mathcal{N}_0}) + D_1|_{\mathcal{N}^+_{\star 1}}(\phi\circ\mathcal{E}, \phi), \tag{344}$$

[34] The finite dimensional commutative version of the equation (341), with a dualiser given by gradient, was discussed in [784].

[35] A conditional expectation $\mathcal{E} : \mathcal{N} \to \mathcal{N}_0 \subseteq \mathcal{N}$ is called ***ω-stable*** for $\omega \in \mathcal{N}^+_\star$ iff $\omega|_{\mathcal{N}_0} \circ \mathcal{E} = \omega$.

(vii) if $\mathcal{N} = \mathcal{N}_1 \oplus \mathcal{N}_2$ and if $\exists\lambda \in]0,1[$ such that for all $x \in \mathcal{N}_1$ and $y \in \mathcal{N}_2$

$$\phi(x \oplus y) = \lambda\phi_1(x) + (1-\lambda)\phi_2(y), \tag{345}$$
$$\omega(x \oplus y) = \lambda\omega_1(x) + (1-\lambda)\omega_2(y), \tag{346}$$

then

$$D_1|_{\mathcal{N}^+_{\star 1}}(\omega,\phi) = \lambda D_1|_{\mathcal{N}^+_{\star 1}}(\omega_1,\phi_1) + (1-\lambda) D_1|_{\mathcal{N}^+_{\star 1}}(\omega_2,\phi_2). \tag{347}$$

(viii) $D_1|_{\mathcal{N}^+_{\star 1}}$ is strictly convex in first variable if it is finite, [568, 231, 232]

(ix) $4\left(d_{L_1(\mathcal{N})}(\omega,\phi)\right)^2 = \|\omega - \phi\|^2_{\mathcal{N}_\star} \leq 2D_1|_{\mathcal{N}^+_{\star 1}}(\omega,\phi)$. [328][36]

Petz [574] characterised $D_1|_{\mathcal{N}^+_{\star 1}}$ as a unique distance on $\mathcal{N}^+_{\star 1}$ for injective[37] W^*-algebras $\mathcal{N}$ that satisfies the conditions 1), 3), (vi) and (vii).

Remark 3.10. The family (299) of quantum γ-distances provides a canonical infinite dimensional noncommutative generalisation of the family (283) of Liese–Vajda γ-distances, and generalises the family (297) of Jenčová–Ojima γ-distances in terms of canonical noncommutative $L_{1/\gamma}(\mathcal{N})$ spaces. These canonical properties, considered together with Propositions 3.3 and 3.6 suggest a quantum analogue of Amari's [19] characterisation of the Liese–Vajda γ-distances. Note that Amari's characterisation holds for $\gamma \in \mathbb{R}$. On the other hand hand, Proposition 5.1 strongly suggests that in quantum case the restriction to $\gamma \in [-1,2]$ is necessary. This leads us to:

Conjecture 3.11. *The family $D_\gamma(\omega,\phi)$ of quantum γ-distances defined by (299) for $\gamma \in [-1,2]$ is the unique family of quantum distances $D(\omega,\phi)$ on $\mathcal{N}^+_\star$ that satisfies the conditions:*

1) $D(\omega,\phi) \geq D(T_\star(\omega),T_\star(\phi)) \quad \forall\omega,\phi \in \mathcal{N}^+_\star \quad \forall T_\star \in \mathrm{Mark}_\star(\mathcal{N}^+_\star)$,

2) $\exists C \subseteq \mathcal{N}^+_\star \ \forall K \subseteq \mathcal{N}^+_\star \ \forall(\phi,\psi) \in K \times C$

$$\mathfrak{P}^D_K(\psi) = \{*\} \quad\Rightarrow\quad D(\phi,\psi) = D(\phi,\mathfrak{P}^D_K(\psi)) + D(\mathfrak{P}^D_K(\psi),\psi), \tag{348}$$

where $\mathfrak{P}^D_K(\psi) := \arg\inf_{\phi\in K}\{D(\phi,\psi)\}$. Moreover, under restriction from $\mathcal{N}^+_\star$ to $\mathcal{N}^+_{\star 1}$, the above conditions are satisfied only by $D_\gamma(\omega,\phi)$ for $\gamma \in \{0,1\}$.

Now let us turn to discussion of the selected results on existence, uniqueness, and properties of the projections $\mathfrak{P}^{D_\gamma}_C(\psi)$, defined as minimisers of quantum γ-distances for $\gamma \in [0,1]$.

The following results on entropic projections based on $D_1|_{\mathcal{N}^+_{\star 1}}$ were obtained by Donald [232] (cf. also [230, 41, 233]). If $\phi,\omega \in \mathcal{N}^+_{\star 1}$, $h \in \mathcal{N}^{\mathrm{ext}}$,[38]then the function

$$c(\omega,h) := \inf_{\phi\in\mathcal{N}^+_{\star 1}}\left\{D_1|_{\mathcal{N}^+_{\star 1}}(\phi,\omega) + \phi(h)\right\} \tag{355}$$

[36]For commutative $\mathcal{N}$ this inequality was established in [587, 177], and in this case it can be sharpened using higher orders of D_1, see [737, 720, 718] and [607] for a review and further improvements.

[37]The injective W^*-algebras were introduced in [466] and were shown in [167] to be equal to approximately finite dimensional W^*-algebras whenever they have a separable predual (for the remaining case there are no counterexamples known).

[38]Given a W^*-algebra $\mathcal{N}$, the ***extended positive cone*** $\mathcal{N}^{\mathrm{ext}}$ is defined as set of maps $m : \mathcal{N}^+_\star \to [0,\infty]$ such that for all $\phi,\psi \in \mathcal{N}^+_\star$ [299]

1) $m(\lambda\phi) = \lambda m(\phi) \ \ \forall\lambda \geq 0$,
2) $m(\phi+\psi) = m(\phi) + m(\psi)$,
3) m is weakly lower semi-continuous, that is,

$$\sup_\iota\{\omega_\iota(x)\} = \omega(x) \quad\Rightarrow\quad m(\omega) \leq \liminf_\iota\{m(\omega_\iota)\} \ \ \forall\omega,\omega_\iota \in \mathcal{N}^+_\star, \tag{349}$$

or, equivalently,

$$\text{the sets } \{\omega \in \mathcal{N}^+_\star \mid m(\omega) > \lambda\} \text{ are weakly open } \forall\lambda \in \mathbb{R}. \tag{350}$$

The set $\mathcal{N}^{\mathrm{ext}}$ can be considered as the 'set of normal weights on $\mathcal{N}_\star$'. It contains $\mathcal{N}^+$, and is closed under addition, multiplication by nonnegative scalars, and increasing limits of nets. For all $m_1,m_2 \in \mathcal{N}^{\mathrm{ext}}$, $x \in \mathcal{N}$, $\phi \in \mathcal{N}^+_\star$ and $\lambda \in \mathbb{R}^+$

is called a ***relative free energy***. We use here the simplified notation $\phi(h) := h(\phi)$, with

$$\phi(PhP) := \begin{cases} 0 & : \ \phi(P) = 0 \\ h\left(\frac{\phi(P \cdot P)}{\phi(P)}\right) & : \ \text{otherwise} \end{cases} \tag{356}$$

for any $P \in \mathrm{Proj}(\mathcal{N})$. If $c(\omega, h) < \infty$, then

$$\exists!\ \omega^h := \arg\inf_{\phi \in \mathcal{N}^+_{\star 1}} \left\{ D_1|_{\mathcal{N}^+_{\star 1}}(\phi, \omega) + \phi(h) \right\}, \tag{357}$$

and it satisfies

$$\omega^h \ll \omega \ \ (\text{with } \omega(h) < \infty \ \Rightarrow \ \mathrm{supp}(\omega^h) = \mathrm{supp}(\omega)), \tag{358}$$
$$\omega^h = \omega^{h+\lambda} = \omega^{PhP} \ \ \forall \lambda \in \mathbb{R}\ \forall P \in \mathrm{Proj}(\mathcal{N}) \text{ such that } P \geq \mathrm{supp}(\omega), \tag{359}$$
$$c(\omega, h) = D_1|_{\mathcal{N}^+_{\star 1}}(\omega^h, \omega) + \omega^h(h), \tag{360}$$
$$D_1|_{\mathcal{N}^+_{\star 1}}(\phi, \omega) + \phi(h) \geq c(\omega, h) + D_1|_{\mathcal{N}^+_{\star 1}}(\phi, \omega^h) \ \ \forall \phi \in \mathcal{N}^+_{\star 1}, \tag{361}$$
$$(\exists \lambda \in \mathbb{R} \ \ \phi \leq \lambda \omega^h) \ \Rightarrow \ D_1|_{\mathcal{N}^+_{\star 1}}(\omega, \phi) + \phi(h) = c(\omega, h) + D_1|_{\mathcal{N}^+_{\star 1}}(\phi, \omega^h) \ \ \forall \phi \in \mathcal{N}^+_{\star 1}, \tag{362}$$
$$\omega(h) < \infty \ \Rightarrow \ D_1|_{\mathcal{N}^+_{\star 1}}(\phi, \omega^h) + D_1|_{\mathcal{N}^+_{\star 1}}(\omega^h, \omega) = D_1|_{\mathcal{N}^+_{\star 1}}(\phi, \omega) + (\omega^h - \phi)(h) \ \ \forall \phi \in \mathcal{N}^+_{\star 1}. \tag{363}$$

Moreover, if $k \in \mathcal{N}^{\mathrm{sa}}$, then

$$(\omega^k)^h = \omega^{k+h}, \tag{364}$$
$$\omega(h) < \infty \ \Rightarrow \ c(\omega, k+h) = c(\omega, k) + c(\omega^k, h). \tag{365}$$

From $c(\omega, h) \leq \omega(h)$ it follows that $\omega(h) < \infty \Rightarrow \exists!\ \omega^h$. From the generalised cosine equation (363) it follows that $D_1|_{\mathcal{N}^+_{\star 1}}$ satisfies the generalised pythagorean equation

$$D_1|_{\mathcal{N}^+_{\star 1}}(\phi, \omega^h) + D_1|_{\mathcal{N}^+_{\star 1}}(\omega^h, \omega) = D_1|_{\mathcal{N}^+_{\star 1}}(\phi, \omega) \tag{366}$$

under the conditions $\omega \in \mathcal{N}^+_{\star 1}$, $h \in \mathcal{N}^{\mathrm{ext}}$, $\omega(h) < \infty$, and the orthogonality condition

$$\omega^h(h) = \phi(h). \tag{367}$$

The special case of equation (362), for $\omega \in \mathcal{N}^+_{\star 01}$ and $h \in \mathcal{N}^{\mathrm{sa}}$ (which removes the need for an assumption $\exists \lambda \in \mathbb{R}\ \phi \leq \lambda \omega^h$), was obtained by Araki in [40], while the special cases of (364)-(365), for $\omega \in \mathcal{N}^+_{\star 01}$ and $h \in \mathcal{N}^{\mathrm{sa}}$, were obtained by him in [33]. Donald [232] showed also that if $\omega, \psi \in \mathcal{N}^+_{\star 1}$ and $\exists \lambda \in \mathbb{R}\ \psi \leq \lambda \omega$, then there exists $h \in \mathcal{N}^{\mathrm{ext}}$ such that

$$\psi = \omega^h, \tag{368}$$
$$D_1|_{\mathcal{N}^+_{\star 1}}(\phi, \omega) + \phi(h) = D_1|_{\mathcal{N}^+_{\star 1}}(\phi, \psi) \ \ \forall \phi \in \mathcal{N}^+_{\star 1}, \tag{369}$$
$$h \geq -(\log \lambda)\mathbb{I}, \tag{370}$$
$$\psi, \omega \in \mathcal{N}^+_{\star 01} \ \Rightarrow \ h \in \mathcal{N}^{\mathrm{sa}}. \tag{371}$$

---

one defines

$$(\lambda m)(\phi) := \lambda m(\phi), \tag{351}$$
$$(m_1 + m_2)(\phi) := m_1(\phi) + m_2(\phi), \tag{352}$$
$$(x^* m x)(\phi) := m(\phi(x \cdot x^*)). \tag{353}$$

Every $m \in \mathcal{N}^{\mathrm{ext}}$ has a unique spectral decomposition

$$m(\phi) = \int_0^\infty \phi(P^m(\lambda)) + \infty \cdot \phi(P^m) \ \ \forall \phi \in \mathcal{N}^+_\star, \tag{354}$$

where $\{P^m(\lambda) \in \mathrm{Proj}(\mathcal{N}) \mid \lambda \in \mathbb{R}^+\}$ is an increasing family which is strongly-$\star$ continuous from the right, and $P^m = \mathbb{I} - \lim_{\lambda \to \infty} P^m(\lambda)$. If $\mathcal{N}$ is commutative and $\mathcal{N} \cong L_\infty(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ then $\mathcal{N}^{\mathrm{ext}} \cong L_0(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}; [0, +\infty])$.

In [33] a special case of this result was given, under an additional assumption that $\exists\lambda_2\in\mathbb{R}\ \omega\leq\lambda_2\psi$, which implies that $h\in\mathcal{N}^{\mathrm{sa}}$, and (368)-(370), as well as $(\log\lambda_2)\mathbb{I}\geq k$, hold. Donald's relative entropic projections ω^h of normalised quantum states ω by means of $h\in\mathcal{N}^{\mathrm{ext}}$ provide a direct generalisation of Araki's [33, 38, 40] perturbations of faithful normalised quantum states by means of $h\in\mathcal{N}^{\mathrm{sa}}$. This means that, for $(\omega,h)\in\mathcal{N}^+_{\star01}\times\mathcal{N}^{\mathrm{sa}}$, (357) can be always determined perturbatively by [232]

$$\begin{aligned}\left[\phi^h:\phi\right]_t&=\Delta^{\mathrm{i}t}_{\phi^h,\phi}\Delta^{-\mathrm{i}t}_{\phi}=\mathrm{e}^{\mathrm{i}tc(\phi,h)}\mathrm{e}^{\mathrm{i}t(\log\Delta_\phi-h)}\Delta^{-\mathrm{i}t}_{\phi}\\&=\mathrm{e}^{\mathrm{i}tc(\phi,h)}\sum_{n=0}^{\infty}(-\mathrm{i})^n\int_0^t\mathrm{d}t_1\int_0^{t_1}\mathrm{d}t_2\cdots\int_0^{t_{n-1}}\mathrm{d}t_n\sigma^\phi_{t_n}(h)\cdots\sigma^\phi_{t_1}(h),\end{aligned}\tag{372}$$

$$\xi_\pi(\phi^h)=\exp\left(\frac{1}{2}\left(\log\Delta_\phi-\mathrm{supp}(\phi)h\,\mathrm{supp}(\phi)+c(\phi,h)\right)\right)\xi_\pi(\phi).\tag{373}$$

For unnormalised perturbed state $\widetilde{\phi^h}:=\mathrm{e}^{-c(\phi,-h)}\phi^{-h}$ this perturbation gives [33]

$$\xi_\pi\left(\widetilde{\phi^h}\right)=\sum_{n=0}^{\infty}(-1)^n\int_0^{1/2}\mathrm{d}t_1\int_0^{t_1}\mathrm{d}t\cdots\int_0^{t_{n-1}}\mathrm{d}t_n\left(\Delta^{t_n}_\phi h\Delta^{-t_n}_\phi\right)\cdots\left(\Delta^{t_1}_\phi h\Delta^{-t_1}_\phi\right)\xi_\pi(\phi),\tag{374}$$

with

$$-c(\omega,-h)=\log\left(\widetilde{\phi^h}(\mathbb{I})\right)=\log\left\|\xi_\pi\left(\widetilde{\phi^h}\right)\right\|^2_{\mathcal{H}_\pi},\tag{375}$$

which corresponds to [361]

$$\widetilde{\phi^h}(\mathbb{I})=\sup_{\omega\in\mathcal{N}^+_\star}\left\{-D_1|_{\mathcal{N}^+_{\star1}}(\omega,\phi)+\omega(h)+\omega(\mathbb{I})\right\}.\tag{376}$$

Petz [569] showed that for $h\in\mathcal{N}^{\mathrm{sa}}$ the equation (357) has a corresponding Fenchel dual

$$D_1|_{\mathcal{N}^+_{\star1}}(\phi,\omega)=\inf_{h\in\mathcal{N}^{\mathrm{sa}}}\left\{c(\omega,h)+\phi(h)\right\}.\tag{377}$$

This duality extends to the Banach dual pair of Banach spaces $(\mathcal{N}^{\mathrm{sa}},(\mathcal{N}^\star)^{\mathrm{sa}})$, if $\mathcal{N}^+_{\star1}$ in (355) and (377) is replaced by $\mathcal{N}^+_\star$. See [572, 550] for some further discussion of the Fenchel duality in this context.

According to the Donald–Petz theorem [231, 570], if $\psi\in\mathcal{N}^+_{\star1}$ and $C\subseteq\mathcal{N}^+_{\star1}$ is nonempty, convex, and weakly-$\star$ closed with $\mathrm{int}(C)\neq\varnothing$ and $\inf_{\phi\in C}\left\{D_1|_{\mathcal{N}^+_{\star1}}(\phi,\psi)\right\}<\infty$, then

$$\arg\inf_{\phi\in C}\left\{D_1|_{\mathcal{N}^+_{\star1}}(\phi,\psi)\right\}=\{*\}.\tag{378}$$

In [572] the same result was provided for $D_1|_{\mathcal{N}^+_{\star1}}(\phi,\psi)$ replaced by $D_1|_{\mathcal{N}^+_{\star1}}(\phi,\psi)+F(\phi)$, with $F:C\to\mathbb{R}\cup\{+\infty\}$ lower semi-continuous, convex, proper. A special case of (378), with $\lambda_1,\ldots,\lambda_n\in\mathbb{R}$, $h_1,\ldots,h_n\in\mathcal{N}^{\mathrm{sa}}$, $n\in\mathbb{N}$, and

$$C=\{\phi\in\mathcal{N}^+_{\star1}\mid\phi(h_i)=\lambda_i\ \ \forall i\in\{1,\ldots,n\}\}\tag{379}$$

was investigated, under some additional assumptions, in [642, 600, 532].

Now let us consider the entropic projections of D_γ given by (299) for $\gamma\in]0,1[$. The following results were obtained first by Jenčová [360] for the Jenčová–Ojima γ-distance and its corresponding dualistic Brègman functional.

Proposition 3.12. *1) if $y\in L_{1/\gamma}(\mathcal{N})$ and $K\subseteq L_{1/\gamma}(\mathcal{N})$ is nonempty, weakly closed, convex, then:*

i) $\bar{\mathfrak{P}}^{\Psi_\gamma}_K(y):=\arg\inf_{x\in K}\left\{\bar{D}_{\Psi_\gamma}(x,y)\right\}=\{*\}$, *iii)*

$$\bar{D}_{\Psi_\gamma}(x,y) \geq \bar{D}_{\Psi_\gamma}(x,\bar{\mathfrak{P}}^{\Psi_\gamma}_K(y)) + \bar{D}_{\Psi_\gamma}(\bar{\mathfrak{P}}^{\Psi_\gamma}_K(y),y) \;\; \forall x \in K, \tag{380}$$

and, equivalently,

$$\mathrm{re}\left[\!\left[x - \bar{\mathfrak{P}}^{\Psi_\gamma}_K(y), \mathbf{L}_{\Psi_\gamma}(y) - \mathbf{L}_{\Psi_\gamma}(\bar{\mathfrak{P}}^{\Psi_\gamma}_K(y))\right]\!\right]_{L_{1/\gamma}(\mathcal{N})\times L_{1/(1-\gamma)}(\mathcal{N})} \leq 0 \;\; \forall x \in K. \tag{381}$$

iv) the equality in (380) *and* (381) *holds if* K *is additionally a vector subspace of* $L_{1/\gamma}(\mathcal{N})$*,*

2) if $\psi \in \mathcal{N}^+_\star$ *and* $C \subseteq \mathcal{N}^+_\star$ *is nonempty,* $\ell_\gamma(C) \subseteq L_{1/\gamma}(\mathcal{N})$ *is convex, and* C *is closed in the topology induced by* $\tilde{\ell}^{-1}_\gamma$ *from the weak topology of* $L_{1/\gamma}(\mathcal{N})$*, then*

i) $\mathfrak{P}^{D_\gamma}_C(\psi) := \arg\inf_{\phi\in C}\{D_\gamma(\phi,\psi)\} = \{*\}$*,*

iii) if $\ell_\gamma(C)$ *is a vector subspace of* $L_{1/\gamma}(\mathcal{N})$*, then the generalised pythagorean equation*

$$D_\gamma(\omega,\psi) = D_\gamma(\omega,\mathfrak{P}^{D_\gamma}_C(\psi)) + D_\gamma(\mathfrak{P}^{D_\gamma}_C(\psi),\psi) \;\; \forall \omega \in C \tag{382}$$

holds.

Proof. Because $\bar{D}_{\Psi_\gamma}$ given by (317) is a Brègman functional in the sense of (B_4), the theorems (P_1) on existence, uniqueness and properties of Brègman projections for definitions (B_3) and (B_4) provided in Section 3.2 apply also in this case. The corresponding results for dualistic Brègman distance D_γ can be obtained by an extension of D_γ to $\hat{D}_\gamma$, defined on the whole space $\mathcal{N}_\star$ by replacing the term $[\![\ell_\gamma(\omega),\ell_{1-\gamma}(\phi)]\!]_\gamma$ in (319) by $\mathrm{re}\left[\!\left[\tilde{\ell}_\gamma(\omega),\tilde{\ell}_{1-\gamma}(\phi)\right]\!\right]_\gamma$. Because $\tilde{\ell}_\gamma$ are homeomorphisms (hence, bijections) between Banach spaces $\mathcal{N}_\star$ and $L_{1/\gamma}(\mathcal{N})$, the theorems on existence, uniquenes, and pythagorean theorem for projections for $\bar{D}_{\Psi_\gamma}$ on $L_{1/\gamma}(\mathcal{N})$ can be translated in terms of topology induced by $\tilde{\ell}^{-1}_\gamma$ on $\mathcal{N}_\star$, turning them into the corresponding theorems on projections for $\hat{D}_\gamma$. The results for D_γ follow then by the restriction of domain of $\hat{D}_\gamma$ to $\mathcal{N}^+_\star$.

Most of the conditions for (P_1) were already verified: $L_{1/\gamma}(\mathcal{N})$ is reflexive, Ψ_γ is lower semicontinuous, Gâteaux differentiable, essentially Gâteaux differentiable and essentially strictly convex on $\mathrm{efd}(\Psi_\gamma) = L_{1/\gamma}(\mathcal{N})$. The strict convexity of Ψ_γ follows from Gâteaux differentiability of $\Psi_{1-\gamma}$. Finally,

$$\lim_{\|x\|_{1/\gamma}\to+\infty}\frac{\Psi_\gamma(x)}{\|x\|_{1/\gamma}} = \frac{\gamma^{1-\gamma}}{1-\gamma}\lim_{\|x\|_{1/\gamma}\to+\infty}\|x\|^{-\gamma}_{1/\gamma} = +\infty \;\; \forall x \in K. \tag{383}$$

□

Remark 3.13. Jenčová [360] proved also that, under the same assumptions as in 1) and 2) above, respectively:

1.ii) $y \mapsto \bar{\mathfrak{P}}^{\Psi_\gamma}_K(y)$ is a continuous function from $L_{1/\gamma}(\mathcal{N})$ with its norm topology to K with the relative weak topology,

2.ii) $\psi \mapsto \mathfrak{P}^{D_\gamma}_C(\psi)$ is a continuous function from $\mathcal{N}^+_\star$ with the topology induced by $\tilde{\ell}^{-1}_\gamma$ from the norm topology of $L_{1/\gamma}(\mathcal{N})$ to C with the relative topology induced by $\tilde{\ell}^{-1}_\gamma$ from the weak topology of $L_{1/\gamma}(\mathcal{N})$.

4 Smooth geometries

A function $f : X \supseteq U \to V \subseteq Y$ between open subsets U, V of Banach spaces X, Y is called ***smooth*** iff it is continuous and Fréchet differentiable on U, its Fréchet derivative is also continuous and Fréchet differentiable on U, and the same holds for all its higher order Fréchet derivatives $\mathfrak{D}^{\mathrm{F}}(\ldots(\mathfrak{D}^{\mathrm{F}}f)\ldots)$. See [216, 141] for a discussion of higher order differentiability in Banach spaces, and [411] for a general treatment of smoothness in infinite dimensional vector spaces. A ***smooth atlas*** on an arbitrary set Z

is defined as a set of pairs $\{(U_i, w_p) \mid i \in I\}$, where I is a set, $U_i \subseteq Z$, $\bigcup_{i\in I} U_i = Z$, $\{Y_i \mid i \in I\}$ is a set of Banach spaces, $w_i : U_i \to w_i(U_i) \subseteq Y_i$ are bijections, $w_i(U_i)$ are open in Y_i, $W_i(U_i \cap U_j)$ is open in Y_i for all $i, j \in I$, and the map

$$w_j \circ w_i^{-1} : w_i(U_i \cap U_j) \to w_j(U_i \cap U_j) \tag{384}$$

is a smooth homeomorphism for all $i, j \in I$. Each pair (U_i, w_i) in a smooth atlas is called a ***chart***. Two smooth atlases on Z are called ***equivalent*** iff their union is a smooth atlas on Z. The set Z equipped with an equivalence class of smooth atlases equivalent to a given atlas $\{(U_i, w_i) \mid i \in I\}$ is called a ***smooth manifold*** modelled on Banach spaces $\{Y_i \mid i \in I\}$.. A subset X of a smooth manifold $(Z, \{(U_i, w_i) \mid i \in I\})$ is called a ***submanifold*** iff for all $x \in X$ there exists a chart (U_j, w_j) such that $x \in U_j$ and there exists a closed vector subspace W_j of Y_j such that

1) there exists a closed vector subspace V_j of Y_j such that $Y_j = W_j \oplus V_j$ and $W_j \cap V_j = \{0\}$,
2) $w_j(X \cap U_j) = W_j \cap w_j(U_j)$.

A collection $\{(U_j, w_j) \mid j \in J \subseteq I\}$ of all such charts defined a smooth atlas on X. As a result, $(X, \{(X \cap U_j, w_j|_{X\cap U_j}) \mid j \in J \subseteq I\})$ is a smooth manifold. If $(Z, \{(U_i, w_i) \mid i \in I\})$ is a smooth manifold, then a ***tangent vector*** at $x \in U_j \subset Z$, $j \in I$, is defined as a pair (x, y), where $y \in Y_j = w_j(U_j)$. A ***tangent space*** of Z at x, denoted $\mathbf{T}_x Z$, is defined as a space of all tangent vectors at x, so it is equal to Y_j, and, equivalently, to any Y_i such that $x \in w_i^{-1}(Y_i)$. Given the union $\bigcup_{i\in I} U_i \times Y_i \times \{i\}$, the quotient set of an equivalence relation

$$(x, v, i) \sim (y, u, j) \iff x = y \text{ and } (\mathfrak{D}^{\mathrm{F}}_{w_j(x)} w_i \circ w_j^{-1})(w) = v \tag{385}$$

is called a ***tangent bundle*** of Z and denoted $\mathbf{T}Z$. Its elements will be denoted $[x, v, i]_Z$. Given

$$\Pi_Z : \mathbf{T}Z \ni [x, v, i]_Z \mapsto x \in Z, \tag{386}$$

$$\mathbf{T} : U_i \mapsto \bigcup_{x\in U_i} \Pi_Z^{-1}(x), \tag{387}$$

$$\mathbf{T}w_i : \mathbf{T}U_i \ni [x, v, i]_Z \mapsto (w_i(x), v) \in w_i(U_i) \times Y_i, \tag{388}$$

the pairs $(\mathbf{T}U_i, \mathbf{T}w_i)$ are charts of a smooth atlas of $\mathbf{T}Z$, with

$$\mathbf{T}w_i \circ (Tw_j)^{-1} : w_j(U_i \cap U_j) \times Y_j \ni (x, v) \mapsto (w_i \circ w_j^{-1}(x), (\mathfrak{D}^{\mathrm{F}}_x w_i \circ w_j^{-1})(v)) \in w_i(U_i \cap U_j) \times Y_i. \tag{389}$$

The map Π_Z is smooth, and $\Pi_Z^{-1}(y) \cong \mathbf{T}_y Z \; \forall y \in Z$. According to the Lindenstrauss–Tzafriri theorem [459], if a Banach space Y has a property that for *every* closed vector subspace V there exists a closed vector subspace W such that $V \oplus W = Y$, then Y is isometrically isomorphic to a Hilbert space. Hence, if one requires that *all* closed vector subspaces of tangent spaces of a smooth manifold $\mathcal{M}$ are tangent spaces of some submanifolds of $\mathcal{M}$, then the tangent bundle of $\mathcal{M}$ *has* to consist of Hilbert spaces.

If $(X, \{(U_i, u_i) \mid i \in I\})$ and $(Z, \{(\widetilde{U}_j, \widetilde{u}_j) \mid j \in J\})$ are smooth manifolds, $f : X \to Z$ is smooth, then, for charts (U_i, u_i) and $(\widetilde{U}_j, \widetilde{u}_j)$ such that $f(U_i) \subseteq \widetilde{U}_j$ and for $v_i \in Y_i$ defined as a representative of $v \in \mathbf{T}_x X$ with $x \in U_i$, the map $\mathbf{T}f : \mathbf{T}_x X \to \mathbf{T}_{f(x)} Z$, defined by

$$[f(x), \mathbf{T}_x f(v), j]_Z = \left(\mathfrak{D}^{\mathrm{F}}_x (\widetilde{u}_j \circ f \circ u_i^{-1})\right)(v) \tag{390}$$

is linear and unique.

In what follows, we will assume that $\mathcal{M}$ is a smooth manifold, not necessarily finite dimensional. The coordinate-dependent representations of the smooth geometric structures will be provided under an (implicit) assumption of $\dim \mathcal{M} < \infty$, or $\mathcal{M}$ modelled on Hilbert spaces, or $\mathcal{M}$ modelled on a Banach space possesing a Schauder basis[39].

[39] A ***Schauder basis*** [653] in a Banach space X over $\mathbb{C}$ is a sequence $\{x_i\} \subseteq X$ such that $\forall x \in X \; \exists! \{\lambda_i\} \subseteq \mathbb{C}$ $x = \sum_{i=1} \lambda_i x_i$, where the convergence holds for the norm topology of X.

4.1 Riemannian and affine geometries

Every smooth manifold $\mathcal{M}$ is canonically equipped with a tangent space $\mathbf{T}_q\mathcal{M}$ at each $q \in \mathcal{M}$. A ***tangent bundle*** $\mathbf{T}\mathcal{M}$ of $\mathcal{M}$ is defined as $\mathbf{T}\mathcal{M} := \bigcup_{q\in\mathcal{M}} \mathbf{T}_q\mathcal{M}$. A ***cotangent space*** $\mathbf{T}^{\circledast}_q\mathcal{M}$ at each $q \in \mathcal{M}$ is defined as a Banach dual $\mathbf{T}^{\circledast}_q\mathcal{M} := (\mathbf{T}_q\mathcal{M})^\star$. A ***cotangent bundle*** $\mathbf{T}^{\circledast}\mathcal{M}$ of $\mathcal{M}$ is defined as $\mathbf{T}^{\circledast}\mathcal{M} := \bigcup_{q\in\mathcal{M}} \mathbf{T}^{\circledast}_q\mathcal{M}$. If $n, m \in \mathbb{N}$, then an (n,m)***-tensor bundle*** is defined as $\bigcup_{q\in\mathcal{M}} \left(\underline{\bigotimes}^n \mathbf{T}_q\mathcal{M}\right) \underline{\otimes} \left(\underline{\bigotimes}^m \mathbf{T}^{\circledast}_q\mathcal{M}\right)$, where $\underline{\otimes}$ denotes the tensor product considered in an algebraic sense (that is, without taking topological completion). A ***riemannian metric*** [619] is defined as a smooth function $\mathbf{g} : \mathbf{T}\mathcal{M} \times \mathbf{T}\mathcal{M} \to \mathbb{R}^+$ that acts pointwisely on fibers of $\mathbf{T}\mathcal{M}$ by

$$\mathcal{M} \ni q \mapsto \mathbf{g}_q : \mathbf{T}_q\mathcal{M} \times \mathbf{T}_q\mathcal{M} \to \mathbb{R} \tag{391}$$

and such that for any $v, u, w \in \mathbf{T}_q\mathcal{M}$, any $q \in \mathcal{M}$, and any $\lambda_1, \lambda_2 \in \mathbb{R}$ it satisfies

(i) $\mathbf{g}_q(u, v) = \mathbf{g}_q(v, u)$,

(ii) $v \neq 0 \;\Rightarrow\; \mathbf{g}_q(v, v) > 0$,

(iii) $\mathbf{g}_q(\lambda_1 u + \lambda_2 v, w) = \lambda_1 \mathbf{g}_q(u, w) + \lambda_2 \mathbf{g}_q(v, w)$.

Hence, $\mathbf{g}$ can be equivalently defined as a smooth section of the $(0,2)$-tensor bundle satisfying the conditions (i)-(iii). These conditions imply linearity in second argument and $\mathbf{g}_q(x, x) = 0 \iff x = 0$. A pair $(\mathcal{M}, \mathbf{g})$ is called a ***riemannian manifold*** or a ***riemannian geometry***.

A riemannian manifold $(\mathcal{M}, \mathbf{g})$ will be called: ***weak*** iff no additional conditions are assumed; ***semi-weak*** iff it uniquely determines a riemannian metric $\mathbf{g}^{\circledast}$ on $\mathbf{T}^{\circledast}\mathcal{M}$ such that, for each $x \in \mathbf{T}^{\circledast}\mathcal{M}$, $\mathbf{g}^{\circledast}(x, \cdot)$ is an injective bundle map [686, 687]; ***self-dual*** iff $\mathbf{T}_x\mathcal{M} \cong (\mathbf{T}_x\mathcal{M})^\star$ $\forall x \in \mathcal{M}$; ***semi-strong*** iff $\mathbf{g}(v, v) = \|v\|^2_{\mathbf{T}_x\mathcal{M}}$ $\forall \mathbf{T}_x\mathcal{M}$ $\forall x \in \mathcal{M}$; ***strong*** iff $\mathbf{T}_x\mathcal{M}$ is a Hilbert space for each $x \in \mathcal{M}$ [428, 393]; ***complete*** iff X is complete as a topological space in the topology induced by the metrical distance $d_{\mathbf{g}}$. The following implications hold in general: strong $\Rightarrow$ semi-strong $\Rightarrow$ self-dual $\Rightarrow$ semi-weak. Moreover, strong $\not\Rightarrow$ complete. If $\dim \mathcal{M} < \infty$, then every weak riemannian manifold satisfies all above properties. Various results of finite dimensional theory of riemannian geometry require different assumptions on the riemannian manifold when their generalisation to infinite dimensions is considered. In particular, the formula for a gradient of a function requires semi-weak structure, while Koszul formula for the Levi-Civita connection requires strong structure [428, 687]. In the remaining part of Section 4 we will avoid discussion of the necessary conditions required for the infinite dimensional riemannian manifold to support the given propositions. However, we will return to this problem in Section 5.1.

A ***length of a vector*** $v \in \mathbf{T}_q\mathcal{M}$ at $q \in \mathcal{M}$ is defined as $\sqrt{\mathbf{g}_q(v, v)}$. A ***curve*** in $\mathcal{M}$ is defined as a smooth map $c : \mathbb{R} \ni t \mapsto c(t) \in \mathcal{M}$. A ***finite curve*** in $\mathcal{M}$ is defined as a smooth map $c : [0,1] \ni t \mapsto c(t) \in \mathcal{M}$. Every curve c induces a vector field $\dot{c}(t) := \frac{\mathrm{d}}{\mathrm{d}t} c(t) \in \mathbf{T}_{c(t)}\mathcal{M}$. A ***length of a finite curve*** $c : [0,1] \ni t \mapsto c(t) \in \mathcal{M}$ connecting points $p := c(1)$ and $q := c(0)$ is defined as

$$d_c(p, q) := \int_0^1 \mathrm{d}t \sqrt{\mathbf{g}_{c(t)}(\dot{c}(t), \dot{c}(t))}. \tag{392}$$

A ***riemannian distance*** between $p \in \mathcal{M}$ and $q \in \mathcal{M}$ (with respect to a riemannian metric $\mathbf{g}$), defined as the length of locally shortest curve among all finite curves connecting p and q,

$$d_{\mathbf{g}}(p, q) := \inf_c \{d_c(p, q) \mid c(0) = q,\; c(1) = p\}, \tag{393}$$

is a metrical (Fréchet) distance.

An ***affine connection*** [447, 325, 761, 760] is defined as a map $\nabla : \mathbf{T}\mathcal{M} \times \mathbf{T}\mathcal{M} \to \mathbf{T}\mathcal{M}$ that acts pointwise on fibers of $\mathbf{T}\mathcal{M}$ by

$$\mathcal{M} \ni q \mapsto \nabla_q : \mathbf{T}_q\mathcal{M} \times \mathbf{T}_q\mathcal{M} \to \mathbf{T}_q\mathcal{M} \tag{394}$$

and satisfies

1. $\nabla(u, \lambda_1 v + \lambda_2 w) = \lambda_1 \nabla(u, v) + \lambda_2 \nabla(u, w)$,
2. $\nabla(u, fw) = u(f)w + f\nabla(u, w)$,
3. $\nabla(fu + hv, w) = f\nabla(u, w) + h\nabla(v, w)$,

for any $\lambda_1, \lambda_2 \in \mathbb{R}$, $f, h : \mathcal{M} \to \mathbb{R}$, $u, v, w \in \mathbf{T}\mathcal{M}$. We will use the common notation $\nabla(u, v) \equiv \nabla_u v$ for $u, v \in \mathbf{T}\mathcal{M}$ as well as $\nabla_q(u, v) \equiv (\nabla_q)_u v$ for $u, v \in \mathbf{T}_q\mathcal{M}$. Let $u(t) \in \mathbf{T}_{c(t)}\mathcal{M}$ be a vector field defined on every point of a curve $c(t)$. If $\nabla_{\dot{c}} u(t) = 0$, then $\mathbf{t}_c^\nabla u \equiv \mathbf{t}_{c(0),c(1)}^\nabla := u(t_1) \in \mathbf{T}_{c(t_1)}\mathcal{M}$ is called a ***parallel transport*** of $u = u(t_0) \in \mathbf{T}_{c(t_0)}\mathcal{M}$ with respect to ∇. The ***covariant derivative*** [616, 617] of u along v is defined as $\nabla_v u$. If $\hat{w} \in \mathbf{T}^{\circledast}\mathcal{M}$ and $v \in \mathbf{T}\mathcal{M}$, then the covariant derivative $\nabla_v \hat{w}$ is defined by

$$(\nabla_v \hat{w})(u) = \nabla_v(\hat{w}(u)) - \hat{w}(\nabla_v u) \ \ \forall u \in \mathbf{T}\mathcal{M}. \tag{395}$$

A curve $c(t)$ is called a ***∇-geodesic*** iff

$$\nabla_{\dot{c}(t)} \dot{c}(t) = 0. \tag{396}$$

The ***Riemann–Christoffel curvature tensor*** [620, 158, 462] of the affine connection ∇ is defined by

$$\begin{gathered} \mathrm{R}^\nabla : \mathbf{T}\mathcal{M} \times \mathbf{T}\mathcal{M} \times \mathbf{T}\mathcal{M} \to \mathbf{T}\mathcal{M}, \\ \mathrm{R}^\nabla(u, v, w) \equiv \mathrm{R}^\nabla(u, v)w := \nabla_u \nabla_v w - \nabla_v \nabla_u w - \nabla_{\{u,v\}} w, \end{gathered} \tag{397}$$

where $\{u, v\}(f) := (\mathfrak{D}^{\mathrm{F}}_f u)(v) - (\mathfrak{D}^{\mathrm{F}}_f v)(u)$ is a Lie bracket on $\mathbf{T}_q\mathcal{M}$ for any $q \in \mathcal{M}$. The ***torsion tensor*** of the connection ∇ is defined by [139, 140]

$$\begin{gathered} \mathfrak{T}^\nabla : \mathbf{T}\mathcal{M} \times \mathbf{T}\mathcal{M} \to \mathbf{T}\mathcal{M}, \\ \mathfrak{T}^\nabla(u, v) := \nabla_u v - \nabla_v u - \{u, v\}. \end{gathered} \tag{398}$$

The affine connection ∇ is called ***torsion-free*** or ***symmetric*** iff $\mathfrak{T}^\nabla(u, v) = 0 \ \forall u, v \in \mathbf{T}\mathcal{M}$. An affine connection ∇ is called ***flat*** iff $\mathrm{R}^\nabla(u, v, w) = 0 \ \forall u, v, w \in \mathbf{T}\mathcal{M}$. If ∇ is flat, then the ∇-parallel transport $\mathbf{t}_c^\nabla$ does not depend on c. If ∇ is flat and torsion-free, then $(\mathcal{M}, \nabla)$ is called an ***affine manifold*** or an ***affine geometry***.

A triple $(\mathcal{M}, \mathbf{g}, \nabla)$, where $\mathbf{g}$ is a riemannian metric and ∇ is an affine connection is called a ***metric–affine geometry***. An affine connection $\hat{\nabla}$ that satisfies any of the equivalent conditions

$$\mathbf{g}(\hat{\nabla}_u v, w) + \mathbf{g}(v, \hat{\nabla}_u w) = u(\mathbf{g}(v, w)) \quad \forall u, v, w \in \mathbf{T}\mathcal{M}, \tag{399}$$

$$\mathbf{g}(\mathbf{t}_c^{\hat{\nabla}} u, \mathbf{t}_c^{\hat{\nabla}} v) = \mathbf{g}(u, v) \ \forall u, v \in \mathbf{T}\mathcal{M} \quad \forall \text{ curves } c : \mathbb{R} \to \mathcal{M}, \tag{400}$$

$$\hat{\nabla}_u \mathbf{g}(v, w) = 0 \quad \forall u, v, w \in \mathbf{T}\mathcal{M}, \tag{401}$$

is called ***metric-compatible*** [325]. For every riemannian manifold $(\mathcal{M}, \mathbf{g})$ there exists a unique affine connection, which is torsion-free and metric-compatible [447]. It is denoted by $\bar{\nabla}$ and called the ***Levi-Civita*** connection. So, each riemannian geometry $(\mathcal{M}, \mathbf{g})$ determines a metric–affine geometry $(\mathcal{M}, \mathbf{g}, \bar{\nabla})$.

If some atlas of coordinate systems (θ^i) on $\mathcal{M}$ is chosen, and if $(\partial_i) \equiv (\frac{\partial}{\partial \theta^i})$ denotes the corresponding choice of the basis in $\mathbf{T}_q\mathcal{M}$, then the ***metric tensor*** of a riemannian metric $\mathbf{g}$ reads

$$\mathbf{g}_{ij}(\theta) \equiv \mathbf{g}_{ij}(q) := \mathbf{g}_{\theta(q)}(\partial_i, \partial_j), \tag{402}$$

the ***Christoffel symbols*** of an affine connection ∇ are [158]

$$\Gamma^\nabla_{ijk}(q) := \mathbf{g}_{\theta(q)}(\nabla_{\partial_i} \partial_j, \partial_k), \tag{403}$$

while the ***Ricci curvature scalar*** of ∇ and $\mathbf{g}$ is [618]

$$\kappa^\nabla(q) := \sum_{i,j,k,l} \mathrm{R}^\nabla_{ijkl}(q) \mathbf{g}^{il}(q) \mathbf{g}^{jk}(q). \tag{404}$$

If ∇ is symmetric, then $\Gamma^{\nabla}_{ijk} = \Gamma^{\nabla}_{jik}$. If ∇ is flat, then for every $q \in \mathcal{M}$ there exists a coordinate system (ζ^i) on an open neighbourhood $U \subseteq \mathcal{M}$ of q such that

$$\mathrm{R}^{\nabla}_{ijkl}(q) = 0, \tag{405}$$

$$(\nabla_q)_{\frac{\partial}{\partial\zeta^i}} \frac{\partial}{\partial\zeta^j} = 0, \tag{406}$$

$$\Gamma^{\nabla}_{ijk}(q) = 0, \tag{407}$$

$$\kappa^{\nabla}(q) = 0. \tag{408}$$

Such coordinate system (ζ^i) is called ***∇-affine***, because every other coordinate system $(\tilde{\zeta}^i)$ for which (406) also holds can be transformed to (ζ^i) by an affine transformation $\tilde{\zeta}^i = \sum_{j=1}^n \mathrm{M}^i_j \zeta^j + \lambda^i$, where M is a constant $n \times n$ matrix, $\lambda^i \in \mathbb{R}$, and $n := \dim \mathcal{M}$. A coordinate system which satisfies

$$\mathbf{g}_{\theta(q)}(\partial_i, \partial_j) = \boldsymbol{\delta}_{ij} \; \forall q \in \mathcal{M} \tag{409}$$

is called ***cartesian*** with respect to $\mathbf{g}$. If $\dim \mathcal{M} < \infty$ then the Levi-Civita connection $\bar{\nabla}$ of $\mathbf{g}$ is flat iff there exists a cartesian coordinate system with respect to $\mathbf{g}$, and in such case it is also $\bar{\nabla}$-affine.

In terms of arbitrary system (ζ^i) of coordinates on $\mathcal{M}$, the condition (396) reads

$$\ddot{\zeta}^k(c(t)) + \sum_{i,j} \dot{\zeta}^i(c(t))\dot{\zeta}^j(c(t)) \Gamma^{\nabla}{}^k_{ij}(\zeta(c(t))) = 0. \tag{410}$$

For every $u_q \in \mathbf{T}_q\mathcal{M}$ the Picard–Lindelöf theorem [585, 458] guarantees local existence and uniqueness of the solution of the differential equation (410) of a ∇-geodesic, which implies an existence of a unique geodesic curve $c^{\nabla}_{u_q}$ such that $c^{\nabla}_{u_q}(0) = q$ and $\dot{c}^{\nabla}_{u_q}(q) = u_q$. In consequence, the ***exponential map*** $\exp^{\nabla}_q : \mathbf{T}_q\mathcal{M} \to \mathcal{M}$ defined by $\exp^{\nabla}_q(u_q) := c^{\nabla}_{u_q} := c^{\nabla}_{u_q}(1)$ satisfies

$$\exp^{\nabla}_q(tu_q) = c^{\nabla}_{u_q}(t). \tag{411}$$

There always exists an open neighbourhood U of $0 \in \mathbf{T}_q\mathcal{M}$ and an open neighbourhood V of $q \in \mathcal{M}$ such that $\exp^{\nabla}_q|_U$ is a diffeomorphism $U \to V$ [82].

4.2 Norden–Sen and Eguchi geometries

A pair $(\nabla, \nabla^\dagger)$ of two affine connections over a smooth manifold $\mathcal{M}$ is called ***Norden–Sen dual*** with respect to a riemannian metric $\mathbf{g}$ on $\mathcal{M}$, iff [541, 665, 666, 542, 667, 543, 668, 669, 544, 545, 670, 546, 664]

$$\mathbf{g}(\nabla_u v, w) + \mathbf{g}(v, \nabla^\dagger_u w) = u(\mathbf{g}(v, w)) \; \forall u, v, w \in \mathbf{T}\mathcal{M}, \tag{412}$$

which is equivalent to

$$\mathbf{g}(\mathbf{t}^{\nabla}_c u, \mathbf{t}^{\nabla^\dagger}_c v) = \mathbf{g}(u, v) \tag{413}$$

for all $u, v \in \mathbf{T}\mathcal{M}$ and for all curves $c : \mathbb{R} \to \mathcal{M}$. Condition (412) is a generalisation of (399) (and, equivalently (413) is a generalisation of (400)). The quadruple $(\mathcal{M}, \mathbf{g}, \nabla, \nabla^\dagger)$ is called a ***Norden–Sen manifold*** or ***Norden–Sen geometry***. From the duality condition (412) it follows that

$$\mathrm{R}^{\nabla}(u, v, w) = \mathrm{R}^{\nabla^\dagger}(u, v, w) \; \forall u, v, w \in \mathbf{T}\mathcal{M}. \tag{414}$$

See [743, 524, 17, 430, 218, 540, 539] for later studies of this geometry.

Eguchi [243, 244, 245] showed that for any smooth manifold $\mathcal{M}$ and any smooth[40] distance D on $\mathcal{M}$ that satisfies

$$\mathfrak{D}^{\mathrm{F}}_v|_p \mathfrak{D}^{\mathrm{F}}_v|_p D(p, q)|_{q=p} \in]0, \infty[\quad \forall p \in \mathcal{M} \; \forall v \in \mathbf{T}_p\mathcal{M} \setminus \{0\}, \tag{415}$$

[40] For the purpose of the equation (416) only, it is sufficient to assume twice differentiability (and this is the case, for example, in [446]). Yet, consideration of $\mathbf{g}$ in (416) as riemannian metric requires to assume smoothness.

the distance D determines a riemannian metric $\mathbf{g}$ and a pair of affine connections $(\nabla, \nabla^\dagger)$ on $\mathcal{M}$, given by the ***Eguchi equations***

$$\mathbf{g}_\phi(u,v) := -\mathfrak{D}^{\mathrm{F}}_u|_\phi \mathfrak{D}^{\mathrm{F}}_v|_\omega D(\phi,\omega)|_{\omega=\phi}, \tag{416}$$

$$\mathbf{g}_\phi((\nabla_u)_\phi v, w) := -\mathfrak{D}^{\mathrm{F}}_u|_\phi \mathfrak{D}^{\mathrm{F}}_v|_\phi \mathfrak{D}^{\mathrm{F}}_w|_\omega D(\phi,\omega)|_{\omega=\phi}, \tag{417}$$

$$\mathbf{g}_\phi(v, (\nabla^\dagger_u)_\phi w) := -\mathfrak{D}^{\mathrm{F}}_u|_\omega \mathfrak{D}^{\mathrm{F}}_w|_\omega \mathfrak{D}^{\mathrm{F}}_v|_\phi D(\phi,\omega)|_{\omega=\phi}. \tag{418}$$

Every quadruple $(\mathcal{M}, \mathbf{g}, \nabla, \nabla^\dagger)$ determined in this way is a Norden–Sen geometry such that both ∇ and $\nabla^\dagger$ are torsion-free. Conversely, Matumoto [488] has shown that, for $\dim \mathcal{M} < \infty$, every torsion-free Norden–Sen geometry $(\mathcal{M}, \mathbf{g}, \nabla, \nabla^\dagger)$ is determined by a smooth distance functional D that is defined globally on $\mathcal{M}$ and satisfies (415). However, this distance is determined uniquely by the quadruple $(\mathcal{M}, \mathbf{g}, \nabla, \nabla^\dagger)$ only up to the third order term of its Taylor expansion. A torsion-free Norden–Sen geometry will be called an ***Eguchi manifold*** or ***Eguchi geometry***, while every pair $(\mathcal{M}, D)$ such that D is smooth, and (415) holds, will be called an ***Eguchi system***.

While in riemannian geometry the affine connection is determined by riemannian metric, in the Eguchi geometry the triple of riemannian metric and two Norden–Sen dual affine connections are determined by distance. Moreover, the Levi-Civita connection $\bar{\nabla}$ of an associated riemannian geometry $(\mathcal{M}, \mathbf{g})$ satisfies $\bar{\nabla} = (\nabla + \nabla^\dagger)/2$. In this sense, the Eguchi geometry provides a generalisation of a riemannian geometry. This way the Eguchi geometry (based on the nonsymmetric distance) provides a generalisation of all main notions of cartesian geometry: distance, length, parallelity and orthogonality. Generalisation of the cartesian distance is provided by the distance D, the induced riemannian metric $\mathbf{g}$ provides the generalisation of orthogonality and length, while the induced torsion-free Norden–Sen dual connections $(\nabla, \nabla^\dagger)$ provide a generalisation of parallelity.[41] The invariance of length under parallel transport that characterises riemannian geometry is weakened to covariance in the sense of (413).

Every Eguchi geometry $(\mathcal{M}, \mathbf{g}, \nabla, \nabla^\dagger)$ allows one to define the family of affine ***ϑ-connections***,

$$\widetilde{\nabla}^\vartheta := (1-\vartheta)\nabla + \vartheta\nabla^\dagger \quad \forall\vartheta \in \mathbb{R}, \tag{419}$$

as well as a completely symmetric ***skewness tensor***,

$$C_{ijk}(q) := \partial_k \mathbf{g}_{ij}(q) = \Gamma^{\nabla^\dagger}_{kij}(q) - \Gamma^{\nabla}_{kji}(q). \tag{420}$$

The pairs $(\widetilde{\nabla}^\vartheta, \widetilde{\nabla}^{1-\vartheta})$ are torsion-free and Norden–Sen dual with respect to $\mathbf{g}$ [524, 430, 431]. Moreover,

$$\widetilde{\nabla}^{1/2} = \frac{1}{2}(\widetilde{\nabla}^\vartheta + \widetilde{\nabla}^{1-\vartheta}) = \bar{\nabla}, \tag{421}$$

that is, $(\vartheta = \frac{1}{2})$-connection is a Levi-Civita connection with respect to $\mathbf{g}$. As shown by Lauritzen [431], the Eguchi geometries $(\mathcal{M}, \mathbf{g}, \widetilde{\nabla}^\vartheta, \widetilde{\nabla}^{1-\vartheta})$ can be characterised by means of the following theorem: for every triple $(\mathcal{M}, \mathbf{g}, C)$, where $C : \mathbf{T}\mathcal{M} \times \mathbf{T}\mathcal{M} \times \mathbf{T}\mathcal{M} \to \mathbb{R}$ is any completely symmetric third-rank tensor on $\mathcal{M}$, there exists a unique affine connection $\widetilde{\nabla}^\vartheta$, which is torsion-free and satisfies

$$\widetilde{\nabla}^\vartheta_u \mathbf{g}(v,w) = (1-2\vartheta)C(u,v,w). \tag{422}$$

The triples $(\mathcal{M}, \mathbf{g}, C)$ will be called ***Lauritzen manifolds***. For $\vartheta = 0$ this theorem sets up a bijection between the Eguchi geometries and the Lauritzen manifolds. The Christoffel symbols of $\widetilde{\nabla}^\vartheta$ read

$$\widetilde{\Gamma}^\vartheta_{ijk}(q) = \bar{\Gamma}_{ijk}(q) + \left(\vartheta - \frac{1}{2}\right) C_{ijk}(q). \tag{423}$$

In coordinate independent terms this corresponds to

$$\widetilde{\nabla}^\vartheta_u v = \bar{\nabla}_u v + \left(\vartheta - \frac{1}{2}\right) C(u,v) \quad \forall u,v \in \mathbf{T}\mathcal{M}, \tag{424}$$

[41] The idea that D should be considered as generalisation of the cartesian distance, while the connection ∇ associated to a projection by means of D should be considered as a proper generalisation of parallelity (at least in the setting of statistical manifolds) is due to Chencov [147, 150].

where $C(u,v)$ is defined by

$$C(u,v,w) = \mathbf{g}(C(u,v),w) \quad \forall u,v,w \in \mathbf{T}\mathcal{M}. \tag{425}$$

Only the distances along ($\vartheta = \frac{1}{2}$)-geodesics are riemannian distances. Other ϑ-geodesics are not the curves of locally minimal length. Hence, the length of a curve can serve as an affine parameter of this curve only for ($\vartheta = \frac{1}{2}$)-geodesics. Note also that $\mathbf{g}_{ij}(q)$ in general depends on q, even in a ϑ-affine coordinate system, due to (422). In consequence, the ϑ-parallel transport generally changes the length of vectors.

In a local coordinate basis (∂_i) at point q, the corresponding ϑ-torsion tensor, ϑ-curvature tensor, and ϑ-curvature scalar read, respectively, [777, 709, 241, 14, 15, 17, 786]

$$\widetilde{\mathfrak{T}}^{\vartheta}_{ijk}(q) = \mathbf{g}_q(\widetilde{\mathfrak{T}}^{\vartheta}(\partial_i,\partial_j),\partial_k) = \widetilde{\Gamma}^{\vartheta}_{ijk}(q) - \widetilde{\Gamma}^{\vartheta}_{jik}(q), \tag{426}$$

$$\widetilde{\mathrm{R}}^{\vartheta}_{ijkm}(q) = \mathbf{g}_q(\widetilde{\mathrm{R}}^{\vartheta}(\partial_i,\partial_j,\partial_k),\partial_m) = \partial_i\widetilde{\Gamma}^{\vartheta}_{jkm}(q) - \partial_j\widetilde{\Gamma}^{\vartheta}_{ikm}(q) + \sum_r \widetilde{\Gamma}^{\vartheta}_{irm}(q)\widetilde{\Gamma}^{\vartheta r}{}_{jk}(q) - \sum_r \widetilde{\Gamma}^{\vartheta}_{jrm}(q)\widetilde{\Gamma}^{\vartheta r}{}_{ik}(q), \tag{427}$$

$$\begin{aligned}\widetilde{\kappa}^{\vartheta}(q) &= \sum_{i,j,k,m} \widetilde{\mathrm{R}}^{\vartheta}_{ijkm}(q)\mathbf{g}^{im}(q)\mathbf{g}^{jk}(q) \\ &= \widetilde{\kappa}^{1/2}(q) + \vartheta(1-\vartheta)\sum_{i,j}\mathbf{g}^{ij}(q)\left(\sum_{k,m} C^m_{ik}(q)C^k_{jm}(q) - \sum_{k,m} C^m_{ij}(q)C^k_{km}(q)\right).\end{aligned} \tag{428}$$

From the Norden–Sen duality between $\widetilde{\nabla}^{\vartheta}$ and $\widetilde{\nabla}^{1-\vartheta}$ it follows that

$$\widetilde{\mathrm{R}}^{\vartheta}(u,v,w) = \widetilde{\mathrm{R}}^{1-\vartheta}(u,v,w), \tag{429}$$

hence [431]

$$\widetilde{\mathrm{R}}^{\vartheta}_{ijkl}(q) = -\widetilde{\mathrm{R}}^{1-\vartheta}_{ijlk}(q). \tag{430}$$

So, $\widetilde{\nabla}^{\vartheta}$ is flat iff $\widetilde{\nabla}^{1-\vartheta}$ is flat. If $\widetilde{\mathfrak{T}}^{\vartheta} = \widetilde{\mathfrak{T}}^{1-\vartheta} = 0$ and $\widetilde{\mathrm{R}}^{\vartheta} = \widetilde{\mathrm{R}}^{1-\vartheta} = 0$, then the quadruple $(\mathcal{M},\mathbf{g},\widetilde{\nabla}^{\vartheta},\widetilde{\nabla}^{1-\vartheta})$ is called a ***ϑ-hessian manifold*** [679, 784]. Note that the Levi-Civita connection (421) on the ($\vartheta \neq \frac{1}{2}$)-hessian manifold may possess nonzero riemannian ($\vartheta = \frac{1}{2}$)-curvature.

4.3 Hessian manifolds

If $\mathfrak{T}^{\nabla} = \mathfrak{T}^{\nabla^\dagger} = 0$ and $\mathrm{R}^{\nabla} = \mathrm{R}^{\nabla^\dagger} = 0$, then the Norden–Sen geometry $(\mathcal{M},\mathbf{g},\nabla,\nabla^\dagger)$ is called a ***dually flat manifold*** [524, 17, 23]. In such case there exists a pair $(\ell,\ell^\dagger)$ of coordinate systems on $\mathcal{M}$, called ***dually flat coordinates***, such that ℓ is a ∇-affine coordinate system, while $\ell^\dagger$ is a $\nabla^\dagger$-affine coordinate system. Note that the flatness of ∇ and $\nabla^\dagger$ does not imply the flatness of $\bar{\nabla}$. The dual flatness of a pair (θ,η) of coordinate systems is equivalent to the orthogonality of their tangent vectors at q with respect to the riemannian metric $\mathbf{g}$ at q,

$$\mathbf{g}_q\left((\mathbf{T}_q\theta)^{-1}\left(\frac{\partial}{\partial\theta^i}\right),(\mathbf{T}_q\eta)^{-1}\left(\frac{\partial}{\partial\eta_j}\right)\right) = \boldsymbol{\delta}_i^j \quad \forall q \in \mathcal{M}. \tag{431}$$

A triple $(\mathcal{M},\mathbf{g},\nabla)$ with a riemannian metric $\mathbf{g}$ and a torsion-free affine connection ∇ is called a ***Codazzi structure*** iff it satisfies

$$(\nabla_u\mathbf{g})(v,w) = (\nabla_v\mathbf{g})(u,w)\ \forall u,v,w \in \mathbf{T}\mathcal{M}. \tag{432}$$

A riemannian metric $\mathbf{g}$ on an affine manifold $(\mathcal{M},\nabla)$ with flat ∇ is said to be ***hessian***, and denoted $\mathbf{g}^{\Phi}$, iff there exists a smooth function $\Phi : \mathcal{M} \to \mathbb{R}$ such that [675, 676, 155]

$$\mathbf{g}(u,v) = (\nabla_u\mathbf{d}\Phi)(v) \quad \forall u,v \in \mathbf{T}\mathcal{M}. \tag{433}$$

Such triple $(\mathcal{M}, \mathbf{g}, \nabla)$ is called a ***hessian manifold*** or a ***hessian structure*** [678] (see also [405, 744, 515, 236]). If (ξ^i) is a ∇-affine coordinate system on $\mathcal{M}$, then (433) is locally equivalent to

$$\mathbf{g}_{ij}(q) = \frac{\partial^2 \Phi}{\partial \xi^i \partial \xi^j} \mathbf{d}\xi^i \otimes \mathbf{d}\xi^j. \tag{434}$$

Given a riemannian manifold $(\mathcal{M}, \mathbf{g})$ and an affine connection ∇ on $\mathcal{M}$, the following conditions are equivalent [677, 679, 678]:

1) $(\mathcal{M}, \mathbf{g}, \nabla)$ is a hessian manifold,

2) $(\mathcal{M}, \mathbf{g}, \nabla)$ is a Codazzi structure and ∇ is flat,

3) $(\mathcal{M}, \mathbf{g}, \nabla, 2\bar{\nabla} - \nabla)$ is a dually flat Norden–Sen geometry, where $\bar{\nabla}$ is the Levi-Civita connection of $\mathbf{g}$,

4) $\mathbf{g}((\bar{\nabla} - \nabla)_u v, w) = \mathbf{g}(v, (\bar{\nabla} - \nabla)_u w)$ $\forall u, v, w \in \mathbf{T}\mathcal{M}$ and for a torsion-free flat ∇.

Hence, hessian structures belong to an intersection of the Norden–Sen geometries with Codazzi structures, and are characterised by the symmetry and flatness of the affine connection. From 3) above it follows that $(\mathcal{M}, \mathbf{g}, \nabla^\dagger)$ for $\nabla^\dagger := 2\bar{\nabla} - \nabla$ is also a hessian manifold. This sets up a bijection between hessian manifolds and dually flat manifolds.

The equation (434) suggests us to introduce the function

$$\Psi := \Phi \circ \ell^{-1} : X \to \mathbb{R}, \tag{435}$$

where X is a codomain vector space of the coordinate system $\ell : \mathcal{M} \to X$. Let $(X, X^{\mathbf{d}}, [\![\cdot,\cdot]\!]_{X \times X^{\mathbf{d}}})$ be a dual pair (in the sense of Section 3.2) such that $(\ell, \ell^\dagger) : \mathcal{M} \to X \times X^{\mathbf{d}}$ is a pair of dually flat coordinates of a dually flat manifold $(\mathcal{M}, \mathbf{g}, \nabla, \nabla^\dagger)$. Using (435) and (433) we obtain

$$\mathbf{g}(u, v) = (\nabla_u \mathbf{d}(\Psi \circ \ell^{-1}))(v) = (\nabla^\dagger_u \mathbf{d}(\tilde{\Psi} \circ (\ell^\dagger)^{-1}))(v). \tag{436}$$

Hence, for a given dually flat manifold, the form of its hessian riemannian metric $\mathbf{g}^\Phi$ is determined by Ψ. This allows to denote $\mathbf{g}^\Phi$ as $\mathbf{g}^\Psi$. For $\dim \mathcal{M} = n < \infty$ the direct application of (434) gives $\tilde{\Psi} = \Psi^{\mathbf{L}}$ [678], where $\Psi^{\mathbf{L}}$ is a Legendre transform of Ψ with respect to the duality (259). We define

$$\Phi^{\mathbf{L}} := \Psi^{\mathbf{L}} \circ \ell^\dagger. \tag{437}$$

The functions Φ and $\Phi^{\mathbf{L}}$ are called ***scalar potentials***. In what follows, we will assume that Ψ is convex (the convexity of $\Psi^{\mathbf{L}}$ follows from the properties of the Legendre transform, independently of the convexity of Φ. For $\dim \mathcal{M} = n < \infty$ one can express $\mathbf{g}$, ∇, and $\nabla^\dagger$ in a coordinate-dependent form

$$\eta_i = \partial_i \Psi(\theta), \tag{438}$$
$$\theta^i = \partial^i \Psi(\eta), \tag{439}$$
$$\mathbf{g}_{ij}(\theta) = \partial_i \partial_j \Psi(\theta), \tag{440}$$
$$\mathbf{g}^{ij}(\eta) = \partial^i \partial^j \Psi^{\mathbf{L}}(\eta), \tag{441}$$
$$\Gamma^{\nabla}_{ijk}(\theta) = \partial_i \partial_j \partial_k \Psi(\theta), \tag{442}$$
$$\Gamma^{\nabla^\dagger ijk}(\eta) = \partial^i \partial^j \partial^k \Psi^{\mathbf{L}}(\eta), \tag{443}$$
$$\Gamma^{\nabla}_{ijk}(\eta) = 0, \tag{444}$$
$$\Gamma^{\nabla^\dagger ijk}(\theta) = 0, \tag{445}$$
$$\sum_{i=1}^{n} \theta^i(p) \eta_i(p) = \Psi(\theta(p)) + \Psi^{\mathbf{L}}(\eta(p)), \tag{446}$$

where $\partial_i := \frac{\partial}{\partial\theta^i}$ and $\partial^i := \frac{\partial}{\partial\eta_i}$. The equation (446) follows from the lower bound of the Young–Fenchel inequality (194). The equations (438)-(439) are the same as (257)-(258). The equations (440)-(443) can be obtained as a result of Eguchi equations (416)-(418) applied to the Brègman distance

$$D_\Psi(p,q) = \Psi(\theta(p)) + \Psi^{\mathbf{L}}(\eta(q)) - \sum_{i=1}^{n}\theta^i(p)\eta_i(q). \tag{447}$$

Equations (444)-(445) are instances of (406), which follow from dual flatness. Conversely [17], if there exists a convex function Φ such that its hessian (matrix of second derivatives) determines pointwise a riemannian metric, then there exists a pair of dually flat coordinate systems, together with a conjugate potential $\Phi^{\mathbf{L}}$, satisfying equations (438)-(446).

The structure of dually flat manifolds determines affine connections up to affine transformations, and determines corresponding scalar potentials up to linear terms. This suggests to use the Eguchi equations in order to reconstruct the unique distance corresponding to the structure of a dually flat manifold, but it is in general impossible, because Eguchi equations determine distance only up to the third order terms. Nevertheless, every dually flat manifold has a naturally associated Brègman distance, which is determined by the Young–Fenchel inequality, and the pair $(\ell,\ell^\dagger)$ of dually flat coordinates. More precisely [524, 422, 218, 423, 487], every n-dimensional smooth manifold equipped with a riemannian metric and a pair of flat and mutually Norden–Sen dual affine connections determines a pair of affine immersions that are related to each other by the Legendre transformation and are unique up to affine transformation of the coordinate codomain space $\mathbb{R}^n$. These affine immersions determine in turn a distance functional (447), which belongs to a class of Brègman distances.

The relationship between Brègman distance and dually flat manifolds can be characterised as follows [23, 383]: if a smooth manifold $\mathcal{M}$ is equipped with two torsion-free affine connections ∇_1 and ∇_2, a riemannian metric $\mathbf{g}$ and a distance D, then $(\mathcal{M},\mathbf{g},\nabla_1,\nabla_2)$ is a dually flat manifold and D is its associated Brègman distance iff

$$D(p,q) + D(q,r) = D(p,r) + \mathbf{g}_q((\exp_q^{\nabla_1})(p),(\exp_q^{\nabla_2})(r)) \quad \forall p,q,r\in\mathcal{M}. \tag{448}$$

The ***generalised cosine equation*** (448) is a special case of the equation (273).

Let us consider further properties of dually flat manifolds [17, 23]. The coordinate system (θ^i) on $\mathcal{Q}$ is called ∇-***affine*** iff all basis vectors fields are ∇-parallel on $\mathcal{Q}$. If, for a given ∇ on $\mathcal{Q}$, there exists a ∇-affine coordinate system, then ∇ is called ***flat***, and $\mathcal{Q}$ is called ∇-***flat*** or ∇-***affine***. The ∇-flatness of $\mathcal{Q}$ is equivalent to the vanishing of the Riemann curvature tensor R^∇. A manifold $\mathcal{Q}$ is called ∇-***autoparallel*** iff $\nabla_u v\in\mathbf{T}\mathcal{Q}$ $\forall u,v\in\mathbf{T}\mathcal{Q}$. If a manifold $\mathcal{M}$ is ∇-flat, then $\mathcal{Q}\subseteq\mathcal{M}$ is ∇-autoparallel iff $\mathcal{Q}$ can be expressed as an affine subspace (or an open subset of an affine subspace) of $\mathcal{M}$ with respect to a ∇-affine coordinate system on $\mathcal{M}$. A ∇-autoparallelity of $\mathcal{Q}$ is equivalent to the vanishing of the Euler–Schouten imbedding curvature tensor [17, 431]. If $\mathcal{Q}\subseteq\mathcal{M}$ is ∇-autoparallel and $\mathcal{M}$ is ∇-flat, then $\mathcal{Q}$ is ∇-flat and ∇-geodesics on $\mathcal{Q}$ have linear equations in ∇-affine coordinates. If $(\mathcal{M},\mathbf{g},\nabla,\nabla^\dagger)$ is dually flat, and $\mathcal{Q}\subseteq\mathcal{M}$ is ∇-autoparallel or $\nabla^\dagger$-autoparallel, then $(\mathcal{Q},\mathbf{g}_\mathcal{Q},\nabla_\mathcal{Q},\nabla^\dagger_\mathcal{Q})$ is also dually flat, with $(\mathbf{g}_\mathcal{Q},\nabla_\mathcal{Q},\nabla^\dagger_\mathcal{Q})$ induced on $\mathcal{Q}$ by $(\mathbf{g},\nabla,\nabla^\dagger)$ [17, 23]. If ∇_1 and ∇_2 are affine connections on $\mathcal{M}$, $\mathbf{g}$ is a riemannian metric on $\mathcal{M}$, and $\mathcal{Q}\subseteq\mathcal{M}$ is ∇_2-autoparallel, then a point $p_\mathcal{Q}\in\mathcal{Q}$ is called a $(\mathbf{g},\nabla_1,\nabla_2)$-***projection*** of $p\in\mathcal{M}$ onto $\mathcal{Q}$ iff the ∇_1-geodesic $c^{\nabla_1}(t)$ connecting p with $p_\mathcal{Q}$ satisfies

$$\mathbf{g}_{p_\mathcal{Q}}(\dot{c}^{\nabla_1}(t),\dot{c}^{\nabla_2}(s)) = 0 \;\; \forall c^{\nabla_2}, \tag{449}$$

where $\dot{c}^{\nabla_1},\dot{c}^{\nabla_2}\in\mathbf{T}_{p_\mathcal{Q}}\mathcal{M}$, and c^{∇_2} varies over all ∇_2-geodesic lines intersecting $p_\mathcal{Q}$ and contained in $\mathcal{Q}$. A set $\mathcal{Q}\subseteq\mathcal{M}$ is called ∇-***convex*** iff for all $p_1,p_2\in\mathcal{Q}$ there exists a unique ∇-geodesic connecting p_1 with p_2 and entirely included in $\mathcal{Q}$.

Let $(\mathcal{M},\mathbf{g},\nabla,\nabla^\dagger)$ be a dually flat manifold, D_Ψ its canonical Brègman distance, $\mathcal{Q}\subseteq\mathcal{M}$, $p_\mathcal{Q}\in\mathcal{Q}$, $p\in\mathcal{M}$. If $\mathcal{Q}$ is $\nabla^\dagger$-autoparallel submanifold of $\mathcal{M}$, then [23]

$$D_\Psi(p,p_\mathcal{Q}) = \inf_{q\in\mathcal{Q}}\{D_\Psi(p,q)\} \tag{450}$$

holds iff $p_{\mathcal{Q}}$ is a $(\mathbf{g}, \nabla, \nabla^{\dagger})$-projection of p onto $\mathcal{Q}$. If $\mathcal{Q}$ is a closed set with a smooth boundary $\partial\mathcal{Q}$, then the $(\mathbf{g}, \nabla, \nabla^{\dagger})$-projection of p onto $\partial\mathcal{Q}$ is unique if $\mathcal{Q}$ is $\nabla^{\dagger}$-convex [17]. Hence, while the existence of $(\mathbf{g}, \nabla, \nabla^{\dagger})$-projection from $\mathcal{M}$ to $\mathcal{Q}$, given by the minimum of Brègman distance, requires $\nabla^{\dagger}$-autoparallelity of $\mathcal{Q}$, its uniqueness requires $\nabla^{\dagger}$-convexity of $\mathcal{Q}$. By the theorem above, if $\mathcal{M}$ is $\nabla^{\dagger}$-flat, then $\mathcal{Q}$ is $\nabla^{\dagger}$-autoparallel iff it is $\nabla^{\dagger}$-affine. So, while in general the existence and uniqueness of $(\mathbf{g}, \nabla, \nabla^{\dagger})$-projection onto $\mathcal{Q}$ requires $\nabla^{\dagger}$-autoparallelity and $\nabla^{\dagger}$-convexity of $\mathcal{Q}$, this requirement is weakened to $\nabla^{\dagger}$-affinity and $\nabla^{\dagger}$-convexity in the case when $\mathcal{M}$ is a dually flat space.

In particular, if $p, q, r \in \mathcal{M}$ such that p and q are points at ∇-geodesic, q and r are points at $\nabla^{\dagger}$-geodesic, and these geodesics are orthogonal at q in the sense of (449) (but without quantifier '$\forall c^{\nabla_2}$'), then the ***generalised pythagorean equation*** [23]

$$D_{\Psi}(p, q) + D_{\Psi}(q, r) = D_{\Psi}(p, r) \tag{451}$$

holds. It is a special case of nonsmooth generalised pythagorean equation (280).

See [23, 784, 510, 20, 787, 21] for a further discussion of dually flat geometries, their relationship with convex Legendre conjugate potentials and Brègman distances, as well as the corresponding geometric description of minimisation problems.

5 Smooth information geometries

For a given finite dimensional information model $\mathcal{M}$, the smooth manifold structure can be introduced using the smooth embeddings into open subsets of $\mathbb{R}^n$. As a result, various differential geometric objects on $\mathcal{M}(\mathcal{A})$ and $\mathcal{M}(\mathcal{N})$ can be studied. For infinite dimensional case, a suitable smooth manifold structure modelled on Banach spaces is required. Following the work of Pistone and Sempi, most of the constructions (with an exception of e.g. [46, 535]) employ a suitably constructed family of (commutative or noncommutative) Orlicz spaces. In general, the passage from the setting of information distances on information models to smooth information geometric structures on information manifolds imposes restriction on integrability to the sets of countably finite W^*-algebras, and restriction of differentiability to the setting of Fréchet derivatives.[42] The former is the price paid for invertibility and chain rule of the Radon–Nikodým quotients, while the latter is the price paid for linearity of derivatives of arbitrary degree.

The consideration of finite dimensional parametric statistical models as smooth manifolds dates back to [601, 351, 13, 147], and the expositions of the subject can be found in [152, 14, 17, 54, 507, 515, 23, 354, 415, 383, 43]. The detailed treatment of the information geometric reformulation of statistics with the key role played by the smooth geometric structures on statistical manifolds is given in [152, 381, 17, 382, 507, 383]. Consideration of multidimensional parametric quantum models dependent on smooth parameters dates back at least to [345, 322], however an explicit study of smooth geometric structures on quantum models was started in [337, 338, 342, 49, 727, 728, 520, 521, 506]. The comprehensive treatment of a smooth geometry of finite dimensional quantum manifolds still waits for its book, but [506, 732, 575, 446, 286, 358, 359, 159, 67, 305, 733] contain some partial overviews of the topic. The smooth manifold structure on nonparametric statistical and quantum models, as well as further smooth geometric structures on them, are a subject of current research, and one can consult [286, 361, 143, 702, 704, 590] for further details.

The main aim of this Section is to provide an overview of the main structures and results of smooth information geometry, in commutative and quantum, as well as in parametric and nonparametric, settings. Our intention is to show how the smooth information geometric structures arise as an approximation to nonsmooth information geometry of $\mathfrak{f}$-distances and generalised Brègman distances (the latter entering through the dually flat geometries). In particular, we prove the smooth parametric quantum analogue of the generalised pythagorean theorem (382).

[42] As one of the consequences, one expects that the dualistic Bregman distances naturally associated with the smooth information manifolds (e.g. by means discussed in Section 4.3) will correspond to Bregman functionals of type (B_4).

5.1 Information manifolds

If a statistical model (resp., probabilistic model) $\mathcal{M}(\mathcal{A})$ can be equipped with a smooth manifold structure, then it is called a ***statistical manifold*** (resp., ***probability manifold***). If a quantum information model $\mathcal{M}(\mathcal{N})$ can be equipped with a smooth manifold structure, then it will be called a ***quantum manifold*** (or ***quantum information manifold***).

Consider a statistical model $\mathcal{M}(\mathcal{A}) \subseteq L_1(\mathcal{A})^+$ with $\dim \mathcal{M}(\mathcal{A}) =: n < \infty$. Assume the following ***regularity conditions***:

(1) $\mathcal{M}(\mathcal{A})$ can be parametrised by finite dimensional vectors $\theta = (\theta^1, \ldots, \theta^n) \in \Theta$, where $\Theta \subset \mathbb{R}^n$ is an open subset, using the smooth and injective mapping

$$p : \Theta \ni \theta \mapsto p(\theta) \in \mathcal{M}(\mathcal{A}); \tag{452}$$

(2) $\mathcal{M}(\mathcal{A})$ is represented as $\mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \subseteq L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+$ for a given representation $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ of $\mathcal{A}$, and with all supports $\mathrm{supp}(p) := \{\chi \in \mathcal{X} \mid p(\chi) > 0\}$ equal to $\mathcal{X}$ for all elements $p \in \mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$. This implies that all $p(\chi)\tilde{\mu}(\chi)$ are mutually absolutely continuous and $\mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \subseteq L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})_0^+$;

(3) integration $\int_{\mathcal{X}} \tilde{\mu}(\chi)$ over $\mathcal{X}$ commutes with differentiation $\partial_i := \frac{\partial}{\partial \theta^i}$ over Θ for all functions on $\mathcal{X} \times \Theta$,

$$\int_{\mathcal{X}} \tilde{\mu}(\chi) \frac{\partial}{\partial \theta^i} f(\chi, \theta) = \frac{\partial}{\partial \theta^i} \int_{\mathcal{X}} \tilde{\mu}(\chi) f(\chi, \theta). \tag{453}$$

Under these conditions, $(\theta^1, \ldots, \theta^n)$ is a global coordinate system on $\mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$, which equips the statistical model $\mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ with a smooth manifold structure. Such $\mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ is called a ***parametric statistical manifold***. Its elements are denoted by $p(\chi, \theta)$. If

$$\mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \subseteq L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})_{01}^+ := \{f \in L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})_0^+ \mid \int \tilde{\mu} f = 1\} \tag{454}$$

is a parametric statistical manifold, then it is called a ***parametric probabilistic manifold***.

The tangent space $\mathbf{T}_{p(\theta)} \mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ of the parametric statistical manifold $\mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ at $p(\theta) := p(x, \theta)$ is defined as a vector space spanned by the basis vectors (∂_i). The Nagaoka–Amari γ-embeddings (294),

$$\ell_\gamma(p(\chi, \theta)) =: \ell_\gamma(\chi, \theta) =: \ell_\gamma(\theta), \tag{455}$$

allow to define a suitable family of representations of this tangent space. If one assumes an additional regularity condition,

(4) for any fixed θ the elements of the set $\{\partial_i \ell_\gamma(x, \theta) \mid i \in \{1, \ldots, \dim(\mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}))\}\}$ are linearly independent,

then the ***γ-representation*** [17] of $\mathbf{T}_{p(\theta)} \mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ is defined as the push-forward $\mathbf{T}_{p(\theta)} \ell_\gamma$ of the γ-embeddings. More explicitly, if $v \in \mathbf{T}_{p(\theta)} \mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ and $c :]-\varepsilon, \varepsilon[\ni t \mapsto c(t) \in \mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ such that $c(0) = p(\theta)$ is a curve in the equivalence class of the vector v, then the push-forward $\mathbf{T}_{p(\theta)} \ell_\gamma$ defines the vector space isomorphism

$$\mathbf{T}_{p(\theta)} \ell_\gamma : \mathbf{T}_{p(\theta)} \mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \to \mathbf{T}_{\ell_\gamma(p(\theta))} L_{1/\gamma}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}), \tag{456}$$

$$v = \sum_{i=1}^n v^i \partial_i \mapsto \frac{\mathrm{d}}{\mathrm{d}t} (\ell_\gamma \circ c)|_{t=0} = \sum_{i=1}^n v^i \partial_i \ell_\gamma(\theta). \tag{457}$$

From (453) and normalisation condition $\int \tilde{\mu} p = 1$ it follows that $\int \tilde{\mu} p(\theta) \partial_i \ell_0(\theta) = 0$ [203], hence one has

$$\begin{aligned} \mathbf{T}_{\ell_0(p)} \mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) &= \{f(\chi) \in L_\infty(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \mid \int \tilde{\mu} p(\theta) \sum_{i=1}^n f^i \partial_i \ell_0(\theta) = 0\} \\ &= \{f(\chi) \in L_\infty(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \mid \int_{\mathcal{X}} \tilde{\mu}(\chi) p(\chi, \theta) f(\chi) = 0\}, \end{aligned} \tag{458}$$

whenever $\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu}) \subseteq L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+_{01}$. If $v = \frac{\partial}{\partial\theta^i}|_q$, then its γ-representation is $\partial_i\ell_\gamma(\theta)$. But since

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{\gamma}p^\gamma\right) = p^\gamma\frac{\mathrm{d}(\log p)}{\mathrm{d}t}, \tag{459}$$

it can be written as $p^\gamma\frac{\partial(\log p)}{\partial\theta^i} = p^\gamma\partial_i\ell_0(\theta)$. Thus,

$$\mathbf{T}_{\ell_\gamma(p)}\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu}) = \{f \in L_{1/\gamma}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu}) \mid \int\tilde{\mu}pf = 0\}. \tag{460}$$

For the case when $\dim\mathcal{M}(\mathcal{A}) = \infty$, the construction based on smooth embeddings into $\mathbb{R}^n$ space is no longer available. Instead, one needs to introduce smooth Banach manifold structure on the models $\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$, using smooth embeddings of the neighbourhoods of points of $\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$ into suitable Banach spaces. Chencov [147, 151, 152] proposed to define a general notion of statistical manifold as an arbitrary statistical model satisfying regularity condition (2), but without assuming parametric finite dimensionality, and using the countably additive ideal instead of choosing specific reference measure. Dawid [203] and Koshevnik and Levit [401] proposed to define a tangent space of $\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu}) \subseteq L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+_{01}$ by

$$\mathbf{T}_p\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu}) := \{f \in L_2(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu}) \mid \int\tilde{\mu}pf = 0\}. \tag{461}$$

However, these ideas do not solve the problem of construction of smooth structure on $\mathcal{M}(\mathcal{A})$. The space (461) does not provide such structure, and neither $L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$ nor $L_\infty(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$ does. The main obstacle encountered when trying to use (461) is the fact that the set of strictly positive elements of its unit sphere[43] has an empty interior, so while it is a riemannian manifold, an embedding $q \mapsto \ell_{1/2}(q) = 2\sqrt{q}$ into it does not define a smooth atlas (however, see [126], where the mapping $q \mapsto \sqrt{\frac{q}{p}} - \int\tilde{\mu}p\sqrt{\frac{q}{p}}$ is used instead).

Pistone and Sempi [594] proposed to consider a suitably defined Orlicz space. Given a localisable measure space $(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$, an ***Orlicz space*** is defined as [554, 555]

$$L_\Upsilon(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu}) := \{f \in L_0(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu};\mathbb{R}\cup\{+\infty\}) \mid \exists\lambda > 0 \int_\mathcal{X}\tilde{\mu}(x)\Upsilon(\lambda f(x)) < \infty\}, \tag{462}$$

where a ***Young function*** [81, 782] $\Upsilon : \mathbb{R} \to \mathbb{R}^+ \cup \{+\infty\}$ is defined by the following conditions

1) $\Upsilon(0) = 0$,
2) $\Upsilon(t) = \Upsilon(-t)\ \forall t \in \mathbb{R}$,
3) $\lim_{t\to+\infty}\Upsilon(t) = +\infty$.

An Orlicz space is a Banach space under several equivalent norms, including the Morse–Transue–Nakano–Luxemburg norm [508, 527, 472]

$$\|f\|_{\Upsilon,\tilde{\mu}} := \inf\{\lambda > 0 \mid \int\tilde{\mu}\Upsilon(\lambda^{-1}f) \leq 1\}. \tag{463}$$

For a detailed treatment of these spaces, see [527, 408, 460, 516, 480, 604, 605]. For any $p \in [1,\infty]$ the $L_p(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$ space is an Orlicz space determined by the Young function $\frac{1}{p}|t|^p$. An Orlicz space determined by the Young function

$$\Upsilon_1(t) := \cosh(t) - 1 = \frac{\mathrm{e}^t + \mathrm{e}^{-t}}{2} - 1 \tag{464}$$

[43] Because every infinite dimensional Hilbert space is diffeomorphic with its unit sphere [70], the same problem is encountered even if normalisation is dropped.

is a nonseparable Banach space. Now, for any element $p \in L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+_{01}$ consider the sets

$$C(p, \tilde{\mu}) := \{f \in L_0(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \mid \exists \epsilon > 0 \; \forall \lambda \in [-\epsilon, \epsilon] \; Z(p, \lambda f) < \infty\}, \tag{465}$$

$$C_0(p, \tilde{\mu}) := \{f \in C(p, \tilde{\mu}) \mid \int \tilde{\mu} p f = 0\}, \tag{466}$$

where

$$Z(p, f) := \int \tilde{\mu} p e^f. \tag{467}$$

(The difference in sign between equations (467) and (591) is because different sign for θ_i is used in (591).) The condition defining $C(p, \tilde{\mu})$ implies that each $f \in C(p, \tilde{\mu})$ satisfies $\int \tilde{\mu} p f \in \mathbb{R}$. Pistone and Rogantin [593] show that

$$\mathrm{span}_{\mathbb{R}} C(p, \tilde{\mu}) = L_{\Upsilon_1}(\mathcal{X}, \mho(\mathcal{X}), p\tilde{\mu}), \tag{468}$$

while [594] show that the MTNL norms $\|\cdot\|_{\Upsilon_1, p\tilde{\mu}}$ and $\|\cdot\|_{\Upsilon_1, q\tilde{\mu}}$ are equivalent for any $f \in C(p, \tilde{\mu}) \cap C(q, \tilde{\mu})$. Hence, $f \in L_{\Upsilon_1}(\mathcal{X}, \mho(\mathcal{X}), p\tilde{\mu})$ is equivalent to existence and finiteness of the moment generating function $\lambda \mapsto Z(p, \lambda f)$ on a neighbourhood of 0. As a result, a tangent space of $L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+_{01}$ at p can be identified with the closed set $C_0(p, \tilde{\mu})$, which parametrises the neighbourhood of p. The probability model $L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+_{01}$ is then equipped with the Banach smooth manifold structure modelled on subsets $C_0(p, \tilde{\mu})$ of the Orlicz space $L_{\Upsilon_1}(\mathcal{X}, \mho(\mathcal{X}), p\tilde{\mu})$ by means of the smooth embeddings given by diffeomorphisms

$$w_p^{-1} : L_{\Upsilon_1}(\mathcal{X}, \mho(\mathcal{X}), p\tilde{\mu}) \supsetneq C_0(p, \tilde{\mu}) \supset U(p) \ni f \mapsto p e^{f - \log(Z(p,f))} \in L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+_{01}, \tag{469}$$

where $p \in L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+_{01}$ and $U(p)$ is an intersection of an open unit ball of $C(p, \tilde{\mu})$ with $C_0(p, \tilde{\mu})$, $U(p) := \{f \in C_0(p, \tilde{\mu}) \mid \|f\|_{\Upsilon_1, p\tilde{\mu}} < 1\}$. The bijectivity of w_p^{-1} follows from the fact that $p e^{f_1 - \log Z(p, f_1)} = p e^{f_2 - \log Z(p, f_2)} \Rightarrow f_1 - f_2 = \mathrm{const}$, and 0 is the only constant element of $U(p)$. The set $C(p, \tilde{\mu})$ can be considered as a nonparametric exponential family at $p\tilde{\mu}$ (see Section 5.3). The restriction of domain of the mapping (469) from $C(p, \tilde{\mu})$ to $C_0(p, \tilde{\mu})$ is required due to normalisation of probability densities and is provided in order to turn this map into a bijection.[44] The inverse of (469) reads

$$w_p : L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+_{01} \ni q \mapsto \log\left(\frac{q}{p}\right) - \int \tilde{\mu} p \log\left(\frac{q}{p}\right) = \log\left(\frac{q}{p}\right) + D_1(p, q) \in C_0(p, \tilde{\mu}), \tag{470}$$

and the maps

$$w_q \circ w_p^{-1}(U(p) \cap U(q)) \ni u \mapsto u + \log\left(\frac{p}{q}\right) - \int \tilde{\mu} q \left(u + \log\left(\frac{p}{q}\right)\right) \in w_q(U(p) \cap U(q)) \tag{471}$$

are smooth and their domains are open sets. The set $\{(w_p^{-1}(U(p)), w_p) \mid p \in L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+_{01}\}$ is a smooth atlas on $L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+_{01}$. Each chart of this atlas is defined using different Banach space, but all of them are isometrically isomorphic [142]. This approach was further developed in [280, 593, 142, 143] (see [590, 591, 592] for recent overviews). An extension of this construction from $L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+_{01}$ to $L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+_0$ was provided in [46]. The tangent space of $p \in L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+_0$ is in such case modelled by $L_{\Upsilon_1}(\mathcal{X}, \mho(\mathcal{X}), p\tilde{\mu})$ instead of $C_0(p, \tilde{\mu})$, but otherwise without the form of the mapping (469). On the other hand, the right hand side of (470) changes, because $D_1(p, q) = \log Z(p, f)$ only at $C_0(p, \tilde{\mu})$. As a result, we can define the ***nonparametric statistical manifold*** as a statistical model $\mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \subseteq L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+_0$ which is a Banach smooth submanifold of $L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+_0$.

[44] In [143] the improved construction is provided, with $U(p)$ replaced by $\widetilde{U}(p) := \mathrm{int}(\mathrm{efd}(\log Z(p, \cdot)))$. In such case, the image $w_p^{-1}(\widetilde{U}(p))$ is called a ***maximal exponential model*** at p. Pistone and Sempi [594] showed that $Z(p, \cdot)$ is Gâteaux differentiable on $\mathrm{int}(\mathrm{efd}(Z(p, \cdot)))$ (for a proof that is analytic on this set, see [766]), and that it is Fréchet differentiable on $\{f \in L_{\Upsilon_1}(\mathcal{X}, \mho(\mathcal{X}), p\tilde{\mu}) \mid \|f\|_{\Upsilon_1, p\tilde{\mu}} < 1\}$, while $\log Z(p, \cdot)$ is analytic on $\mathrm{efd}(\log Z(p, \cdot)) \cap C_0(p, \tilde{\mu})$.

Following the idea of Pistone [589], Grasselli [286, 287, 289] developed a modification of the Pistone–Sempi approach, with $L_{\Upsilon_1}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ and $C_0(p, \tilde{\mu})$ in (469) replaced, respectively, by

$$B_{\Upsilon_1}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) := \{f \in L_0(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}; \mathbb{R} \cup \{+\infty\}) \mid \forall \lambda \in \mathbb{R} \int_{\mathcal{X}} \tilde{\mu}(x)\Upsilon_1(\lambda f(x)) < \infty\}, \tag{472}$$

$$B^0_{\Upsilon_1}(p, \tilde{\mu}) := \{f \in B_{\Upsilon_1}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \mid \int \tilde{\mu} p f = 0\}. \tag{473}$$

The space $B_{\Upsilon_1}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ coincides with the closure of $L_\infty(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ in the topology generated by the norm $\|\cdot\|_{\Upsilon_1, \tilde{\mu}}$. The condition in (472) reads explicitly

$$\int \tilde{\mu} p \left(\frac{\mathrm{e}^{\lambda f} + \mathrm{e}^{-\lambda f}}{2} - 1 \right) < \infty \quad \forall \lambda \in \mathbb{R}, \tag{474}$$

which implies that $Z(p, f) \in \mathbb{R}\ \forall f \in B^0_{\Upsilon_1}(p, \tilde{\mu})$.

The finite dimensional parametric quantum manifolds are constructed in the way analogous to the commutative parametric case. The standard approach is based on the choice of a *finite dimensional* Hilbert space $\mathcal{H}$ and some parametric family $\theta \mapsto \rho(\theta)$ of invertible positive trace class operators that act on $\mathcal{H}$, forming the subset

$$\mathcal{M}(\mathcal{H}, \Theta) \subseteq \{\rho(\theta) \in \mathfrak{G}_1(\mathcal{H})^{\mathrm{sa}} \mid \rho(\theta) > 0,\ \theta \in \Theta \subseteq \mathbb{R}^m \text{ open},\ \dim \mathcal{H} =: n < \infty\} \subseteq \mathfrak{G}_1(\mathbb{C}^n)^+_0 \cong \mathrm{M}_n(\mathbb{C})^+_0. \tag{475}$$

Usually, the additional condition $\mathrm{tr}(\rho(\theta)) = 1$ is imposed on the elements of $\mathcal{M}(\mathcal{H}, \Theta)$. From the algebraic perspective, this construction is a choice of a particular 'dominating' $\omega \in \mathcal{N}^+_\star$ on a W^*-algebra $\mathcal{N} \cong \mathfrak{B}(\mathcal{H})$, and the choice of a subset $\mathcal{M}(\mathcal{N}, \omega)$ of the faithful normal algebraic states that belong to the *folium* of ω, which means that the space $\mathcal{M}(\mathcal{N}, \omega)$ can be represented as a parametric family (475) with $\mathcal{H}$ given by the GNS Hilbert space $\mathcal{H}_\omega$ of ω. Yet, the condition $\dim \mathcal{H} < \infty$ restricts the considerations to finite dimensional W^*-algebras $\mathfrak{B}(\mathcal{H})$ of type I_n, with $\mathcal{M}(\mathcal{H}) \subseteq \mathfrak{B}(\mathcal{H})^+_{\star 0} \cong \mathrm{M}_n(\mathbb{C})^+_0$, which inherits the structure of smooth manifold as an open subset of the real vector space $\mathrm{M}_n(\mathbb{C})^{\mathrm{sa}}$. Generalisation of this construction to semi-finite W^*-algebras $\mathcal{N}$ requires replacing selection of the subset of a folium of a given algebraic state by selection of the set of states on $\mathcal{N}$ that satisfy the noncommutative analogue of the Radon–Nikodým theorem with respect to a given faithful normal semi-finite trace τ on $\mathcal{N}^+$. Hence, for a given semi-finite W^*-algebra $\mathcal{N}$ and a faithful normal semi-finite trace τ on $\mathcal{N}$, we define the ***parametric quantum manifold*** $\mathcal{M}(\mathcal{N}, \tau, \Theta)$ as an image of a smooth map

$$\rho : \mathbb{R}^m \supsetneq \Theta \ni \theta \mapsto \rho(\theta) \in L_1(\mathcal{N}, \tau)^+_0 \subseteq \mathscr{M}(\mathcal{N}, \tau) \tag{476}$$

on an open set $\Theta \subsetneq \mathbb{R}^m$, $m \in \mathbb{N}$. Each $\rho(\theta)$ is a Dye–Segal density of τ with respect to some element of $\mathcal{N}^+_{\star 0}$. Note that the above definition does not encapsulate the analogue of the condition (453), so the relationship between differentiation $\frac{\partial}{\partial \theta^i}$ and integration $\tau(\cdot)$, required for an explicit representation of the tangent space of $\mathcal{M}(\mathcal{N}, \tau, \Theta)$ as a space of operators, remains to be clarified in further applications of this definition.

A tangent space $\mathbf{T}_\rho \mathrm{M}_n(\mathbb{C})^+_0$ is the real vector space of all Fréchet derivatives in the directions of smooth curves in $\mathrm{M}_n(\mathbb{C})^+_0$ that pass through ρ, so it can be identified with a restriction of $\mathrm{M}_n(\mathbb{C})^{\mathrm{sa}}$. A restriction of domain of ρ to $\mathrm{M}_n(\mathbb{C})^+_{01}$ implies a restriction of the tangent vectors to the space $\{x \in \mathrm{M}_n(\mathbb{C})^{\mathrm{sa}} \mid \mathrm{tr}(x) = 0\}$. Constructions of the γ-representations for the parametric quantum manifolds $\mathcal{M}(\mathcal{H}, \Theta)$ are, as in the commutative case, based on push-forwards of the γ-embeddings (295). In the case of $\gamma \in]0, 1]$, the codomains of γ-embeddings are $\mathfrak{G}_{1/\gamma}(\mathcal{H})$ spaces. In the case $\gamma = 0$ the mapping

$$\ell_0 \equiv \log : \mathrm{M}_n(\mathbb{C})^+_0 \ni \rho \mapsto \log \rho \in \mathrm{M}_n(\mathbb{C})^{\mathrm{sa}} \tag{477}$$

is a diffeomorphism, allowing to identify $\mathrm{M}_n(\mathbb{C})^+_0$ with $\mathrm{M}_n(\mathbb{C})^{\mathrm{sa}}$. In particular, any n-dimensional submanifold $\mathcal{Q}_n$ of $\mathrm{M}_n(\mathbb{C})^{\mathrm{sa}}$ corresponds to n-dimensional submanifold $\exp(\mathcal{Q}_n) =: \mathcal{M}(\mathrm{M}_n(\mathbb{C}), \Theta) \subseteq \mathrm{M}_n(\mathbb{C})^+_0$ for some $\Theta \subseteq \mathbb{R}^n$. If

$$H : \mathbb{R}^n \supseteq \mathcal{O} \ni (x^1, \ldots, x^n) =: x \mapsto H(x) \in \mathcal{U} \subseteq \mathcal{Q}_n \tag{478}$$

is a diffeomorphism of open subsets $\mathcal{O} \subset \mathbb{R}^n$ and $\mathcal{U} \subseteq \mathcal{Q}_n$, then the inverse map H^{-1} is a coordinate system on $\mathcal{U}$.

Any $x \in \mathrm{M}_n(\mathbb{C})^{\mathrm{sa}}$ can be decomposed as $x = \tilde{x} + \mathrm{i}[\rho, H]$ for $[\tilde{x}, \rho] = 0$ and $H \in \mathrm{M}_n(\mathbb{C})^{\mathrm{sa}}$. If $f \in \mathrm{C}(]0, \infty[)$, then [72]

$$\frac{\mathrm{d}}{\mathrm{d}\rho} f(x) = f'(\rho)\tilde{x} + \mathrm{i}[f(\rho), H], \tag{479}$$

where f' is a derivative of f. This follows from

$$\frac{\mathrm{d}}{\mathrm{d}t}|_{t=0} f(x + ty + \mathrm{i}t[x, z]) = yf'(x) + \mathrm{i}[f(x), z] \quad \forall x, y, z \in \mathrm{M}_n(\mathbb{C})^{\mathrm{sa}}, \tag{480}$$

which holds if $[x, y] = 0$. An infinitesimal transformation $\rho \mapsto \rho + \eth\rho$ can be decomposed as [313]

$$\eth\rho := \widetilde{\eth}\rho + [\rho, W] = \sum_{i=1}^{n} \left(\frac{\partial \rho(\theta)}{\partial \theta^i} + [\rho, W_i] \right) \mathrm{d}\theta^i, \tag{481}$$

where $\widetilde{\eth}\rho = \sum_{i=1}^{n} \frac{\partial \rho(\theta)}{\partial \theta^i} \mathrm{d}\theta^i$ is defined by $[\widetilde{\eth}\rho, \rho] = 0$, while $W = \sum_{i=1}^{n} W_i \mathrm{d}\theta^i$ is an antiself-adjoint operator. The mappings $\eth$, $\widetilde{\eth}$ and $[\,\cdot\,, W]$ are derivations on $\mathfrak{B}(\mathcal{H})$. This determines a decomposition of tangent space at ρ into the direct product of the corresponding subspaces, and allows to write the ***γ-representation*** of $v = \sum_{i=1}^{n} v_i \frac{\partial}{\partial \theta^i} \in \mathbf{T}_\rho \mathfrak{B}(\mathcal{H})^+_{\star 01}$, corresponding to (295), as [313, 314]

$$\mathbf{T}_\rho \ell_\gamma(v) = \begin{cases} \sum_{i=1}^{n} v_i \left(\rho^\gamma \frac{\partial \log \rho}{\partial \theta^i} + \frac{1}{\gamma}[\rho^\gamma, W_i] \right) & : \ \gamma \notin \{0, 1\} \\ \sum_{i=1}^{n} v_i \frac{\partial \log(\rho(\theta))}{\partial \theta^i} & : \ \gamma = 0 \\ \sum_{i=1}^{n} v_i \frac{\partial \rho(\theta)}{\partial \theta^i} & : \ \gamma = 1. \end{cases} \tag{482}$$

Note that $\mathbf{T}_\rho \ell_\gamma(x) = \left(\mathfrak{D}^{\mathrm{F}}_\rho (\ell_\gamma(\rho)) \right)(x)$. For $\gamma \in \{0, 1\}$ the codomain spaces of these mappings are given by

$$\mathbf{T}_\rho \ell_0 : \mathbf{T}_\rho \mathfrak{G}_1(\mathcal{H})^+_{01} \to \{x \in \mathfrak{B}(\mathcal{H})^{\mathrm{sa}} \mid \mathrm{tr}(\rho x) = 0\}, \tag{483}$$
$$\mathbf{T}_\rho \ell_1 : \mathbf{T}_\rho \mathfrak{G}_1(\mathcal{H})^+_{01} \to \{x \in \mathfrak{B}(\mathcal{H})^{\mathrm{sa}} \mid \mathrm{tr}(x) = 0\}. \tag{484}$$

The Banach smooth manifold structure on the set[45]

$$\mathfrak{B}(\mathcal{H})^+_{\star 01} \cap \left(\bigcup_{0<p<1} \mathfrak{G}_p(\mathcal{H}) \right) \tag{485}$$

for $\dim \mathcal{H} = \infty$ was introduced and studied by Streater [698, 696, 697, 699, 700, 701, 704, 703, 702] (in partial collaboration with Grasselli [290, 286]). He used the Rellich–Kato theory [608, 609, 610, 611, 384, 385] of perturbation of operators by *semi-bounded* forms. All variants of this construction depend on the choice of an underlying Hilbert space representation. An alternative Banach smooth manifold structure was introduced by Jenčová [361, 362] on $\mathcal{N}^+_{\star 01}$ for an arbitrary W^*-algebra $\mathcal{N}$. She used the Araki–Donald theory [33, 39, 38, 41, 230, 232, 233] of relative entropic perturbations of quantum states by *bounded* self-adjoint operators (see Section 3.4). This way Jenčová's approach follows the Pistone–Grasselli approach, while Streater's approach follows the Pistone–Sempi approach.[46] In both cases the central role is played by suitably defined noncommutative analogue of an Orlicz space, such that the resulting quantum manifolds are quantum models $\mathcal{M}(\mathcal{N})$ with local neighbourhood of any quantum state $\phi \in \mathcal{M}(\mathcal{N})$ consisting only of such quantum states which have finite Araki distance to ϕ. Streater's approach is stronger, because his quantum manifolds also make the absolute von Neumann entropy (601) finite, but the price paid is the restriction of the class of states and class of algebras under consideration. We consider Jenčová's approach to be more suitable as a *general* setting

[45] If $0 < p < 1$, then $\mathfrak{G}_p(\mathcal{H})$ space is defined as a subset of all $x \in \mathfrak{G}_0(\mathcal{H})$ such that $|x|^p \in \mathfrak{G}_1(\mathcal{H})$.

[46] The W^*-algebras $\mathcal{N}$ which are not countably finite do not allow faithful quantum states ($\mathcal{N}^+_{\star 0} = \varnothing$), so one cannot introduce any of the above Banach smooth manifold structures on quantum models $\mathcal{M}(\mathcal{N})$ over such algebras.

for quantum manifolds, because it allows to deal with arbitrary countably finite W^*-algebras, and is manifestly Hilbert space independent.

Jenčová's approach [361] starts from a generalisation of the notion of ***Young function*** to such $\Upsilon : X \to \mathbb{R}^+ \cup \{+\infty\}$ over a real Banach space X, that

i) $\Upsilon(0) = 0$,

ii) $\Upsilon(x) = \Upsilon(-x)\ \forall x \in X$,

iii) $x \neq 0 \Rightarrow \lim_{t\to+\infty} \Upsilon(tx) = +\infty$.

If Υ is also convex and lower semi-continuous, then the corresponding Orlicz space is defined as a Banach space $B_\Upsilon(X)$ arising from the Cauchy completion of the real vector space

$$L_\Upsilon(X) := \{f \in X \mid \exists\lambda > 0\ \Upsilon(\lambda f) < \infty\} \tag{486}$$

in the norm

$$\|f\|_\Upsilon := \inf\{\lambda > 0 \mid \Upsilon(\lambda^{-1} f) \leq 1\}. \tag{487}$$

For an arbitrary W^*-algebra $\mathcal{N}$ and $\phi \in \mathcal{N}^+_\star$, Jenčová considers the Young function on $\mathcal{N}^{\rm sa}$ given by

$$\Upsilon_\phi : \mathcal{N}^{\rm sa} \ni h \mapsto \Upsilon_\phi(h) := \frac{\widetilde{\phi^h}(\mathbb{I}) + \widetilde{\phi^{-h}}(\mathbb{I})}{2} - 1 = \frac{\mathrm{e}^{-c(\phi,-h)} + \mathrm{e}^{-c(\phi,h)}}{2} - 1 \in \mathbb{R}^+, \tag{488}$$

as well as the Young function on $\{x \in \mathcal{N}^{\rm sa} \mid \phi(x) = 0\}$,

$$\widetilde{\Upsilon}_\phi : \{x \in \mathcal{N}^{\rm sa} \mid \phi(x) = 0\} \ni h \mapsto -\frac{c(\phi,-h) + c(\phi,h)}{2} = -\frac{D_1|_{\mathcal{N}^+_{\star 1}}(\phi,\phi^h) + D_1|_{\mathcal{N}^+_{\star 1}}(\phi,\phi^{-h})}{2} \in \mathbb{R}. \tag{489}$$

The function Υ_ϕ coincides for $\mathcal{N} = \mathfrak{B}(\mathcal{H})$ with the Young function introduced by Streater in [698, 700]. The analogues of (472) and (473) are defined, respectively, as

$$B_{\Upsilon_\phi}(\mathcal{N}) := \overline{L_{\Upsilon_\phi}(\mathcal{N})}^{\|\cdot\|_{\Upsilon_\phi}}, \tag{490}$$

$$B^0_{\Upsilon_\phi}(\mathcal{N}) := \{x \in B_{\Upsilon_\phi}(\mathcal{N}) \mid \phi(x) = 0\}. \tag{491}$$

The space $B_{\Upsilon_\phi}(\mathcal{N})$ is an example of a noncommutative Orlicz space. (For other approaches to the theory of noncommutative Orlicz spaces see [511, 512, 228, 420, 4, 424, **?**, 5, 47, **?**, 645].) The norms $\|\cdot\|_{\Upsilon_\phi}$ and $\|\cdot\|_{\widetilde{\Upsilon}_\phi}$ are equivalent on $\{x \in \mathcal{N}^{\rm sa} \mid \phi(x) = 0\}$, the inclusion $\mathcal{N}^{\rm sa} \subseteq B_{\Upsilon_\phi}(\mathcal{N})$ is continuous, and

$$B^0_{\widetilde{\Upsilon}_\phi}(\mathcal{N}) := \{x \in B_{\widetilde{\Upsilon}_\phi} \mid \phi(x) = 0\} = \overline{\{x \in \mathcal{N}^{\rm sa} \mid \phi(x) = 0\}}^{\|\cdot\|_{\Upsilon_\phi}} = B^0_{\Upsilon_\phi}(\mathcal{N}). \tag{492}$$

Jenčová proves that the quantum model $\mathcal{N}^+_{\star 01}$ can be equipped with the smooth Banach manifold structure modelled on closed subsets $B^0_{\Upsilon_\phi}(\mathcal{N})$ of Banach spaces $B_{\Upsilon_\phi}(\mathcal{N})$ by means of the smooth embeddings given by inverses of diffeomorphisms

$$w_\phi^{-1} : B^0_{\Upsilon_\phi}(\mathcal{N}) \supseteq U(\phi) \ni h \mapsto \phi^h \in \mathcal{N}^+_{\star 01}, \tag{493}$$

where $\phi \in \mathcal{N}^+_{\star 01}$ and $U(\phi) := \{x \in B^0_{\Upsilon_\phi}(\mathcal{N}) \mid \|x\|_{\Upsilon_\phi} < 1\}$ (i.e., $U(\phi)$ is an open unit ball of $B^0_{\Upsilon_\phi}(\mathcal{N})$). The set $\{(w_\phi^{-1}(U(\phi)), w_\phi) \mid \phi \in \mathcal{N}^+_{\star 01}\}$ is a smooth atlas on $\mathcal{N}^+_{\star 01}$. This provides a quantum generalisation of the Pistone–Grasselli manifold structure, and coincides with it for $\mathcal{N} = L_\infty(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$. We conjecture that Jenčová's construction can be extended to $\mathcal{N}^+_{\star 0}$ analogously to the commutative case, by replacing $B^0_{\Upsilon_\phi}$ by B_{Υ_ϕ} in (493) and in $U(\phi)$. Under this conjecture, we define a ***nonparametric quantum manifold*** as a quantum model $\mathcal{M}(\mathcal{N}) \subseteq \mathcal{N}^+_{\star 0}$ that is a Banach smooth submanifold of $\mathcal{N}^+_{\star 0}$.

In [362] it is noted that the proper quantum analogue of the Pistone–Sempi manifold structure could be provided by an extension of the map w_ϕ^{-1} to the space $\{x \in (B^0_{\Upsilon_\phi}(\mathcal{N}))^{\star\star} \mid \|x\| < 1\}$. We propose to provide it by defining a vector space

$$\mathcal{N}^{\pm\rm ext}_\phi := \{x \in \mathrm{span}_{\mathbb{R}}(\mathcal{N}^{\rm ext}) \mid \phi(x) \in \mathbb{R}\}, \tag{494}$$

considering the extension of the domain of Υ in (488) to $\mathcal{N}_\phi^{\pm\mathrm{ext}}$, on which (488) is a pseudomodular function (its convexity follows from joint convexity of $D_1|_{\mathcal{N}^+_{\star 1}}(\omega,\phi)$). This allows us to define

$$(\mathcal{N}_\phi^{\pm\mathrm{ext}})_{\Upsilon_\phi} := \{h \in \mathcal{N}_\phi^{\pm\mathrm{ext}} \mid \lim_{\lambda\to 0}\Upsilon_\phi(\lambda h) = 0\}, \tag{495}$$

$$L_{\Upsilon_\phi}(\mathcal{N}_\phi^{\pm\mathrm{ext}}) := \overline{(\mathcal{N}_\phi^{\pm\mathrm{ext}})_{\Upsilon_\phi}}^{\|\cdot\|_{\Upsilon_\phi}}, \tag{496}$$

$$U(\phi) := \{x \in L_{\Upsilon_\phi}(\mathcal{N}_\phi^{\pm\mathrm{ext}}) \mid \|\cdot\|_{\Upsilon_\phi} < 1\}, \tag{497}$$

$$w_\phi^{-1} : L_{\Upsilon_\phi}(\mathcal{N}_\phi^{\pm\mathrm{ext}}) \supseteq U(\phi) \ni h \mapsto \phi^h \in \mathcal{N}^+_{\star 0}. \tag{498}$$

We use here open unit ball $U(\phi)$ instead of $\widetilde{U}(\phi) := \mathrm{int}(\mathrm{efd}(c(\phi,\cdot)))$, because from $\phi(x) \geq c(\phi,x)$ $\forall x \in \mathcal{N}^{\mathrm{ext}}$ it follows that $\mathcal{N}_\phi^{\pm\mathrm{ext}} \subseteq \mathrm{efd}(c(\phi,\cdot))$, so $\mathrm{int}(\mathrm{efd}(c(\phi,\cdot)))$ may be too large for our purposes. Note that $\mathcal{N}_\phi^{\pm\mathrm{ext}}$ is *not* a Banach space. We will discuss this generalisation in full detail elsewhere. The main reason why we will not use our generalisation of the noncommutative Orlicz spaces, $L_\Upsilon(\mathcal{N})$, for the above purpose is that their elements belong to the space $\mathscr{M}(\widetilde{\mathcal{N}},\widetilde{\tau})$, while the strict results for D_1 projections are known for the elements of $\mathcal{N}^{\mathrm{ext}}$. In general, $\mathscr{M}(\mathcal{N},\tau) \subseteq \mathrm{aff}(\mathcal{N})$, $\mathscr{M}(\widetilde{\mathcal{N}},\widetilde{\tau}) \subseteq \mathrm{aff}(\widetilde{\mathcal{N}})$, and $\mathrm{aff}(\mathcal{N}) \subseteq \mathcal{N}^{\mathrm{ext}}$. So, without establishing more direct relationship between $\mathcal{N}^{\mathrm{ext}}$ and $\mathscr{M}(\widetilde{\mathcal{N}},\widetilde{\tau})$, or without restriction to semi-finite algebras with some choice of a faithful normal semifinite trace τ and restriction of allowed vectors from $\mathcal{N}^{\mathrm{ext}}$ to $\mathscr{M}(\mathcal{N},\tau)$, it is unclear how to apply the spaces $L_\Upsilon(\mathcal{N})$ in this specific case. Once again, this shows the very specific character of D_1 distance and its projections.

5.2 Smooth $\mathfrak{f}$- and γ-geometries

A ***riemannian quantum manifold*** is defined as a riemannian manifold $(\mathcal{M}(\mathcal{N}),\mathbf{g})$, where $\mathcal{M}(\mathcal{N})$ is a quantum manifold. A ***riemannian statistical manifold*** is defined as a riemannian manifold $(\mathcal{M}(\mathcal{A}),\mathbf{g})$, where $\mathcal{M}(\mathcal{A})$ is a statistical manifold. For early discussions of various finite dimensional riemannian statistical manifolds, see [152, 647, 14, 45, 124, 125].

A ***Fisher matrix*** [255] on a finite dimensional statistical manifold $\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$,

$$\begin{aligned}\mathbf{g}^{\mathrm{FRJ}}_{ij}(\theta) :&= \int_{\mathcal{X}} \tilde{\mu}(\chi)p(\chi,\theta)\frac{\partial}{\partial\theta^i}\log p(\chi,\theta)\frac{\partial}{\partial\theta^j}\log p(\chi,\theta) = \int_{\mathcal{X}}\tilde{\mu}(\chi)\partial_i\ell_\gamma(\chi,\theta)\partial_j\ell_{1-\gamma}(\chi,\theta)\\ &= \int_{\mathcal{X}}\tilde{\mu}(\chi)p(\chi,\theta)\partial_i\ell_0(\chi,\theta)\partial_j\ell_0(\chi,\theta) = -\int_{\mathcal{X}}\tilde{\mu}(\chi)p(\chi,\theta)\partial_i\partial_j\ell_0(\chi,\theta),\end{aligned} \tag{499}$$

defines an inner product

$$\mathbf{g}^{\mathrm{FRJ}}_{\theta(p)}(\cdot,\cdot) : \mathbf{T}_{\theta(p)}\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu}) \times \mathbf{T}_{\theta(p)}\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu}) \ni (\partial_i,\partial_j) \mapsto \mathbf{g}^{\mathrm{FRJ}}_{ij}(\theta) \in \mathbb{R}, \tag{500}$$

which is symmetric ($\mathbf{g}^{\mathrm{FRJ}}_{ij}(\theta) = \mathbf{g}^{\mathrm{FRJ}}_{ji}(\theta)$) and positive definite,

$$\mathbf{g}^{\mathrm{FRJ}}_{\theta(p)}(u,u) = \sum_{i=1}^n\sum_{j=1}^n \mathbf{g}^{\mathrm{FRJ}}_{ij}(\theta)u^iu^j > 0 \;\forall u \neq 0. \tag{501}$$

The last property follows from the regularity condition (4). Hence, as Rao [601, 602] and Jeffreys [351] have independently observed, (500) defines a riemannian metric

$$\mathbf{g}^{\mathrm{FRJ}} : \mathbf{T}\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu}) \times \mathbf{T}\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu}) \to \mathbb{R}, \tag{502}$$

which will be called the ***Fisher–Rao–Jeffreys metric***. In consequence, $(\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu}),\mathbf{g})$ is a riemannian manifold. Further study of (502) as a riemannian metric was carried by Kozlov [406]. A riemannian distance $d_{\mathbf{g}^{\mathrm{FRJ}}}$ of $\mathbf{g}^{\mathrm{FRJ}}$ is given by

$$d_{\mathbf{g}^{\mathrm{FRJ}}}(p,q) = 2D_{\mathrm{KH}}(p,q) \tag{503}$$

for $p,q \in L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+_0$, and by $2\arccos(D_{\mathrm{B}}(p,q))$ for $p,q \in L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+_{01}$.

The Christoffel symbols with respect to the FRJ metric given by

$$\begin{aligned}\mathbf{g}^{\mathrm{FRJ}}_{\theta(p)}(\nabla^{\gamma}_{\partial_i}\partial_j,\partial_k)=\Gamma^{\gamma}_{ijk}(\theta)&=\int\tilde{\mu}p(\theta)\partial_i\partial_j\ell_{\gamma}(\theta)\partial_k\ell_{1-\gamma}(\theta)\\&=\int\tilde{\mu}p(\theta)\left(\partial_i\partial_j\ell_0(\theta)+\gamma\partial_i\ell_0(\theta)\partial_j\ell_0(\theta)\partial_k\ell_0(\theta)\right),\end{aligned}\tag{504}$$

determine the family $\{\nabla^{\gamma}\mid\gamma\in\mathbb{R}\}$ of affine connections on $\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$, called the ***Chencov–Amari*** $\boldsymbol{\gamma}$***-connections*** [152, 14].[47] The Levi-Civita connection of the FRJ metric is equal to $\nabla^{1/2}$. The Chencov–Amari γ-connections are torsion free ($\mathfrak{T}^{\gamma}_{ijk}=0$) [17, 18, 505, 506] and Norden–Sen dual with respect to the FRJ metric: $(\nabla^{\gamma})^{\dagger}=\nabla^{1-\gamma}$ [524]. The connection ∇^{γ} is flat on $L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^{+}_{0}$ for every $\gamma\in[0,1]$, but it is flat on $L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^{+}_{01}$ only for $\gamma\in\{0,1\}$. If ∇^{γ} is flat, then its dually flat coordinates are given by the Nagaoka–Amari γ-embeddings $(\ell_{\gamma},\ell_{1-\gamma})$ specified by (294). The dually flat geometries $(\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu}),\mathbf{g}^{\mathrm{FRJ}},\nabla^{\gamma},\nabla^{1-\gamma})$ were first considered explicitly in [524] and will be called the ***Chencov–Amari–Nagaoka geometries***. They can be characterised as Lauritzen manifolds $(\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu}),\mathbf{g}^{\mathrm{FRJ}},C)$ for $\vartheta=\gamma$ and for

$$C_{ijk}(\theta)=\int\tilde{\mu}p(\theta)\partial_i\ell_0(\theta)\partial_j\ell_0(\theta)\partial_k\ell_0(\theta).\tag{505}$$

An explicit characterisation of the ∇^{γ}-geodesics was provided in [506]. The relationship between Riemann curvature tensors for $\gamma\in[0,1]$ reads [18, 505, 506]

$$\mathrm{R}^{\nabla^{\gamma}}=4\gamma(1-\gamma)\mathrm{R}^{\nabla^{1/2}}.\tag{506}$$

Chencov [151, 152] (see also [136, 137]) showed that the FRJ metric $\mathbf{g}^{\mathrm{FRJ}}$ is a unique (up to a multiplicative constant) riemannian metric $\mathbf{g}$ on submanifolds $\mathcal{M}(\mathcal{X},\mho(\mathcal{X}))$ of a finite probability simplex $L_1(\mathcal{X},\mho(\mathcal{X}))^{+}_{1}$ such that its riemannian distance $d_{\mathbf{g}}$ is ***monotone*** with respect to finite coarse grainings (that is, with respect to arrows in $\mathbf{ProbMod}_{\mathrm{fin}}(\mathcal{X})$),

$$d_{\mathbf{g}}(p,q)\geq d_{\mathbf{g}}(T_{\star}(p),T_{\star}(q))\quad\forall p,q\in\mathcal{M}(\mathcal{X},\mho(\mathcal{X}))\;\;\forall T_{\star}\in\mathrm{Mark}_{\star}(\mathcal{M}(\mathcal{X},\mho(\mathcal{X}))).\tag{507}$$

He has also shown [152] that the Chencov–Amari γ-connections are unique (up to a multiplicative constant) affine connections ∇ on submanifolds $\mathcal{M}(\mathcal{X},\mho(\mathcal{X}))$ such that the image of each ∇-geodesic line on $\mathcal{M}(\mathcal{X},\mho(\mathcal{X}))$ belongs to a ∇-geodesic line on $T_{\star}(\mathcal{M}(\mathcal{X},\mho(\mathcal{X})))$ as its interval or its point.[48]

Eguchi [243, 244] proved that for arbitrary Csiszár–Morimoto distance $D_{\mathfrak{f}}$, with smooth $\mathfrak{f}$ and on finite dimensional probabilistic manifold $\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})\subseteq L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^{+}_{01}$, the riemannian metric $\mathbf{g}^{\mathfrak{f}}$ associated to $D_{\mathfrak{f}}$ by means of (416) takes the form

$$\mathbf{g}^{\mathfrak{f}}=\mathbf{g}^{\mathrm{FRJ}}\mathfrak{f}''(1),\tag{508}$$

while the $\nabla^{\mathfrak{f}}$-connections derived by means of (417) coincide with ∇^{γ}-connections with

$$1-2\gamma=2\mathfrak{f}'''(1)+3\mathfrak{f}''(1),\tag{509}$$

with $\mathfrak{f}'''$ denoting third derivative of $\mathfrak{f}$. For the higher order Taylor approximations of the Csiszár–Morimoto $\mathfrak{f}$-distance, see [55]. The Eguchi equations (416)-(418) applied to any Liese–Vajda γ-distance

[47] An affine connection ∇^0 was introduced by Chencov [147], following the suggestion of Morozova (see [507]), and it was rediscovered later by Dawid [203, 204]. In [152] Chencov characterised a family (504) of γ-connections, with $\gamma\in\mathbb{R}$, which was rediscovered later by Amari [14, 15]. The expressions for Γ^{γ}_{ijk} have appeared for the first time in the works of Hartigan [309, 310, 311], but without realising their differential geometric meaning.

[48] The more general ***Amari conjecture*** [17], stating that (the scalar multiples of) FRJ metric and the Chencov–Amari γ-connections are the unique riemannian metric and unique affine connections which are invariant under *any* coordinate transformations of the sample space $\mathcal{X}$ and the parameter space Θ, still waits for a proof, although there are some partial results for continuous $\mathcal{X}$, see [586, 23]. Campbell [137] has shown that the extension of $\mathbf{g}^{\mathrm{FRJ}}$ from $L_1(\mathcal{A})^{+}_{1}$ to $L_1(\mathcal{A})^{+}$ is not unique under Markov morphisms. Zhu [789] attempted to provide an extension of the Chencov uniqueness theorem to arbitrary dimensional $L_1(\mathcal{A})^{+}$ by restricting the class of morphisms under consideration. For a recent work on these problems, see [46].

or ACR γ-distance determine the FRJ metric and the Chencov–Amari γ-connection, with $(\nabla^{\gamma})^{\dagger} = \nabla^{1-\gamma}$. Thus, CAN geometries are Eguchi geometries induced from Liese–Vajda γ-distances. The third order Taylor expansion of D_{γ} can be expressed completely in geometric terms. Let $p(\theta+\mathbf{d}\theta) = p+\mathbf{d}p$ be expressed as $p+\sum_{i=1}^{n}\frac{\partial\log p(\chi,\theta)}{\partial\theta^i}\mathbf{d}\theta^i$ in terms of 0-representation $\mathbf{T}_{\ell_0(p)}\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$ of a vector space $\mathbf{T}_p\mathcal{M}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$, where $\partial_i\log p(\theta) = \partial_i\ell_0(\theta)$ represents ∂_i. Then, up to cubic terms,

$$D_{\gamma}(p(\theta+\mathbf{d}\theta),p(\theta)) = \frac{1}{2}\sum_{i,j}\mathbf{g}^{\mathrm{FRJ}}_{ij}(\theta)\mathbf{d}\theta^i\otimes\mathbf{d}\theta^j + \sum_{i,j,k}\frac{1}{6}(\Gamma^0_{ijk}+\Gamma^{\gamma}_{kij}+\Gamma^1_{jki})\mathbf{d}\theta^i\otimes\mathbf{d}\theta^j\otimes\mathbf{d}\theta^k. \quad (510)$$

For $p_1(\chi),p_2(\chi)\in L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+_{01}$ the ∇^0-geodesic curve is given by

$$p(\chi,t) = \exp((1-t)p_1(\chi)+tp_2(\chi)-c(t)), \quad (511)$$

with factor $c(t)$ arising from normalisation condition $\int_{\mathcal{X}}\tilde{\mu}(\chi)p(\chi,t) = 1$, while the ∇^1-geodesic is given by

$$p(\chi,t) = (1-t)p_1(\chi)+tp_2(\chi). \quad (512)$$

For this reason the ∇^0-geodesics are called ***exponential***, while the ∇^1-geodesics are called ***affine***. The vector $u(\chi)\in\mathbf{T}_{\ell_0(p)}L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+_{01}$ is not equal to itself after ∇^0-parallel transport to $\tilde{u}(\chi)\in\mathbf{T}_{\ell_0(q)}L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+_{01}$ for $p\neq q$:

$$\mathbf{t}^{\nabla^0}_{p,q}(u) = \tilde{u} \iff \tilde{u}(\chi) = u(\chi) - \int_{\mathcal{X}}\tilde{\mu}(\chi)q(\chi)u(\chi). \quad (513)$$

However, for

$$v(\chi)\in\mathbf{T}_{\ell_1(q)}L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+_{01} := \{v(\chi)\mid\int_{\mathcal{X}}\tilde{\mu}(\chi)v(\chi) = 0\}, \quad (514)$$

the ∇^1-parallel transport *does not* change the vector:

$$\mathbf{t}^{\nabla^1}_{p,q}(v) = \tilde{v} \iff \tilde{v}(\chi) = v(\chi). \quad (515)$$

As shown in [17, 23], if the statistical manifold $L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+_0$ is finite dimensional, then it is dually flat for every $\gamma\in[0,1]$. Hence, as follows from the results discussed in Section 4.3, for any closed and $\nabla^{1-\gamma}$-convex subset $\mathcal{Q}\subseteq L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+_0$, $q\in\mathcal{Q}$, and $p\in L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+_0$, the solution of the minimisation problem

$$\mathfrak{P}^{D_{\gamma}}_{\mathcal{Q}}(p) := \arg\inf_{q\in\mathcal{Q}}\{D_{\gamma}(q,p)\} \quad (516)$$

is a singleton set. In particular, $L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+_0$ is ∇^{γ}-convex for every $\gamma\in[0,1]$. However, under restriction to the submanifold $L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+_{01}$ the γ-connections are dually flat only for $\gamma\in\{0,1\}$, and $L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+_{01}$ is ∇^{γ}-convex only for $\gamma=1$.

The generalised cosine equation (273) has two important special cases. Let $c_1(t)$ and $c_2(t)$ be two curves $[0,1]\to\mathcal{M}$, which are, respectively, a ∇^{γ}-geodesic and a $\nabla^{1-\gamma}$-geodesic, and satisfy $c_1(0)=q=c_2(0)$, $c_1(1)=p$, $c_2(1)=r$ with $c_1^i(t) := \theta^i(c_1(t))-\theta^i(q)$ and $c_2^i(t) := \theta^i(c_2(t))-\theta^i(q)$. Then for $t\to^+ 0$ one obtains [23]

$$D_{\gamma}(c_1(t),q)+D_{\gamma}(q,c_2(t)) = D_{\gamma}(c_1(t),c_2(t))+\mathbf{g}^{\mathrm{FRJ}}_q(\dot{c}_1(t),\dot{c}_2(t))\cdot t^2+\mathcal{O}(t^3). \quad (517)$$

If c_1 and c_2 intersect at q orthogonally with respect to $\mathbf{g}^{\mathrm{FRJ}}$, then the ***generalised pythagorean equation*** holds [150, 179, 524][49]:

$$D_{\gamma}(p,q)+D_{\gamma}(q,r) = D_{\gamma}(p,r). \quad (518)$$

[49]Chencov [150] proved it for $\gamma=1$ and $L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+_{01}$, Csiszár [179] proved it for $\gamma=0$ and $L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+_{01}$, while Nagaoka and Amari [524] proved it for $\gamma\in\mathbb{R}$ and $L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})^+_0$. All these proofs were provided in finite dimensional parametric setting.

Since ∇-geodesics are always ∇-convex, the equation (516) implies that (518) is a special case of the generalised pythagorean equation (451), and, as a result, also of (382).

In the noncommutative parametric case the garden of smooth Eguchi geometries arising from the Kosaki–Petz $\mathfrak{f}$-distances is essentially richer. A ***Morozova–Chencov function*** is defined as a function $\mathfrak{c}(\lambda_1, \lambda_2) \in [0, \infty[$ acting on the set $\lambda_1 \in [0, \infty[$, $\lambda_2 \leq \lambda_1 + \lambda_2 \leq 1$, and satisfying [506]

i) $\mathfrak{c}(\lambda_1, \lambda_2) = \mathfrak{c}(\lambda_2, \lambda_1)$,

ii) $\exists\lambda \in]0, \infty[\quad \mathfrak{c}(\lambda_1, \lambda_2) = \lambda\lambda_1^{-1}$,

iii) $\forall\lambda_3 \in]0, \infty[\quad \mathfrak{c}(\lambda_3\lambda_1, \lambda_3\lambda_2) = \frac{1}{\lambda_3}\mathfrak{c}(\lambda_1, \lambda_2)$.

A ***Petz function*** is defined as a function $\mathfrak{h} : [0, \infty[\to [0, \infty[$ satisfying [575]

i) $\mathfrak{h}(\lambda) = \lambda\mathfrak{h}(\lambda^{-1}) \;\forall\lambda > 0$,

ii) $\mathfrak{h}$ is operator monotone increasing.

To each Petz function $\mathfrak{h}$ there corresponds a unique Morozova–Chencov function $\mathfrak{c}_{\mathfrak{h}}$ [575]

$$\mathfrak{c}_{\mathfrak{h}}(\lambda_1, \lambda_2) = (\mathfrak{h}(\lambda_1\lambda_2^{-1})\lambda_2)^{-1} = \frac{\lambda_1\mathfrak{h}\left(\frac{\lambda_2}{\lambda_1}\right) + \lambda_2\mathfrak{h}\left(\frac{\lambda_1}{\lambda_2}\right)}{2\lambda_1\lambda_2\mathfrak{h}\left(\frac{\lambda_2}{\lambda_1}\right)\mathfrak{h}\left(\frac{\lambda_1}{\lambda_2}\right)} \quad \forall\lambda_1, \lambda_2 > 0. \tag{519}$$

The converse relationship is $\mathfrak{h}(\lambda) = (\mathfrak{c}_{\mathfrak{h}}(\lambda, 1))^{-1}$. Every Morozova–Chencov function $\mathfrak{c}_{\mathfrak{h}}$ is operator convex. A riemannian metric $\mathbf{g}$ on $\mathbf{T}\mathfrak{B}(\mathcal{H})^+_{\star01}$ for $\dim\mathcal{H} = n < \infty$ is called:

(1) ***symmetric*** iff $\mathbf{g}_\rho((\mathbf{T}_\rho\ell_1)^{-1}(x), (\mathbf{T}_\rho\ell_1)^{-1}(y)) = \mathbf{g}_\rho((\mathbf{T}_\rho\ell_1)^{-1}(y^*), (\mathbf{T}_\rho\ell_1)^{-1}(x^*))$,

(2) ***monotone*** iff $\mathbf{g}_{T_\star(\rho)}(\hat{T}_\star(u), \hat{T}_\star(v)) \leq \mathbf{g}_\rho(u, v) \;\forall T_\star \in \mathrm{Mark}_\star(\mathfrak{B}(\mathcal{H})^+_{\star01})$,

(3) ***normalised*** iff $\mathbf{g}_\rho((\mathbf{T}_\rho\ell_1)^{-1}(\mathbb{I}), (\mathbf{T}_\rho\ell_1)^{-1}(\mathbb{I})) = \mathrm{tr}(\rho^{-1})$,

where

$$\hat{T}_\star(u) := \left(\mathbf{T}_{T_\star(\rho)}\ell_1\right)^{-1}\left(T_\star(\mathbf{T}_{T_\star(\rho)}\ell_1(u))\right). \tag{520}$$

According to Petz's characterisation theorem [575], which followed earlier work by Morozova and Chencov [506] and Petz [573, 579], there exists a bijection between the set of monotone symmetric riemannian metrics on $\mathbf{T}\mathfrak{B}(\mathcal{H})^+_{\star01}$ (or, equivalently, on $\cup_{\mathrm{M}_n(\mathbb{C})^+_{01}}\mathrm{M}_n(\mathbb{C})$, see below) and the set of Petz functions. It is given by

$$\mathbf{g}^{\mathfrak{h}}_\rho(u, v) = \left\langle \mathbf{T}_\rho\ell_1(u), \mathfrak{J}^{\mathfrak{h}}_\rho(\mathbf{T}_\rho\ell_1(v)) \right\rangle_{\mathfrak{G}_2(\mathcal{H})}, \tag{521}$$

with

$$\mathfrak{J}^{\mathfrak{h}}_\rho(x) := (\mathfrak{h}(\Delta_\rho)\mathfrak{R}_\rho)^{-1} = \mathfrak{R}_\rho^{-1}(\mathfrak{h}(\Delta_\rho))^{-1} = (\mathfrak{R}_\rho^{1/2}\mathfrak{h}(\Delta_\rho)\mathfrak{R}_\rho^{1/2})^{-1}. \tag{522}$$

The equation (519) allows to write each (521) as

$$\mathbf{g}^{\mathfrak{h}}_\rho(u, v) = \mathrm{tr}\left((\mathbf{T}_\rho\ell_1(u))^*\mathfrak{c}_{\mathfrak{h}}(\mathfrak{L}_\rho, \mathfrak{R}_\rho)\mathbf{T}_\rho\ell_1(v)\right). \tag{523}$$

The riemannian metrics given by equivalent formulas (521) and (523) will be called the ***Morozova–Chencov–Petz metrics***. If $[x, \rho] = 0$, then $\mathfrak{J}^{\mathfrak{h}}_\rho(x) = \rho^{-1}x$. The additional condition $\mathfrak{h}(1) = 1$ is equivalent to normalisation of $\mathbf{g}^{\mathfrak{h}}_\rho$, and corresponds to the condition $\mathfrak{c}_{\mathfrak{h}}(1, 1) = 1$. Hansen [306] showed that every normalised Morozova–Chencov function (519) admits a canonical representation

$$\mathfrak{c}_{\mathfrak{h}}(\lambda_1, \lambda_2) = \int_0^1 \tilde{\mu}(\lambda)\frac{1 + \lambda}{2}\left(\frac{1}{\lambda_1 + \lambda\lambda_2} + \frac{1}{\lambda\lambda_1 + \lambda_2}\right), \tag{524}$$

where $\tilde{\mu} : [0,1] \to [0,1]$ is a probability measure (see also [305]). The set of all normalised Morozova–Chencov functions (519) is a Bauer simplex[50]. The set of all normalised Petz functions is convex. A riemannian distance determined by (521) according to (393) reads [446]

$$d_{\mathbf{g}^{\mathfrak{h}}}(\rho_0,\rho_1) = \inf_c \left\{ \int_0^1 \mathrm{d}\lambda \sqrt{\left\langle \mathbf{T}_\rho \ell_1(\dot{c}(t)), \mathfrak{J}^{\mathfrak{h}}_{c(t)} \mathbf{T}_\rho \ell_1(\dot{c}(t)) \right\rangle_{\mathfrak{G}_2(\mathcal{H})}} \mid c(0) = \rho_0,\ c(1) = \rho_1 \right\}, \tag{525}$$

and it satisfies [446]

$$d_{\mathbf{g}^{\mathfrak{h}}}(\phi,\omega) \geq d_{\mathbf{g}^{\mathfrak{h}}}(T_\star(\phi), T_\star(\omega)) \quad \forall T \in \mathrm{Mark}_\star(\mathcal{M}(\mathcal{N}))\ \ \forall \phi,\omega \in \mathcal{M}(\mathcal{N}), \tag{526}$$

where $T_\star$ varies over coarse grainings into parametric quantum submanifolds of $\mathfrak{B}(\mathcal{H})^+_{\star 01}$, while $\mathcal{M}(\mathcal{N})$ is one of these manifolds and $\mathcal{N} = \mathfrak{B}(\mathcal{H}) \cong \mathrm{M}_n(\mathbb{C})$.

The domain of $\mathbf{g}^{\mathfrak{h}}_\rho(\cdot,\cdot)$ can be extended to $\mathrm{M}_n(\mathbb{C})$ by extension of domain of $\mathfrak{J}^{\mathfrak{h}}_\rho$ to traceless matrices $\{x \in \mathrm{M}_n(\mathbb{C}) \mid \mathrm{tr}(x) = 0\}$ by the complexification $\mathbf{T}_\rho\mathfrak{B}(\mathcal{H})^+_{\star 01} \ni x_1, x_2 \Rightarrow x = x_1 + \mathrm{i}x_2$, and further extension to $\mathrm{M}_n(\mathbb{C}) \cong \mathbb{C}^n \times \mathbb{C}^n$ provided by linearity and $\mathfrak{J}^{\mathfrak{h}}_\rho(\mathbb{I}) = \rho^{-1}\mathbb{I}$. This was done already in [575] for $\rho \in \mathrm{M}_n(\mathbb{C})^+_{01}$, where $\mathbf{g}^{\mathfrak{h}}_\rho(\cdot,\cdot)$ was characterised among all ρ-dependent inner products $\langle\cdot,\cdot\rangle_\rho : \mathrm{M}_n(\mathbb{C}) \times \mathrm{M}_n(\mathbb{C}) \to \mathbb{C}$ by the conditions

($\tilde{1}$) $(x,y) \mapsto \langle x,y\rangle_\rho$ is linear in y and antilinear in x $\quad\forall \rho \in \mathrm{M}_n(\mathbb{C})^+_{01}$,

($\tilde{2}$) $\rho \mapsto \langle x,x\rangle_\rho$ is continuous on $\mathrm{M}_n(\mathbb{C})^+_0$ $\quad\forall x \in \mathrm{M}_n(\mathbb{C})^+_{01}$,

($\tilde{3}$) $\langle x,x\rangle_\rho \geq 0$,

($\tilde{4}$) $\langle x,x\rangle_\rho = 0 \iff x = 0$,

($\tilde{5}$) $\langle T(x),T(y)\rangle_{T(\rho)} \leq \langle x,y\rangle_\rho \quad \forall T \in \mathrm{Mark}_\star(\mathrm{M}_n(\mathbb{C}))$,

($\tilde{6}$) $\langle x,y\rangle_\rho = \langle y^*,x^*\rangle_\rho$.

This allows to drop the notation $\mathbf{T}_\rho\ell_1$ above. However, we prefer to keep it, because we are interested in riemannian metrics more than in inner products, and we also want to provide a clear separation between geometric and representational properties of quantum smooth geometries. An extension of the Petz characterisation theorem to $\rho \in \mathrm{M}_n(\mathbb{C})^+_0$ (corresponding to Campbell's extension [137] of Chencov's characterisation theorem in the commutative case) was carried out by Kumagai [419]. The resulting class of metrics reads

$$\mathbf{g}^{\mathfrak{h}}_\rho(x,y) = m(\mathrm{tr}(\rho))\mathrm{tr}(x^*)\mathrm{tr}(y) + \lambda \left\langle x, \mathfrak{J}^{\mathfrak{h}}_\rho(y) \right\rangle_{\mathfrak{G}_2(\mathcal{H})}, \tag{527}$$

where $m : \mathbb{R}^+ \to \mathbb{R}^+$ and $\lambda > 0$.

According to the Lesniewski–Ruskai theorem [446], the Eguchi equation (416) applied to the Kosaki–Petz $\mathfrak{f}$-distance (186) yields the MCP metric (521) with the Petz function given by

$$\mathfrak{h}_{\mathfrak{f}}(\lambda) = \frac{(\lambda-1)^2}{\mathfrak{f}(\lambda) + \mathfrak{f}^{\mathbf{c}}(\lambda)}. \tag{528}$$

This theorem does not require to assume $\mathfrak{f}(0) \leq 0$. Moreover, due to finite dimensionality of the problem, the additional assumption of the smoothness of $\mathfrak{f}$ is also not required. The set of all functions $\mathfrak{f} :]0,\infty[\to \mathbb{R}$ that are operator convex on $]0,\infty[$ and satisfy $\mathfrak{f}(1) = 0$ will be denoted $\mathfrak{F}$. An $\mathfrak{f} \in \mathfrak{F}$ will be called ***symmetric*** iff $\mathfrak{f} = \mathfrak{f}^{\mathbf{c}}$. A subset of all symmetric elements of $\mathfrak{F}$ will be denoted $\mathfrak{F}^{\mathrm{sym}}$. The convex combinations

$$\bar{\mathfrak{f}}_t(\lambda) := t\mathfrak{f}(\lambda) + (1-t)\mathfrak{f}^{\mathbf{c}}(\lambda) = t\mathfrak{f}(\lambda) + (1-t)\lambda\mathfrak{f}(\lambda^{-1}) \in \mathfrak{F} \tag{529}$$

[50] A ***Bauer simplex*** is a nonempty convex compact subset of a locally convex space that is a Choquet simplex and such that the set of its extreme elements is closed [57, 10].

yield the same function $\mathfrak{h}_{\tilde{\mathfrak{f}}_t}$ for all $t \in [0,1]$. Thus, several different Kosaki–Petz $\mathfrak{f}$-distances lead to the same MCP metric. On the other hand, each $\mathfrak{h}$ defines a unique symmetric $\mathfrak{f}$, given by

$$\mathfrak{f}(\lambda) = \frac{(\lambda-1)^2}{2\mathfrak{h}(\lambda)}. \tag{530}$$

The normalisation condition $\mathfrak{h}(1)=1$ is equivalent to the condition $\mathfrak{f}''(1)=1$. A set of elements $\mathfrak{f}\in\mathfrak{F}$ satisfying this condition will be denoted $\mathfrak{F}_1$, and we will also use the set $\mathfrak{F}_1^{\mathrm{sym}} := \mathfrak{F}_1 \cap \mathfrak{F}^{\mathrm{sym}}$, which consists of all operator convex functions $\mathfrak{f}:]0,\infty[\to \mathbb{R}$ satisfying $\mathfrak{f}=\mathfrak{f}^{\mathfrak{c}}$, $\mathfrak{f}(1)=0$, and $\mathfrak{f}''(1)=1$. As shown in [413],

$$\mathfrak{R}_\rho^{-1} + \mathfrak{L}_\omega^{-1} \geq \mathfrak{R}_\rho^{-1}\mathfrak{h}_{\mathfrak{f}}^{-1}(\Delta_{\omega,\rho}) \geq (\mathfrak{R}_\rho + \mathfrak{L}_\omega)^{-1} \;\; \forall \rho,\omega \in \mathfrak{B}(\mathcal{H})^+_{\star 01} \;\; \forall \mathfrak{f}\in\mathfrak{F}_1. \tag{531}$$

The symmetry condition on $\mathbf{g}^{\mathfrak{h}}_\rho$ is equivalent to

$$\mathfrak{h}(\mathfrak{L}_\rho\mathfrak{R}_\rho^{-1})\mathfrak{R}_\rho = \mathfrak{h}(\mathfrak{R}_\rho\mathfrak{L}_\rho^{-1})\mathfrak{L}_\rho, \tag{532}$$

and it follows that

$$\mathfrak{h}(\mathfrak{L}_\rho\mathfrak{R}_\rho^{-1}) = \mathfrak{h}(\mathfrak{R}_\rho\mathfrak{L}_\rho^{-1})\mathfrak{L}_\rho\mathfrak{R}_\rho^{-1} = \mathfrak{h}(\Delta_\rho^{-1})\Delta_\rho. \tag{533}$$

Thus, the theorems of Petz and Lesniewski–Ruskai imply that there is a bijection between the sets of:

(i) normalised MCP metrics,

(ii) normalised Petz functions,

(iii) elements of $\mathfrak{F}_1^{\mathrm{sym}}$.

The key examples of the MCP metrics are:

1) The ***Mori–Kubo–Bogolyubov metric*** (introduced by Mori [499] and Kubo [414], and shown to be real and positive definite by Bogolyubov [87])

$$\mathbf{g}_\rho^{\mathrm{MKB}}(x,y) = \mathrm{tr}\left(\int_0^\infty \mathrm{d}\lambda\, x^* \frac{1}{\lambda\mathbb{I}+\rho} y \frac{1}{\lambda\mathbb{I}+\rho}\right), \tag{534}$$

which corresponds to

$$\mathfrak{h}(\lambda) = \frac{\lambda-1}{\log\lambda} = \int_0^1 \mathrm{d}t\lambda^t, \quad \mathfrak{c}_{\mathfrak{h}}(\lambda_1,\lambda_2) = \frac{\log\lambda_1 - \log\lambda_2}{\lambda_1-\lambda_2}, \tag{535}$$

and

$$\mathfrak{J}^{\mathfrak{h}}_\rho(x) = \int_0^\infty \mathrm{d}\lambda \frac{1}{\lambda\mathbb{I}+\mathfrak{L}_\rho} x \frac{1}{\lambda\mathbb{I}+\mathfrak{R}_\rho}. \tag{536}$$

In [338] the MKB metric (534) was derived using the Eguchi equation (416) from the Umegaki distance (308), which is a Kosaki–Petz $\mathfrak{f}$-distance for $\mathfrak{f}(\lambda) = -\log\lambda$. Early studies of $\mathbf{g}^{\mathrm{MKB}}$ were carried in [252, 581]. The monotonicity of (534) was proved in [573]. The formula (534) is specified in terms of $x, y \in \mathrm{M}_n(\mathbb{C})$, which corresponds to an implicit use of a 'mixture' ($\gamma = 1$)-representation of $\mathbf{g}^{\mathrm{MKB}}_\rho(u,v)$ by means of $x = \mathbf{T}_\rho\ell_1(u)$ and $y = \mathbf{T}_\rho\ell_1(v)$. A coordinate transformation to an 'exponential' ($\gamma=0$)-representation,

$$\mathbf{T}_\rho\ell_1(u) = u \mapsto \mathbf{T}_\rho\ell_0(u) = u - \mathrm{tr}(\rho u)\mathbb{I} = \int_0^\infty \mathrm{d}\lambda \frac{1}{\rho+\lambda\mathbb{I}} u \frac{1}{\rho+\lambda\mathbb{I}} = \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0}\log(\rho+tu), \tag{537}$$

is a functional form of the Fréchet derivative of $\log\rho$ in the u direction, and is called the ***Kubo transform*** (if $[\rho,u]=0$, then the Kubo transform reduces to $u\mapsto\rho^{-1}u$). This allows to write the MKB metric as

$$\mathbf{g}_\rho^{\mathrm{MKB}}(u,v) = \int_0^1 \mathrm{d}\lambda \mathrm{tr}(\rho^\lambda \mathbf{T}_\rho\ell_0(u)\rho^{1-\lambda}\mathbf{T}_\rho\ell_0(v)) = \mathrm{tr}(\mathbf{T}_\rho\ell_1(u)\mathbf{T}_\rho\ell_0(v)). \tag{538}$$

This form is suitable to describe the MKB metric on arbitrary submanifold $\mathcal{Q}_n \subseteq \mathrm{M}_n(\mathbb{C})^{\mathrm{sa}}$:

$$\mathbf{g}_h^{\mathrm{MKB}}(u,v) = \int_0^1 \mathrm{d}\lambda \mathrm{tr}\left(\mathrm{e}^{\lambda h} u \mathrm{e}^{(1-\lambda)h} v\right), \tag{539}$$

where $h \in \mathcal{Q}_n$ and $u, v \in \mathbf{T}_h\mathcal{Q}_n \subset \mathbf{T}_h\mathrm{M}_n(\mathbb{C})^{\mathrm{sa}} \cong \mathrm{M}_n(\mathbb{C})^{\mathrm{sa}}$. Using the parametrisation H of $\mathcal{Q}_n$ defined by (478), we can locally express matrix elements of the metric tensor (539) as

$$\mathbf{g}_{ij}^{\mathrm{MKB}}(x) = \int_0^1 \mathrm{d}\lambda \mathrm{tr}\left(\mathrm{e}^{\lambda H(x)}(\partial_i H)(x)\mathrm{e}^{(1-\lambda)H(x)}(\partial_j H)(x)\right). \tag{540}$$

This can be simplified to

$$\mathbf{g}_{ij}^{\mathrm{MKB}}(x) = \partial_i\partial_j \mathrm{tr}\left(\mathrm{e}^{H(x)}\right) - \mathrm{tr}\left((\partial_i\partial_j H)(x)\mathrm{e}^{H(x)}\right), \tag{541}$$

which follows from

$$\begin{aligned} \partial_i \mathrm{e}^H &= \int_0^1 \mathrm{d}\lambda \mathrm{e}^{\lambda H}(\partial_i H)\mathrm{e}^{(1-\lambda)H}, & (542)\\ \partial_i\partial_j \mathrm{tr}\left(\mathrm{e}^H\right) &= \partial_i \mathrm{tr}\left(\mathrm{e}^H \partial_j H\right) \\ &= \int_0^1 \mathrm{d}\lambda \mathrm{tr}\left(\mathrm{e}^{\lambda H}(\partial_i H)\mathrm{e}^{(1-\lambda)H}(\partial_j H)\right) + \mathrm{tr}\left(\mathrm{e}^H \partial_i\partial_j H\right) \\ &= \mathbf{g}_{ij}^{\mathrm{MKB}} + \mathrm{tr}\left((\partial_i\partial_j H)\mathrm{e}^H\right). & (543) \end{aligned}$$

2) The ***Wigner–Yanase metric*** [770]

$$\mathbf{g}_\rho^{\mathrm{WY}}(x,y) = 4\mathrm{tr}\left(\left((\sqrt{\mathfrak{L}_\rho} + \sqrt{\mathfrak{R}_\rho})^{-1}(x)\right)^* \left((\sqrt{\mathfrak{L}_\rho} + \sqrt{\mathfrak{R}_\rho})^{-1}(y)\right)\right), \tag{544}$$

which corresponds to

$$\mathfrak{h}(\lambda) = \frac{1}{4}(1+\sqrt{\lambda})^2, \quad \mathfrak{c}_{\mathfrak{h}}(\lambda_1,\lambda_2) = \left(\frac{2}{\sqrt{\lambda_1}+\sqrt{\lambda_2}}\right), \tag{545}$$

and

$$\mathfrak{J}_\rho^{\mathfrak{h}}(x) = 4(\sqrt{\mathfrak{L}_\rho} + \sqrt{\mathfrak{R}_\rho})^{-2}(x), \tag{546}$$

and can be derived using the Eguchi equation (416) from the distance

$$D(\rho_1,\rho_2) = 4\mathrm{tr}\left(\rho_1 - \rho_1^{1/2}\rho_2^{1/2}\right) \tag{547}$$

which is a Kosaki–Petz $\mathfrak{f}$-distance for $\mathfrak{f}(\lambda) = 4(1-\sqrt{\lambda})$. The riemannian distance of (544) reads [278]

$$d_{\mathbf{g}^{\mathrm{WY}}}(\rho_1,\rho_2) = 2\arccos\left(\mathrm{tr}(\sqrt{\rho_1}\sqrt{\rho_2})\right). \tag{548}$$

3) The ***Yuen–Lax metric*** [780, 781], known also as the *right/left logarithmic derivative metric*,

$$\mathbf{g}_\rho^{\mathrm{YL}}(x,y) = \frac{1}{2}\mathrm{tr}\left(\rho^{-1}(x^*y + yx^*)\right), \tag{549}$$

which corresponds to

$$\mathfrak{h}(\lambda) = \frac{2}{1+\lambda}, \quad \mathfrak{c}_{\mathfrak{h}}(\lambda_1,\lambda_2) = \frac{\lambda_1+\lambda_2}{2\lambda_1\lambda_2}, \tag{550}$$

and

$$\mathfrak{J}_\rho^{\mathfrak{h}}(x) = \frac{1}{2}(\rho^{-1}x + x\rho^{-1}). \tag{551}$$

The Yuen–Lax metric can be derived from the distance (187), which is a Kosaki–Petz $\mathfrak{f}$-distance for $\mathfrak{f}(\lambda) = (\lambda-1)^2$.

4) The ***Wigner–Yanase–Dyson–Hasegawa metric*** [770, 313] for $\gamma \in [-1,2] \setminus \{0,1\}$

$$\mathbf{g}^{\gamma}_{ij}(\theta(\rho)) = \mathrm{tr}\left(\rho \frac{\partial \log\rho}{\partial\theta^i}\frac{\partial\log\rho}{\partial\theta^j}\right) + \frac{1}{\gamma(1-\gamma)}\mathrm{tr}\left([\rho^{\gamma}, W_i][\rho^{1-\gamma}, W_j]\right), \tag{552}$$

where

$$\mathbf{g}^{\gamma}_{ij}(\theta(\rho)) = \mathrm{tr}\left(\mathbf{T}_{\rho}\ell_{\gamma}\left(\frac{\partial}{\partial\theta^i}\right)\mathbf{T}_{\rho}\ell_{1-\gamma}\left(\frac{\partial}{\partial\theta^j}\right)\right) = \mathbf{g}^{1-\gamma}_{ij}(\theta(\rho)). \tag{553}$$

Given decomposition (481), we have

$$\mathbf{g}^{\gamma}_{\rho}(\mathrm{i}[\rho,W],\mathrm{i}[\rho,W]) = \frac{1}{\gamma(1-\gamma)}\mathrm{tr}([\rho^{\gamma},\mathrm{i}W][\rho^{1-\gamma},\mathrm{i}W]) =: -\frac{2}{\gamma(1-\gamma)}\mathbf{I}^{\mathrm{WYD}}_{\gamma}(\rho,\mathrm{i}W). \tag{554}$$

The former part of the metric $\mathbf{g}^{\gamma}_{ij}(\theta(\rho))$ is a direct quantum analogue of the FRJ metric (499), while $\mathbf{I}^{\mathrm{WYD}}_{\gamma}(\rho,\mathrm{i}W)$ is called the ***Wigner–Yanase–Dyson skew information***, and it arises due to $[\rho, \eth\rho] \neq 0$. The WYDH metric (552) corresponds to [578, 317]

$$\mathfrak{h}_{\gamma}(\lambda) = \gamma(1-\gamma)\frac{(\lambda-1)^2}{(\lambda^{\gamma}-1)(\lambda^{1-\gamma}-1)}, \quad \mathfrak{c}_{\mathfrak{h}_{\gamma}}(\lambda_1,\lambda_2) = \frac{1}{\gamma(1-\gamma)}\frac{(\lambda_1^{\gamma}-\lambda_2^{\gamma})(\lambda_1^{1-\gamma}-\lambda_2^{1-\gamma})}{(\lambda_1-\lambda_2)^2}. \tag{555}$$

The metric (552) was obtained in [313] by applying the Eguchi equation (416) to the Hasegawa distance (305), which is a Kosaki–Petz $\mathfrak{f}$-distance for

$$\mathfrak{f}(\lambda) = \frac{1}{\gamma(1-\gamma)}(1-\lambda^{\gamma}). \tag{556}$$

The restriction of the allowed domain of γ from $\mathbb{R}\setminus\{0,1\}$ to $[-1,2]\setminus\{0,1\}$ follows from the fact that (556) is operator convex only for the latter range [578, 317, 316]. The Wigner–Yanase metric (544) is the WYDH metric with $\gamma = \frac{1}{2}$, the Yuen–Lax metric (549) is the WYDH metric with $\gamma = 2$, while the MKB metric (534) arises as a limit of the WYDH family for $\gamma \to 0$ or $\gamma \to 1$, which can be obtained for (521) with (555), or for (186) with (556). These limits turn (552) into [313]

$$\mathbf{g}^{\mathrm{MKB}}_{ij}(\theta(\rho)) = \mathrm{tr}\left(\rho\frac{\partial\log\rho}{\partial\theta^i}\frac{\partial\log\rho}{\partial\theta^j}\right) + \mathrm{tr}\left([W_i,\log\rho][W_j,\rho]\right). \tag{557}$$

5) The ***Helstrom–Uhlmann metric*** [322, 323, 324, 730, 731] (see also [521, 332, 729, 219, 220]), known also as the *symmetric logarithmic derivative metric*,

$$\mathbf{g}^{\mathrm{HU}}_{\rho}(x,y) = \mathrm{tr}\left(x^*(\mathfrak{R}_{\rho}+\mathfrak{L}_{\rho})^{-1}(y)\right), \tag{558}$$

which corresponds to

$$\mathfrak{h}(\lambda) = \frac{1+\lambda}{2}, \quad \mathfrak{c}_{\mathfrak{h}}(\lambda_1,\lambda_2) = \frac{2}{\lambda_1+\lambda_2}, \tag{559}$$

and

$$\mathfrak{J}^{\mathfrak{h}}_{\rho}(x) = \frac{1}{\mathfrak{R}_{\rho}+\mathfrak{L}_{\rho}}(x) = 2\int_0^{\infty}\mathrm{d}\lambda \mathrm{e}^{-\lambda\rho}x\mathrm{e}^{-\lambda\rho}. \tag{560}$$

The Helstrom–Uhlmann metric can be derived from the distance

$$D(\rho_1,\rho_2) = \mathrm{tr}\left((\rho_1-\rho_2)(\mathfrak{R}_{\rho_2}+\mathfrak{L}_{\rho_1})^{-1}(\rho_1-\rho_2)\right), \tag{561}$$

which is a Kosaki–Petz $\mathfrak{f}$-distance with $\mathfrak{f}(\lambda) = \frac{(\lambda-1)^2}{\lambda+1}$. Uhlmann [730, 731] showed that a riemannian distance of (558) satisfies

$$d_{\mathbf{g}^{\mathrm{HU}}}(\rho_1,\rho_2) = 4d_{\mathrm{Bures}}(\rho_1,\rho_2) = \sqrt{\mathrm{tr}(\rho_1)+\mathrm{tr}(\rho_2)-\mathrm{tr}\sqrt{\rho_1^{1/2}\rho_2\rho_1^{1/2}}}, \tag{562}$$

where d_{Bures} is a Bures distance (161). The monotonicity of (558) was first established in [117]. An explicit formula for (558) is given in [222].

From (531) it follows that

$$\mathfrak{L}_\rho^{-1}+\mathfrak{R}_\rho^{-1}\geq\mathfrak{J}_\rho^{\mathfrak{h}_\mathfrak{f}}\geq(\mathfrak{R}_\rho+\mathfrak{L}_\rho)^{-1}\quad\forall\mathfrak{f}\in\mathfrak{F}_1^{\mathrm{sym}},\tag{563}$$

hence

$$\mathbf{g}_\rho^{\mathrm{YL}}(x,y)\geq\mathbf{g}_\rho^{\mathfrak{h}_\mathfrak{f}}(x,y)\geq\mathbf{g}^{\mathrm{HU}}(x,y)\quad\forall\mathfrak{f}\in\mathfrak{F}_1^{\mathrm{sym}}\;\;\forall x,y\in\mathrm{M}_n(\mathbb{C}).\tag{564}$$

An extension of the MCP metrics to the pure states ($\rho^2=\rho$) was studied in [221, 707, 580]. Among the above examples, only $\mathbf{g}^{\mathrm{HU}}$ admits such extension, and it coincides on the pure states with the Fubini–Study metric [267, 705].

The equation (554) was generalised by Hansen [306] to the notion of ***skew information*** of $x\in\mathrm{M}_n(\mathbb{C})$ with respect to ρ,

$$\mathbf{I}^{\mathfrak{h}}(\rho,x):=\frac{\mathfrak{h}(0)}{2}\mathbf{g}_\rho^{\mathfrak{h}}(\mathrm{i}[\rho,x],\mathrm{i}[\rho,x])=\frac{\mathfrak{h}(0)}{2}\mathrm{tr}\left(\mathrm{i}[\rho,x^*]\mathfrak{c}_{\mathfrak{h}}(\mathfrak{L}_\rho,\mathfrak{R}_\rho)\mathrm{i}[\rho,x]\right),\tag{565}$$

for such $\mathfrak{h}$ that satisfy $\mathfrak{h}(0)>0$. This is not the case for (554) if $\gamma\in]1,2]$, for which another extension of $\mathbf{I}_\gamma^{\mathrm{WYD}}$ has been constructed, see [313, 365, 135].

Early study [522, 523, 314, 315, 23] of the affine connections on $\mathrm{M}_n(\mathbb{C})_0^+$ which are the Norden–Sen dual with respect to a given MCP metric showed that, unlike in commutative case, the Norden–Sen dual connections can satisfy $\mathrm{R}^l_{ijk}=0$ and $\mathfrak{T}^k_{ij}\neq0$ (which, as remarked in [23], allows to study 'distant parallelism' on quantum manifolds). The quantum analogues of ($\gamma=0$)- and ($\gamma=1$)-connections can be defined by means of parallel transports

$$\mathbf{t}^{\nabla^0}_{\rho,\omega}:\mathbf{T}_{\ell_0(\rho)}\mathfrak{B}(\mathcal{H})^+_{\star01}\to\mathbf{T}_{\ell_0(\omega)}\mathfrak{B}(\mathcal{H})^+_{\star01},\tag{566}$$
$$\mathbf{t}^{\nabla^1}_{\rho,\omega}:\mathbf{T}_{\ell_1(\rho)}\mathfrak{B}(\mathcal{H})^+_{\star01}\to\mathbf{T}_{\ell_1(\omega)}\mathfrak{B}(\mathcal{H})^+_{\star01},\tag{567}$$

satisfying the conditions

$$\mathbf{t}^{\nabla^0}_{\rho,\omega}(x)=\tilde{x}\iff\tilde{x}=x-\mathrm{tr}(\omega x),\tag{568}$$
$$\mathbf{t}^{\nabla^1}_{\rho,\omega}(x)=\tilde{x}\iff\tilde{x}=x.\tag{569}$$

Both $\mathbf{t}^{\nabla^0}$ and $\mathbf{t}^{\nabla^1}$ are independent of the choice of the curve connecting ρ and ω, so ∇^0 and ∇^1 have vanishing Riemann curvature tensors. The ($\gamma=0$)-representation of the $\nabla^{\gamma=0}$-covariant derivative applied to $\frac{\partial}{\partial\theta^i}$ is

$$\mathbf{T}_\rho\ell_0\left(\nabla^{\gamma=0}_{\partial_i}\frac{\partial}{\partial\theta^j}\right)=\frac{\partial^2\log\rho}{\partial\theta^i\partial\theta^j}-\mathrm{tr}\left(\rho\frac{\partial^2\log\rho}{\partial\theta^i\partial\theta^j}\right).\tag{570}$$

Let $x_1,\ldots,x_n\in\mathrm{M}_n(\mathbb{C})^{\mathrm{sa}}$. If ρ is parametrised in terms of the ($\gamma=0$)-affine system,

$$\rho=\mathrm{e}^{\sum_{i=1}^n\theta^ix_i-(\log Z)\mathbb{I}},\tag{571}$$

with $\log Z$ determined by the condition $\mathrm{tr}(\rho)=1$, then (570) becomes equal to zero, which shows that $\mathfrak{B}(\mathcal{H})^+_{\star01}$, as well as the parametric exponential families, are $\nabla^{\gamma=0}$-flat. On the other hand, the ($\gamma=1$)-representation of the $\nabla^{\gamma=1}$-covariant derivative applied to $\frac{\partial}{\partial\theta^i}$ reads

$$\mathbf{T}_\rho\ell_1\left(\nabla^{\gamma=1}_{\partial_i}\frac{\partial}{\partial\theta^j}\right)=\frac{\partial^2\rho}{\partial\theta^i\partial\theta^j}.\tag{572}$$

If ρ is parametrised in terms of the ($\gamma=1$)-affine system,

$$\rho=\sum_{i=1}^n\eta^ix_i,\tag{573}$$

then the right hand side of (572) becomes equal to zero, which shows that $\mathfrak{B}(\mathcal{H})^+_{\star01}$, as well as any of its linear subspaces, is $\nabla^{\gamma=1}$-flat.

The family of torsion-free γ-connections on $\mathrm{M}_n(\mathbb{C})_0^+$ for $\gamma \in \mathbb{R}$ was introduced and studied by Jenčová [355, 356], who defined them by means of γ-representation. She showed that the Riemann curvature tensor $\mathrm{R}^{\nabla^\gamma}$ is equal to zero for all $\gamma \in [-1, 2]$ on $\mathrm{M}_n(\mathbb{C})_0^+$ and for $\gamma \in \{0, 1\}$ on $\mathrm{M}_n(\mathbb{C})_{01}^+$. The affine connections $(\nabla^\gamma)^{\dagger_{\mathfrak{h}}}$, defined as the Norden–Sen duals of ∇^γ with respect to a given MCP metric $\mathbf{g}^{\mathfrak{h}}$ have vanishing Riemann curvature tensor, but are not necessarily torsion-free (hence, are not flat). The geodesics of these connections are determined by the equation

$$x = \ell_\gamma^{-1}\left(\mathfrak{J}^{\mathfrak{h}}_{\rho(t)}(\dot{\rho}(t))\right), \tag{574}$$

where $x \in \mathrm{M}_n(\mathbb{C})^{\mathrm{sa}}$ and $\rho \in \mathrm{M}_n(\mathbb{C})_0^+$. In general, $(\nabla^\gamma)^{\dagger_{\mathfrak{h}}} \neq \nabla^{1-\gamma}$. The condition that $(\nabla^\gamma)^{\dagger_{\mathfrak{h}}}$ is torsion-free implies that $\gamma \in [-1, 2]$ and $\mathfrak{h} = \mathfrak{h}_\gamma$, where $\mathfrak{h}_\gamma$ is given by (555). In such case

$$(\nabla^\gamma)^{\dagger_{\mathfrak{h}_\gamma}} = \nabla^{1-\gamma}. \tag{575}$$

Using an integral representation of $\mathfrak{f} \in \mathfrak{F}$,

$$\mathfrak{f}(\lambda) = c_1(\lambda - 1) + c_2(\lambda - 1)^2 + c_3\frac{(\lambda - 1)^2}{\lambda} + \int_0^\infty \tilde{\mu}(t)(\lambda - 1)^2\frac{1 + t}{\lambda + t}, \tag{576}$$

where $c_2, c_3 \geq 0$, $c_1 \in \mathbb{R}$, and $\tilde{\mu} :]0, \infty[\to \mathbb{R}^+$ is a measure satisfying $\int_0^\infty \tilde{\mu}(t) \in \mathbb{R}$ [446], Jenčová [359] showed that the $\mathfrak{f}$-***connections***, defined by the Eguchi equation (417) applied to the Kosaki–Petz $\mathfrak{f}$-distance (186), have the form

$$\mathbf{g}^{\mathfrak{h}_\mathfrak{f}}_\rho(\nabla^\mathfrak{f}_x y, z) = 2\int_0^\infty \tilde{\mu}(\lambda)\mathrm{re}\,(\widetilde{C}(\lambda, z, x, y)) - 2\int_0^\infty \tilde{\mu}(\lambda^{-1})\left(\mathrm{re}\,(\widetilde{C}(\lambda, y, x, z) + \mathrm{re}\,(\widetilde{C}(\lambda, y, x, z))\right), \tag{577}$$

where

$$\widetilde{C}(\lambda, x, y, z) := (1 + \lambda)\mathrm{tr}\left(x\frac{1}{\lambda\mathfrak{R}_\rho + \mathfrak{L}_\rho}(y)\frac{1}{\mathfrak{R}_\rho + \lambda\mathfrak{L}_\rho}(z)\right). \tag{578}$$

The connections determined from $D_\mathfrak{f}$ by the equation (418) are equal to $\nabla^{\mathfrak{f}^{\mathbf{c}}}$. The connections $\nabla^\mathfrak{f}$ and $\nabla^{\mathfrak{f}^{\mathbf{c}}}$ are torsion-free and mutually Norden–Sen dual with respect to $\mathbf{g}^{\mathfrak{h}_\mathfrak{f}} = \mathbf{g}^{\mathfrak{h}_{\mathfrak{f}^{\mathbf{c}}}}$. Using this setting, Jenčová [358, 359] proved that:

1) $\nabla^\mathfrak{f}$ is flat iff $\nabla^\mathfrak{f} = \nabla^\gamma$ for $\gamma \in [-1, 2]$, and in such case $\mathfrak{f} = \mathfrak{f}_\gamma$, where $\mathfrak{f}_\gamma$ is given by (286),
2) for a given $\mathfrak{f} \in \mathfrak{F}_1$ the MCP metric $\mathbf{g}^{\mathfrak{h}_\mathfrak{f}}$ admits a pair of dually flat affine connections iff $\mathfrak{h}_\mathfrak{f} = \mathfrak{h}_\gamma$ for $\gamma \in [-1, 2]$.

In such case, the corresponding dually flat connections are uniquely determined and are given by $(\nabla^\gamma, \nabla^{1-\gamma})$. Hence, the family of quantum Norden–Sen geometries $(\mathrm{M}_n(\mathbb{C})_0^+, \mathbf{g}^{\mathfrak{h}_\gamma}, \nabla^\gamma, \nabla^{1-\gamma})$ for $\gamma \in [-1, 2]$ is characterised as the dually flat Eguchi geometry arising from the Kosaki–Petz $\mathfrak{f}$-distances.[51] Under restriction to $\mathrm{M}_n(\mathbb{C})_{01}^+$ this condition characterises the unique quantum Norden–Sen geometry,

$$(\mathrm{M}_n(\mathbb{C})_{01}^+, \mathbf{g}^{\mathrm{MKB}}, \nabla^1, \nabla^0). \tag{579}$$

For any $\mathfrak{f} \in \mathfrak{F}_1 \setminus \mathfrak{F}_1^{\mathrm{sym}}$ the triple $(\mathrm{M}_n(\mathbb{C})_0^+, \mathbf{g}^{\mathfrak{h}_\mathfrak{f}}, C)$, where

$$C(x, y, z) := \mathbf{g}^{\mathfrak{h}_\mathfrak{f}}\left((\nabla^\mathfrak{f}_x y - \nabla^{\mathfrak{f}^{\mathbf{c}}}_x y), z\right) \tag{580}$$

is a Lauritzen manifold [358, 359]. Every convex mixture

$$\mathfrak{f}_\vartheta := \vartheta\mathfrak{f} + (1 - \vartheta)\mathfrak{f}^{\mathbf{c}} \quad \forall\vartheta \in [0, 1] \tag{581}$$

[51] See [317, 316, 288] for related but weaker characterisations. The first result of this type was obtained by Grasselli and Streater [292, 291, 286], who showed that for $\mathrm{M}_n(\mathbb{C})_{01}^+$ the condition that the pair (∇^1, ∇^0) of affine connections should be Norden–Sen dual with respect to the given riemannian metric $\mathbf{g}$ characterises $\mathbf{g}^{\mathrm{MKB}}$ among all MCP metrics, which selects a single Norden–Sen geometry, $(\mathrm{M}_n(\mathbb{C})_{01}^+, \mathbf{g}^{\mathrm{MKB}}, \nabla^1, \nabla^0)$.

determines a torsion-free connection $\nabla^{\mathfrak{f}_\vartheta}$, which is a ∇^ϑ connection for $(\mathbf{g}^{\mathfrak{h}_\mathfrak{f}}, \nabla^\mathfrak{f}, \nabla^{\mathfrak{f}^\mathfrak{c}})$ in the sense of (419), and $\nabla^{\vartheta=\frac{1}{2}}$ is a Levi-Civita connection of $\mathbf{g}^{\mathfrak{h}_\mathfrak{f}}$. The family of ∇^ϑ-connections associated this way to $\mathfrak{f}_\gamma$ with $\gamma \in \{0,1\}$ is torsion-free, but

$$\nabla^{\gamma=r} \neq \nabla^{\vartheta=r} \quad \forall r \in]0,1[, \tag{582}$$

and ∇^ϑ is not flat unless $\vartheta \in \{0,1\}$ (see also a discussion in [291, 286]). The Ricci curvature scalars of the Levi-Civita connections of the various MCP metrics (and under various assumptions) were calculated in [219, 573, 223, 224, 495, 25, 278]. The Riemann curvature tensors of various $\nabla^\mathfrak{f}$ connections were calculated in [355, 356, 358, 359].

A quantum analogue of the generalised cosine equation (517) for $(\mathrm{M}_n(\mathbb{C})^+_{01}, \mathbf{g}^{\mathrm{MKB}}, \nabla^1, \nabla^0)$ was obtained independently by Petz [573] and Nagaoka [522, 523] and reads

$$D_1|_{\mathcal{N}^+_{\star 1}}(\phi,\omega) + D_1|_{\mathcal{N}^+_{\star 1}}(\omega,\psi) = D_1|_{\mathcal{N}^+_{\star 1}}(\phi,\psi) + \mathbf{g}^{\mathrm{MKB}}_\omega\left(\dot{c}^{\nabla^{\gamma=0}}_{\omega,\phi}(0), \dot{c}^{\nabla^{\gamma=1}}_{\omega,\psi}(0)\right). \tag{583}$$

Under orthogonality assumption, defined as vanishing of the last term at right hand side, (583) turns to a generalised pythagorean equation, which is both a quantum analogue of Chencov's generalised pythagorean equation [150] and a smooth geometric special case of Donald's quantum generalised pythagorean equation (366). The above results of Jenčová allow us to generalise the Nagaoka–Petz generalised pythagorean equation to a direct quantum counterpart of (518):

Proposition 5.1. *If $\phi, \omega, \psi \in \mathrm{M}_n(\mathbb{C})^+_0$, $\gamma \in [-1,2]$, $c^{\nabla^\gamma}_{\omega,\phi}$ is a ∇^γ-geodesic curve, $c^{\nabla^{1-\gamma}}_{\omega,\psi}$ is a $\nabla^{1-\gamma}$-geodesic curve, and these curves intersect at ω while satisfying*

$$\mathbf{g}^{\mathfrak{h}_\gamma}_\omega\left(\dot{c}^{\nabla^\gamma}_{\omega,\phi}(0), \dot{c}^{\nabla^{1-\gamma}}_{\omega,\psi}(0)\right) = 0, \tag{584}$$

then

$$D_\gamma(\phi,\omega) + D_\gamma(\omega,\phi) = D_\gamma(\phi,\psi). \tag{585}$$

Proof. Follows directly from dual flatness of $(\mathrm{M}_n(\mathbb{C})^+_0, \mathbf{g}^{\mathfrak{h}_\gamma}, \nabla^\gamma, \nabla^{1-\gamma})$, (448) and (451). □

An early study of an extension of the Fisher–Rao metric and the Chencov–Amari γ-connections for the nonparametric commutative models was carried out by Amari [18, 22] in terms of bundles of Hilbert spaces $L_2(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ and their subspaces given by (461). Eguchi [244, 245] investigated the equations (416)-(418) for statistical distances over nonparametric models, including γ-connections, using Gâteaux and Fréchet derivatives. However, these approaches lacked a definite underlying smooth manifold structure. Pistone and Rogantin [593] (see also [143]) used the Pistone–Sempi smooth manifold structure, and introduced a positive definite, symmetric, bilinear inner product on $C_0(p,\tilde{\mu})$,

$$C_0(p,\tilde{\mu}) \times C_0(p,\tilde{\mu}) \ni (u,v) \mapsto \int \tilde{\mu} puv \in \mathbb{R}^+. \tag{586}$$

From a generalised Rogers–Hölder inequality for Orlicz spaces it follows that

$$\exists \lambda \in ??? \;\; \forall u,v \in C_0(p,\tilde{\mu}) \quad \left| \int \tilde{\mu} puv \right| \leq \lambda \|u\|_{\Upsilon_1, p\tilde{\mu}} \|v\|_{\Upsilon_1, p\tilde{\mu}}, \tag{587}$$

so (586) is continuous in p. Moreover, it arises as a hessian of the WGKL distance, so it can be considered as a generalisation of the FRJ riemannian metric to nonparametric statistical manifold $\mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$. The same is true for $C_0(p,\tilde{\mu})$ replaced by $B_0(p,\tilde{\mu})$ [286]. A generalisation of the Chencov–Amari family of γ-connections was developed in [280, 281, 142] for the Pistone–Sempi manifolds, and in [286, 287, 289] for the Pistone–Grasselli manifolds. In particular, for $r, q \in w_p^{-1}(U(p))$ or $r, q \in w_p^{-1}(\widetilde{U}(p))$, then $(\gamma = 0)$- and $(\gamma = 1)$-connections are defined in terms of their corresponding parallel transports,

$$\mathbf{t}^{\nabla^0}_{pq} : C_0(p,\tilde{\mu}) \ni u \mapsto u - \int \tilde{\mu} qu \in C_0(q,\tilde{\mu}), \tag{588}$$

$$\mathbf{t}^{\nabla^1}_{pq} : C_0(p,\tilde{\mu}) \ni u \mapsto \frac{p}{q} u \in C_0(q,\tilde{\mu}). \tag{589}$$

The quantities (586), (588), and (589) satisfy the analogue of the Norden–Sen duality (413). However, since tangent spaces $C_0(p, \tilde{\mu})$ and $B_0(p, \tilde{\mu})$ are neither self-dual nor reflexive, these generalisations are somewhat problematic: the Pistone–Gibilisco γ-connections are not defined on tangent spaces but on vector bundles (for $\gamma \in]0, 1[$ these are given by $L_{1/\gamma}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ spaces), while Grasselli's γ-connections do not determine a covariant derivative that would be defined everywhere on $\mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$. For finite dimensional manifolds all these problems disappear and the above settings coincide with the parametric one (see [286, 142] for an explicit discussion).

The infinite dimensional nonparametric generalisation of the Wigner–Yanase metric was introduced by Connes and Størmer [171], while the infinite dimensional nonparametric generalisation of the Wigner–Yanase–Dyson skew information was introduced by Kosaki [397], following earlier work by Pusz and Woronowicz [598]. Construction of the WYDH riemannian metrics on nonparametric quantum manifolds induced by embeddings from noncommutative $L_{1/\gamma}$ spaces was considered in [277], while the construction of various inequivalent families of γ-connections in infinite dimensional noncommutative case was provided in [276, 286, 699, 360, 361].

Let us note, following [203, 593], the distinguished role played by the generalisation of the Chencov–Dawid ∇^0 connection in the nonparametric setting (both commutative and quantum). When smooth manifold structure is introduced through the embedding into the Orlicz space L_{Υ_1}, then it determines each tangent vector by means of an equivalence class of curves, that are one dimensional exponential models: $p \exp(\lambda f - \log Z(p, \lambda f))$ in the commutative case of (469) and $\phi^{\lambda h}$ in the noncommutative case of (493). Hence, the smooth structure equips information model with local ∇^0-geodesics, which determine the ∇^0-connection.

5.3 Exponential models

The most prominent example of an information manifold is an ***exponential family*** defined as an n-dimensional parametric probabilistic manifold [200, 395, 597]

$$\mathcal{M}_{\exp}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) := \{p(x, \theta) := \exp(-\log Z(\theta) - \textstyle\sum_{i=1}^n \theta^i f_i(x)) \mid \theta := (\theta^1, \ldots, \theta^n) \in \Theta \subseteq \mathbb{R}^n\}, \quad (590)$$

where $f_i : \mathcal{X} \to \mathbb{R}$ are assumed to be arbitrary functions, linearly independent of each other and of the constant function 1 (this guarantees that $\theta \mapsto p(\theta)$ is one-to-one and that the matrix $\mathbf{g}_{ij}$ is invertible [758]),

$$\log Z(\theta) := \log \int_{\mathcal{X}} \tilde{\mu}(x) \exp\left(-\sum_{i=1}^n \theta^i f_i(x)\right) \quad (591)$$

is a factor arising from normalisation condition $\int_{\mathcal{X}} \tilde{\mu}(x) p(x, \theta) = 1$, called a ***partition function***, a ***cumulant function***, or a ***Massieu–Planck functional***, while $\Theta \subseteq \mathbb{R}^n$ is supposed to be such open set that the integral in (591) converges. (This definition can be also provided in measure space independent terms of an mcb-algebra $\mathcal{A}$, giving rise to $\mathcal{M}_{\exp}(\mathcal{A})$.) Components of smooth diffeomorphism

$$(\theta^i) : \mathcal{M}_{\exp}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \ni p \mapsto \theta(p) \in \Theta \subseteq \mathbb{R}^n \quad (592)$$

are called ***exponential coordinates***, while the components of smooth diffeomorphism

$$(\eta_i) : \mathcal{M}_{\exp}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \ni p \mapsto \eta(p) := \left(\int_{\mathcal{X}} \tilde{\mu}(x) p(x) f_i(x)\right) \in \Xi \subseteq \mathbb{R}^n \quad (593)$$

are called ***mixture coordinates***. The study of geometric properties of this family provided an original stimulus for development of information geometry [149, 152, 242, 15]. In particular, Chencov found [149] that the finite dimensional exponential families are geodesic surfaces of ∇^0-connections and admit the generalised pythagorean equation (518) for the WGKL distance (290) [150]. Finite dimensional exponential families are analysed in detail in [149, 152, 53, 123]. The infinite dimensional generalisation of (590) is studied in [594, 593, 143, 335, 336].

If $\dim \mathcal{X} =: m < \infty$, then $\int_{\mathcal{X}} \tilde{\mu}(x) k(x) = \sum_{j=1}^{m} k(x_j)$ for any $k : \mathcal{X} \to \mathbb{R}$. In such case $\mathcal{M}_{\exp}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ can be characterised in terms of the Gibbs–Jaynes [275, 344] procedure of maximisation of the ***Gibbs–Shannon entropy*** [275, 672, 673]

$$\mathbf{S}_{\mathrm{GS}}(p) := -\sum_{j=1}^{m} p(x_j) \log p(x_j) \tag{594}$$

subject to constraints $F(p)$ given by

$$\begin{cases} \sum_{j=1}^{m} p(x_j) 1 = 1, \\ \sum_{j=1}^{m} p(x_j) f_i(x_j) = \eta_i, \end{cases} \tag{595}$$

with $\eta := (\eta_i) \in \Xi \subseteq \mathbb{R}^n$. The space of solutions of this variational problem,

$$p(x, \eta) := \underset{p \in L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+}{\arg\max} \left\{ \mathbf{S}_{\mathrm{GS}}(p) + \theta^0 \left(\sum_{j=1}^{m} p(x_j) 1 - 1 \right) + \sum_{i=1}^{n} \theta^i \left(\sum_{j=1}^{m} p(x_j) f_i(x_j) - \eta_i \right) \right\}, \tag{596}$$

attained for all possible (η_i), is given by the special case of the family (590),

$$p(x, \theta) = \mathrm{e}^{-\log Z(\theta) - \sum_{i=1}^{n} \theta^i f_i(x)}. \tag{597}$$

In this case the function (591) takes the form

$$Z(\theta^1, \ldots, \theta^n) = \sum_j \mathrm{e}^{-\sum_{i=1}^{n} \theta^i f_i(x_j)}, \tag{598}$$

with the Lagrange multipliers $(\theta^i) \in \Theta \subseteq \mathbb{R}^n$ determined by

$$\eta_i = -\frac{\partial}{\partial \theta^i} \log Z(\theta^1, \ldots, \theta^n). \tag{599}$$

The maximum value attained by $\mathbf{S}_{\mathrm{GS}}$ for a given (η_i) (or, equivalently, for a given (θ^i)), reads

$$\mathbf{S}_{\mathrm{GS}}(p(\theta)) = \log Z(\theta) + \sum_{i=1}^{n} \theta^i \eta_i. \tag{600}$$

This way the Lagrange multipliers (θ^i) act as 'potentials' of the constraints (595): the greater the value of θ^i, the stronger the impact of the corresponding i-th constraint on the maximum value of the entropy. If η_i in (600) is substituted by (599), then (600) is called a (generalised) ***Gibbs–Helmholtz equation***. See [105, 341, 248, 249] for a detailed treatment of the above procedure including the infinite dimensional case (it is a nontrivial generalisation). As stressed by Jaynes on many occasions (see e.g. [344, 348, 350]), the procedure of maximisation of constrained *absolute* entropy has an interpretative virtue allowing to justify the choice of an exponential model as an information model that is maximally noncommital to any other information than this which is specified by the constraints (595). The conceptual problem associated with this interpretation is to justify why it is $\mathbf{S}_{\mathrm{GS}}$ and not some other lower semi-continuous concave functional on $L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+$ that should be maximised under given constraints (see e.g. [723] for a discussion).

The analogous results can be obtained when the functions f_i are replaced by self-adjoint linear (not necessarily bounded) operators on a Hilbert space $\mathcal{H}$ of arbitrary dimension, p is replaced by $\rho \in \mathfrak{G}_1(\mathcal{H})_1^+$, and $\mathbf{S}_{\mathrm{GS}}$ is replaced by the ***von Neumann entropy***

$$\mathbf{S}_{\mathrm{vN}}(\rho) := -\mathrm{tr}(\rho \log \rho), \tag{601}$$

see [345, 340, 765, 65, 547, 339, 758] for details. While the properties of the noncommutative case with $\dim \mathcal{H} < \infty$ are analogous to the commutative one with $\dim \mathcal{X} < \infty$, the treatments of $\dim \mathcal{H} = \infty$ case require several additional, and not necessary optimal, conditions. A generalisation of (601) to

$L_1(\mathcal{N},\tau)$ spaces for semi-finite W^*-algebras $\mathcal{N}$ was proposed by Segal in [661], and was further studied in [526, 641, 548, 556, 550]. It amounts to replacing the standard trace tr on $\mathfrak{B}(\mathcal{H})$ by a faithful normal semi-finite trace τ on $\mathcal{N}$, and replacement of the Landau–von Neumann density operator $\rho \in \mathfrak{B}(\mathcal{H})^+_\star$ by the Dye–Segal density $\rho \in L_1(\mathcal{N},\tau)$ (see Section 2.2) .[52] However, for all W^*-algebras $\mathcal{N}$ of type different from I_n, the problem of avoiding quantum states with infinite absolute entropy becomes crucial: for any $\phi \in L_1(\mathcal{N})$, its open neighbourhood defined in terms of the metrical distance (160), $d_{L_1(\mathcal{N})}(\phi,\omega) := \frac{1}{2}\|\phi-\omega\|_{L_1(\mathcal{N})}$, contains a dense set of states with infinite Araki distance with respect to ϕ, and a dense set of states with infinite Segal entropy (whenever $\mathcal{N}$ is semi-finite), see [758]. A suitable construction of quantum manifold that rules out those states was provided only recently by Streater [700, 702, 703] (see Section 5.1 for more details, as well as [5, 479] for further developments in this spirit).

An alternative approach, proposed by Haag, Hugenholtz and Winnink [295], is to consider quantum states satisfying the KMS condition as a generalisation of the exponential quantum states with *one* self-adjoint operator. These objects coincide for finite dimensional Hilbert spaces. Araki [34] proved that *for spin systems* the KMS condition is equivalent with the maximum von Neumann entropy condition (see also [626, 427] for earlier proof that in this case the KMS condition is implied by the maximum von Neumann entropy procedure).

Finally, Jenčová and Petz [363, 364] proposed to define a ***quantum exponential family*** as a quantum model

$$\mathcal{M}_{\mathrm{exp}}(\mathcal{N},\omega) := \{\omega^{\sum_{i=1}^n \theta^i x_i} \mid \theta := (\theta^1,\ldots,\theta^n) \in \Theta \subseteq \mathbb{R}^n,\ x_1,\ldots,x_n \in \mathcal{N}^{\mathrm{sa}}\} \subseteq \mathcal{N}^+_{\star 1}, \tag{602}$$

where $\omega \in \mathcal{N}^+_{\star 1}$ is an arbitrary prior (reference) quantum state, while $\omega^{\sum_{i=1}^n \theta^i x_i}$ is an Araki–Donald perturbation (357). This definition follows earlier work by Neirotti and Raggio [532] (see also [642]), and provides a noncommutative counterpart to the approaches of Good [285] and Jaynes [346, 347, 349], who proposed to apply the WGKL distance minimisation with a fixed prior measure as a method of statistical model construction (as opposed to the method of statistical inference) in the case when $(\mathcal{X},\mho(\mathcal{X}))$ is not finite.

Regardless whether the model $\mathcal{M}_{\mathrm{exp}}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$ is introduced by postulate or by the solution of constrained maximisation problem (596), the Gibbs–Shannon entropy $\mathbf{S}_{\mathrm{GS}} : \mathcal{M}_{\mathrm{exp}}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu}) \to \mathbb{R}$ induces a function $\Psi^{\mathbf{L}} : \Xi \to \mathbb{R}$ defined by $\Psi^{\mathbf{L}} := \mathbf{S}_{\mathrm{GS}} \circ (\eta_i)^{-1}$. The function $\Psi^{\mathbf{L}}(\eta)$ is a Legendre dual of a function $\Psi(\theta) := -\log Z(\theta)$ with respect to a vector space duality (259). To simplify the notation, we will write $\mathbf{S}_{\mathrm{GS}}(\eta) \equiv \Psi^{\mathbf{L}}(\eta) := \mathbf{S}_{\mathrm{GS}} \circ (\eta_i)^{-1}(\eta)$. The Legendre transformations (257) and (258) between these two functions are given by

$$(\eta_i) = \mathbf{L}_\Psi(\theta^i) = \left(-\frac{\partial \log Z(\theta)}{\partial \theta^i}\right), \tag{603}$$

$$(\theta^i) = \mathbf{L}_\Psi^{-1}(\eta_i) = \left(\frac{\partial \mathbf{S}_{\mathrm{GS}}(\eta)}{\partial \eta_i}\right). \tag{604}$$

The function Ψ determines the Brègman functional $\bar{D}_\Psi$ on $(\mathbb{R}^n,\mathbb{R}^n)$, given by (267), which induces a Brègman distance $D_\Psi(\cdot,\cdot) := \bar{D}_\Psi(\theta(\cdot),\eta(\cdot))$ on $\mathcal{M}_{\mathrm{exp}}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$ given by (447), which turns out to be equal to the WGKL distance (290).

Given *arbitrary* parametric statistical manifold $\mathcal{M}(\mathcal{A})$ and an *arbitrary* Brègman distance D_Ψ on $\mathcal{M}(\mathcal{A})$, a riemannian metric $\mathbf{g}^\Psi$ associated to D_Ψ by means of (416) has two standard representations in terms of two coordinate-dependent choices of basis in $\mathbf{T}_p\mathcal{M}(\mathcal{A})$,

$$\mathbf{g}^\Psi{}_{ij}(\theta) := \mathbf{g}^\Psi_{\theta(p)}\left(\frac{\partial}{\partial\theta^i},\frac{\partial}{\partial\theta^j}\right) = \frac{\partial^2\Psi(\theta)}{\partial\theta^i\partial\theta^j}, \tag{605}$$

$$\mathbf{g}^{\Psi ij}(\eta) := \mathbf{g}^\Psi_{\eta(p)}\left(\frac{\partial}{\partial\eta_i},\frac{\partial}{\partial\eta_j}\right) = \frac{\partial^2\Psi^{\mathbf{L}}(\eta)}{\partial\eta_i\partial\eta_j}. \tag{606}$$

As shown in [261, 201, 601, 173],

[52] Further generalisation of (601) to weights was carried by Naudts in [528].

i) the ***covariance matrix*** $\mathbf{K}$,

$$\mathbf{K}_{ij}(p) := \int_{\mathcal{X}} \tilde{\mu}(x)p(x)\left(f_i(x) - \int_{\mathcal{X}} \tilde{\mu}(x)p(x)f_i(x)\right)\left(f_j(x) - \int_{\mathcal{X}} \tilde{\mu}(x)p(x)f_j(x)\right), \tag{607}$$

satisfies the ***information inequality***

$$\mathbf{K} \geq (\mathbf{g}^{\mathrm{FRJ}})^{-1} \tag{608}$$

(which means that the matrix $\mathbf{K} - (\mathbf{g}^{\mathrm{FRJ}})^{-1}$ is positive semi-definite) for any $\mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \subseteq L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+_1$,

ii) the equality $\mathbf{K} = (\mathbf{g}^{\mathrm{FRJ}})^{-1}$ *characterises* the model $\mathcal{M}_{\exp}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ among other probabilistic models in $L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+_{01}$.

For exponential family $\mathcal{M}_{\exp}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$, (607) takes the form

$$\mathbf{K}_{ij}(p) = \int_{\mathcal{X}} \tilde{\mu}(x)p(x)(f_i(x) - \eta_i)(f_j(x) - \eta_j), \tag{609}$$

while $\mathbf{g}^{\Psi}_p$ takes the form

$$\mathbf{g}^{\Psi ij}(\theta) = \mathbf{g}^{\Psi ij}(\eta) = \frac{\partial^2 \mathbf{S}_{\mathrm{GS}}(\eta)}{\partial \eta_i \partial \eta_j} = -\frac{\partial \theta^i}{\partial \eta_j} = -\frac{\partial \theta^j}{\partial \eta_i}. \tag{610}$$

Under these conditions, $\mathbf{K} = (\mathbf{g}^{\mathrm{FRJ}})^{-1}$, so one can identify:

$$\mathbf{g}^{\Psi}_{ij}(\theta) = \mathbf{K}_{ij}(\theta) = \frac{\partial^2 \log Z(\theta)}{\partial \theta^i \partial \theta^j} = -\frac{\partial \eta_i}{\partial \theta^j} = -\frac{\partial \eta_j}{\partial \theta^i}. \tag{611}$$

Hence, $\mathbf{K}_{ij}$ and $\mathbf{g}^{\Psi ij}$ are two equivalent coordinate dependent representations of a single FRJ metric $\mathbf{g}^{\mathrm{FRJ}}$.

The ***normal model*** is an example of exponential family given by

$$\mathcal{M}_{\mathrm{norm}}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) := \left\{ p(x, (m,s)) = \frac{1}{\sqrt{2\pi}s} \mathrm{e}^{-\frac{(x-m)^2}{2s^2}} \;\middle|\; (m,s) \in \Theta \subseteq \mathbb{R} \times \mathbb{R}^+ \right\}, \tag{612}$$

where $\dim \mathcal{X} = 1$. It can be obtained by the procedure of maximisation of $\mathbf{S}_{\mathrm{GS}}$ under constraints

$$\begin{cases} \int_{\mathcal{X}} \tilde{\mu}(x)p(x) = 1 \\ \int_{\mathcal{X}} \tilde{\mu}(x)p(x)x = m \\ \int_{\mathcal{X}} \tilde{\mu}(x)p(x)(x-m)^2 = s^2. \end{cases} \tag{613}$$

It can be equipped with the dual pair of mixture ∇^1-affine coordinates

$$(\eta_1, \eta_2) = \left(\int_{\mathcal{X}} \tilde{\mu}(x)p(x,(m,s))x, \int_{\mathcal{X}} \tilde{\mu}(x)p(x,(m,s))x^2\right) = (m, m^2 + s^2) \tag{614}$$

and exponential ∇^0-affine coordinates

$$(\theta^1, \theta^2) = \left(\frac{m}{s^2}, -\frac{1}{2s^2}\right), \tag{615}$$

while the scalar curvature of ∇^γ satisfies [431]

$$\kappa^\gamma = -2\gamma(1-\gamma), \tag{616}$$

so the Ricci scalar curvature of the Levi-Civita connection reads $\kappa^{\gamma=1/2} = -\frac{1}{2}$ [13, 15] (see also [778, 777]). This means that the space of all normal probability densities is the Lobachevskiĭ–Bolyai space [464, 465, 90] of constant negative (scalar and metrical) curvature. It follows that the model $\mathcal{M}_{\mathrm{norm}}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$ can be identified with the symmetric space $\mathrm{SL}(2,\mathbb{R})/\mathrm{SO}(2)$ [715] and that its group of fractional isometric transformations is isomorphic to [235]

$$\mathrm{SU}(1,1)/\mathbb{Z}_2 \cong \mathrm{SL}(2,\mathbb{R})/\mathbb{Z}_2 \cong \mathrm{SO}^{\uparrow}(1,2), \tag{617}$$

that is, to 'ortochronous' subgroup of the Vogt–Lorentz group of transformations of the $(2+1)$-dimensional Minkowski space-time. The 1-parameter subgroups of isometries of the Lie algebra $\mathbf{sl}(2,\mathbb{R})/\mathbb{Z}_2$, generated by the function $\exp(tu)$, with

$$u = \left(\begin{array}{cc} \lambda_1 & \lambda_2 \\ \lambda_3 & -\lambda_1 \end{array}\right) \in \mathbf{sl}(2,\mathbb{R})/\mathbb{Z}_2, \quad \lambda_1,\lambda_2,\lambda_3 \in \mathbb{R}, \tag{618}$$

and $u^2 = (\lambda_1^2+\lambda_2\lambda_3)\mathbb{I}$ are the groups of location transformation (for $\lambda_1=\lambda_3=0$, $\lambda_2=1$), scale transformation (for $\lambda_2=\lambda_3=0$, $\lambda_1=1$), euclidean rotations (for $\lambda_1=0$, $\lambda_2=1$, $\lambda_3=-1$) and hyperbolic rotations (for $\lambda_1=0$, $\lambda_2=\lambda_3=1$) [634]. Moreover, while the transformation $\mathrm{SL}(2,\mathbb{R})$ does not leave the $\mathcal{M}_{\mathrm{norm}}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$ invariant, there exists a unique (up to a multiplicative constant number) metric on this probabilistic manifold which is invariant under the action of $\mathrm{SL}(2,\mathbb{R})$, which is given by the FRJ metric, and a unique scalar product on $\mathbf{T}_{p(\mathcal{X},(0,1))}\mathcal{M}_{\mathrm{norm}}(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})$ that is invariant under the action of SO(2) [515]. The investigation of the information geometric properties of normal model with $\dim\mathcal{X}=2$ was provided in [647]. For an analysis of geometric properties of normal models of with $\dim\mathcal{X}\geq 2$, see [685, 776, 469, 26].

Acknowledgments

I would like to thank Ingemar Bengtsson, Paolo Gibilisco, Carlos Guedes, Anna Jenčová, Wojciech Kamiński, Dimtriĭ Pavlov, David Sherman, Raymond Streater, Stanisław Woronowicz, and Karol Życzkowski for valuable discussions and correspondence. This research was supported in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research and Innovation. Projekt został częściowo sfinansowany ze środków Narodowego Centrum Nauki w ramach projektu N N202 343640.

References

Note: All links provided in references link to the free access files. Files and digital copies that are subject of any sort of restricted access rules were not linked. See michaelnielsen.org/polymath1/index.php?title=Journal_publishing_reform for the reasons why.

[1] Aczél J., 1984, *Measuring information beyond communication theory – why some generalized information measures may be useful, others not*, Aequat. Math. **27**, 1. ↑ 27.

[2] Aczél J., 1987, *Characterizing information measures: approaching the end of an era*, in: Bouchon–Mennier B., Yager R.R. (eds.), *Uncertainty in knowledge-based systems*, LNCS **286**, Springer, Berlin, p.359. ↑ 27.

[3] Aczél J., Daròczy Z., 1975, *On measures of information and their characterizations*, Academic Press, New York. ↑ 27.

[4] Al-Rashed M.H.A., Zegarliński B., 2007, *Noncommutative Orlicz spaces associated to a state*, Studia Math. **180**, 199. ↑ 73.

[5] Al-Rashed M.H.A., Zegarliński B., 2011, *Noncommutative Orlicz spaces associated to a state II*, Lin. Alg. Appl. **435**, 2999. libgen.org:scimag/get.php?doi=10.1016/j.laa.2010.10.027. ↑ 73, 87.

[6] Alber Ya.I., 1996, *Metric and generalized projection operators in Banach spaces: properties and applications*, in: Kartsatos A.G. (ed.), *Theory and applications of nonlinear operators of accretive and monotone type*, Dekker, New York, p.15. ↑ 37.

[7] Alber Ya.I., 1998, *Generalized projections, decompositions, and the pythagorean-type theorem in Banach spaces*, Appl. Math. Lett. **11**, 115. ↑ 37.

[8] Alber Ya.I., 2007, *Young–Fenchel transformation and some new characteristics of Banach spaces*, in: Jarosz K. (ed.), *Function Spaces: Fifth Conference on Function Spaces, May 16-20, 2006, Southern Illinois University, Edwardsville, Illinois*, Contemp. Math. **435**, American Mathematical Society, Providence, p.1. ↑ 37, 38, 40.

[9] Alber Ya.I., Butnariu D., 1997, *Convergence of Bregman projection methods for solving consistent convex feasibility problems in reflexive Banach spaces*, J. Optim. Theor. Appl. **92**, 33. ↑ 36, 37, 38.

[10] Alfsen E.M., 1971, *Compact convex sets and boundary integrals*, Springer, Berlin. libgen.org:EC849C9F200202BC0F57C1830368DF3B. ↑ 78.

[11] Ali S.M., Silvey S.D., 1966, *A general class of coefficients of divergence of one distribution from another*, J. Roy. Stat. Soc. B **28**, 131. ↑ 3, 29.

[12] Altun Y., Smola A., 2006, *Unifying divergence minimization and statistical inference via convex duality*, in: Lugosi G., Simon H.U. (eds.), *Learning theory. 19th Annual Conference on Learning Theory, COLT 2006, Pittsburgh, PA, USA, June 22-25, 2006, Proceedings*, LNCS **4005**, Springer, Berlin, p.139. ↑ 32.

[13] Amari S.-i., 1959, unpublished. ↑ 67, 88.

[14] Amari S.-i., 1980, *Theory of information spaces – a differential–geometrical foundation of statistics*, Post Res. Assoc. Appl. Geom. Mem. Report **106**. ↑ 64, 67, 74, 75.

[15] Amari S.-i., 1982, *Differential geometry of curved exponential families – curvatures and information loss*, Ann. Statist. **10**, 357. euclid:aos/1176345779. ↑ 64, 75, 85, 88.

[16] Amari S.-i., 1983, *A foundation of information geometry*, Electr. Commun. Japan A **66**, 1. ↑ 50.

[17] Amari S.-i., 1985, *Differential–geometrical methods in statistics*, Lecture Notes in Statistics **28**, Springer, Berlin. ↑ 50, 62, 64, 66, 67, 68, 75, 76.

[18] Amari S.-i., 1987, *Dual connections on the Hilbert bundles of statistical models*, in: Dodson C.T.J. (ed.), *Geometrization of statistical theory*, Univ. Lancaster Dept. Math. Publ., Lancaster, p.123. ↑ 75, 84.

[19] Amari S.-i., 2009, *α-divergence is unique, belonging to both f-divergence and Bregman divergence classes*, IEEE Trans. Inf. Theor. **55**, 4925. ↑ 2, 28, 40, 47, 55.

[20] Amari S.-i., 2009, *Information geometry and its applications: convex function and dually flat manifold*, in: Nielsen F. (ed.), *Emerging trends in visual computing*, LNCS **5416**, Springer, Berlin, p.75. ↑ 40, 67.

[21] Amari S.-i., Cichocki A., 2010, *Information geometry of divergence functions*, Bull. Polon. Acad. Sci. Tech. Sci. **58**, 183. ↑ 67.

[22] Amari S.-i., Kumon M., 1988, *Estimation in the presence of infinitely many nuisance parameters: geometry of estimating functions*, Ann. Statist. **16**, 1044. ↑ 84.

[23] Amari S.-i., Nagaoka H., 1993, *Jōhō kika no hōhō*, Iwanami Shoten, Tōkyō (Engl. transl. rev. ed.: 2000, *Methods of information geometry*, American Mathematical Society, Providence). ↑ 40, 48, 64, 66, 67, 75, 76, 82.

[24] Amemiya I., 1953, *A genaralization of Riesz–Fischer's theorem*, J. Math. Soc. Japan **5**, 353. euclid:jmsj/1261415457. ↑ 4.

[25] Andai A., 2003, *Monotone riemannian metrics on density matrices with nonmonotone scalar curvature*, J. Math. Phys. **44**, 3675. arXiv:math-ph/0305060. ↑ 84.

[26] Andai A., 2009, *On the geometry of generalized gaussian distributions*, J. Multivar. Anal. **100**, 777. arXiv:0706.0606. ↑ 89.

[27] Andō T., 1969, *Banachverbände und positive Projektionen*, Math. Z. **109**, 121. gdz.sub.uni-goettingen.de/dms/load/img/?PPN=GDZPPN002404192&IDDOC=168945. ↑ 7.

[28] Andō T., 1978, *Topics on operator inequalities*, Lecture notes, Hokkaidō University, Sapporo. ↑ 31.

[29] Andō T., 1979, *Concavity of certain maps on positive definite matrices and applications to Hadamard products*, Lin. Alg. Appl. **26**, 203. ↑ 54.

[30] Ara P., Goldstein D., 1993, *A solution to the matrix problem for Rickart C^*-algebras*, Math. Nachr. **164**, 259. ↑ 23.

[31] Araki H., 1968, *Multiple time analyticity of a quantum statistical state satisfying the KMS boundary condition*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **4**, 361. dx.doi.org/10.2977/prims/1195194880. ↑ 19.

[32] Araki H., 1972, *Bures distance function and a generalization of Sakai's noncommutative Radon–Nikodym theorem*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **8**, 335. dx.doi.org/10.2977/prims/1195193113. ↑ 27.

[33] Araki H., 1973, *Relative hamiltonian for faithful normal states of von Neumann algebra*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **9**, 165. dx.doi.org/10.2977/prims/1195192744. ↑ 15, 19, 56, 57, 72.

[34] Araki H., 1974, *On the equivalence of the KMS condition and the variational principle for quantum lattice systems*, Commun. Math. Phys. **38**, 1. euclid:cmp/1103859963. ↑ 87.

[35] Araki H., 1974, *Some properties of the modular conjugation operator of von Neumann algebras and a non-commutative Radon–Nikodym theorem with a chain rule*, Pacific J. Math. **50**, 309. euclid:pjm/1102913224. ↑ 27.

[36] Araki H., 1976, *Introduction to relative hamiltonian and relative entropy*, in: Guerra F., Robinson D.W., Stora R. (eds.), *Les methodes mathematiques de la theorie quantique des champs, Marseille, 23-27 juin 1975*, Colloques Internationaux C.N.R.S. **248**, Éditions du C.N.R.S., Paris, p.782. ↑ 50.

[37] Araki H., 1976, *Positive cone, Radon–Nikodym theorems, relative hamiltonian and the Gibbs condition in statistical mechanics*, in: Kastler D. (ed.), *C^*-algebras and their applications to statistical mechanics and quantum field theory*, North-Holland, Amsterdam. ↑ 50.

[38] Araki H., 1976, *Recent developments in the theory of operator algebras and their significance in theoretical physics*, in: *Convegno sulle algebre C^* e loro applicazioni in fisica teoria (12-18 Marzo 1975, Convegno sulla teoria degli operatori indice e teoria K (13-18 Ottobre 1975), del Istituto Nazionale di Alta Matematica, Roma*, Symposia Mathematica **20**, Academic Press, New York, p.395. ↑ 57, 72.
[39] Araki H., 1976, *Relative entropy for states of von Neumann algebras I*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **11**, 809. dx.doi.org/10.2977/prims/1195191148. ↑ 27, 50, 72.
[40] Araki H., 1977, *Relative entropy for states of von Neumann algebras II*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **13**, 173. dx.doi.org/10.2977/prims/1195190105. ↑ 15, 27, 50, 54, 56, 57.
[41] Araki H., 1987, *Recent progress on entropy and relative entropy*, in: Mebkhout M., Sénéor R. (eds.), *Proceedings of VIIIth International Congress on Mathematical Physics*, World Scientific, Singapore, p.354. ↑ 54, 55, 72.
[42] Araki H., Masuda T., 1982, *Positive cones and L_p-spaces for von Neumann algebras*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **18**, 339. dx.doi.org/10.2977/prims/1195183577. ↑ 15, 19, 20, 48.
[43] Arwini K.A., Dodson C.T.J., 2008, *Information geometry: near randomness and near independence*, LNM **1953**, Springer, Berlin. ↑ 67.
[44] Ascoli G., 1932, *Sugli spazi lineari metrici e loro varietà lineari*, Ann. Mat. Pura Appl. **10**, 33, 203. ↑ 34.
[45] Atkinson C., Mitchell A.F.S., 1981, *Rao's distance measure*, Sankhyā **43**, 345. ↑ 74.
[46] Ay N., Jost J., Lê H.V., Schwachhöfer L., 2012, *Information geometry and sufficient statistics*, arXiv:1207.6736. ↑ 67, 70, 75.
[47] Ayupov Sh.A., Chilin V.I., Abdullaev R.Z., 2012, *Orlicz spaces associated with a semi-finite von Neumann algebra*, Comment. Math. Univ. Carolinae **53**, 519. arXiv:1108.3267. ↑ 73.
[48] Baire R., 1905, *Leçons sur les fonctions discontinues, professées au collège de France*, Gauthiers-Villars, Paris. ↑ 29.
[49] Balian R., Alhassid Y., Reinhardt H., 1986, *Dissipation in many-body systems: a geometrical approach based on information theory*, Phys. Rep. **131**, 1. ↑ 67.
[50] Banerjee A., Guo X., Wang H., 2005, *On the optimality of conditional expectation as a Bregman predictor*, IEEE Trans. Inf. Theor. **51**, 2664. ↑ 40.
[51] Banks D., 1993, *A conversation with I.J. Good*, Statist. Sci **11**, 1. ↑ 47.
[52] Barbu V., Precupanu T., 1978, *Convexity and optimisation in Banach spaces*, Sÿthoff & Noordhoff, Alphen aan den Rijn (4th ed., 2012, Springer, Berlin). ↑ 32.
[53] Barndorff-Nielsen O.E., 1978, *Information and exponential families in statistical theory*, Wiley, New York. ↑ 85.
[54] Barndorff-Nielsen O.E., 1988, *Parametric statistical models and likelihood*, Lecture Notes in Statistics **50**, Springer, Berlin. ↑ 67.
[55] Barnett N.S., Carone P., Dragomir S.S., Sofo A., 2002, *Approximating Csiszár f-divergence by the use of Taylor's formula with integral reminder*, Math. Ineq. Appl. **5**, 417. rgmia.org/papers/Csiszar/ACfDTFIRInt2.pdf. ↑ 30, 75.
[56] Bauer H., 1958, *Minimalstellen von Funktionen und Extremalpunkte*, Arch. Math. **9**, 389. ↑ 44.
[57] Bauer H., 1961, *Schilowsche Rand und Dirichletsches Problem*, Ann. Inst. Fourier Grenoble **11**, 89. ↑ 78.
[58] Bauschke H.H., 2003, *Duality for Bregman projections onto translated cones and affine subspaces*, J. Approx. Theor. **121**, 1. ↑ 39.
[59] Bauschke H.H., Borwein J.M., 1997, *Legendre functions and the method of random Bregman projections*, J. Conv. Anal. **4**, 27. www.emis.de/journals/JCA/vol.4_no.1/j86.ps.gz. ↑ 38, 39, 41.
[60] Bauschke H.H., Borwein J.M., 2001, *Joint and separate convexity of the Bregman distance*, Stud. Comp. Math. **8**, 23. people.ok.ubc.ca/bauschke/Research/c04.ps. ↑ 42.
[61] Bauschke H.H., Borwein J.M., Combettes P.L., 2001, *Essential smoothness, essential strict convexity, and Legendre functions in Banach spaces*, Commun. Cont. Math. **3**, 615. ↑ 2, 35, 36, 37, 38, 40.
[62] Bauschke H.H., Borwein J.M., Combettes P.L., 2003, *Bregman monotone optimization algorithms*, Soc. Industr. Appl. Math. J. Contr. Optim. **42**, 596. ↑ 38.
[63] Bauschke H.H., Combettes P.L., 2003, *Constructions of best Bregman approximations in reflexive Banach spaces*, Proc. Amer. Math. Soc. **131**, 3757. ↑ 38.
[64] Bauschke H.H., Lewis A.S., 2000, *Dykstra's algorithm with Bregman projections: a convergence proof*, Optimization **48**, 409. ↑ 38.
[65] Bayer W., Ochs W., 1973, *Quantum states with maximum information entropy I*, Z. Naturf. **28a**, 693. ↑ 86.
[66] Bendat J., Sherman S., 1955, *Monotone and convex operator functions*, Trans. Amer. Math. Soc. **79**, 58. ↑ 31.
[67] Bengtsson I., Życzkowski K., 2006, *Geometry of quantum states: an introduction to quantum entanglement*, Cambridge University Press, Cambridge. ↑ 67.
[68] Beran R., 1977, *Minimum Hellinger distance estimates for parametric models*, Ann. Statist. **5**, 445. ↑ 32.
[69] Bercher J.-F., 1995, *Développement de critères de nature entropique pour la résolution des problèmes inverses linéaires*, Ph.D. thesis, Université de Paris-Sud, Orsay. ↑ 32.
[70] Bessaga C., 1966, *Every infinite-dimensional Hilbert space is diffeomorphic with its unit sphere*, Bull. Acad. Polon. Sci. Sér. Math. **14**, 27. ↑ 69.
[71] Beukers F., Huijsmans C.B., de Pagter B., 1983, *Unital embedding and complexification of f-algebras*, Math. Z. **183**, 131. gdz.sub.uni-goettingen.de/dms/load/img/?PPN=GDZPPN002427753&IDDOC=171006. ↑ 21.
[72] Bhatia R., 1997, *Matrix analysis*, Springer, Berlin. ↑ 31, 54, 72.
[73] Bhattacharyya A., 1943, *On a measure of divergence between two statistical populations defined by their probability distributions*, Bull. Calcutta Math. Soc. **35**, 99. ↑ 27, 51.
[74] Bhattacharyya A., 1943, *On discrimination and divergence*, in: **?**, *Proceedings of 29th Indian Scientific Congress*, part **III**, p.13. ↑ 27, 51.
[75] Bhattacharyya A., 1946, *On a measure of divergence between two multinomial populations*, Sankhyā **7**, 401. ↑ 27, 51.
[76] Bhattacharyya A., 1946, *On some analogues to the amount of information and their uses in statistical estimation*, Sankhyā **8**, 1. ↑ 27, 51.
[77] Birkhoff G., 1938, *Dependent probabilities and spaces* (L), Proc. Nat. Acad. Sci. U.S.A. **24**, 154. www.ncbi.nlm.nih.gov/pmc/articles/PMC1077053/pdf/pnas01791-0050.pdf. ↑ 5.
[78] Birkhoff G., 1940, *Lattice theory*, American Mathematical Society, Providence (3rd rev. ed., 1967). libgen.org:E25AFE588D80C627BD3013ABA587DC06. ↑ 4.
[79] Birkhoff G., Pierce R.S., 1956, *Lattice-ordered rings*, An. Acad. Brasil. Ciênc. **28**, 41. ↑ 4.
[80] Birnbaum Z.W., Orlicz W., 1930, *Über Approximation im Mittel*, Studia Math. **2**, 197. matwbn.icm.edu.pl/ksiazki/sm/sm2/sm2117.pdf. ↑ 33.

[81] Birnbaum Z.W., Orlicz W., 1931, *Über die Verallgemeinerung des Begriffes der zueinander konjugierten Potenzen*, Studia Math. **3**, 1. matwbn.icm.edu.pl/ksiazki/sm/sm3/sm311.pdf. ↑ 69.
[82] Bishop R.L., Crittenden R.J., 1964, *Geometry of manifolds*, Academic Press, New York. ↑ 62.
[83] Blackwell D.A., 1951, *Comparison of experiments*, in: Neyman J. (ed.), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, p.93. euclid:bsmsp/1200500222. ↑ 22, 24.
[84] Blackwell D.A., 1953, *Equivalent comparisons of experiments*, Ann. Math. Statist. **24**, 265. ↑ 24.
[85] Blackwell D.A., Girshick M.A., 1954, *Theory of games and statistical decisions*, Wiley, New York. ↑ 25.
[86] Boţ R.I., 2010, *Conjugate duality in convex optimization*, Springer, Berlin. ↑ 32.
[87] Bogolyubov N.N., 1961, *Kvazisrednie v zadachakh statisticheskoĭ mekhaniki*, D-**871**, Ob'yedinyennyĭ Institut Yadernykh Issledovaniĭ, Laboratoriya Teoreticheskoĭ Fiziki, Dubna (Germ. transl. 1962, *Quasimittelwerte in Problemen der Statistischen Mechanik I, II*, Physik. Abhandl. Sowjetunion **6**, 1, 229). ↑ 79.
[88] Bohnenblust H.F., 1940, *An axiomatic characterisation of L_p-spaces*, Duke Math. J. **6**, 627. ↑ 5, 7.
[89] Bohnenblust H.F., Kakutani S., 1941, *Concrete representation of (M)-spaces*, Ann. Math. **42**, 1025. libgen.org:scimag/get.php?doi=10.2307/1968779. ↑ 7.
[90] Bolyai de Bolya J., 1832, *Appendix: scientiam spatii absolute veram exhibens: a veritate aut falsitate axiomatis XI. Euclidei (a priori haud unquam decidenda) independentem: adjecta ad casum falsitatis, quadratura circuli geometrica*, in: Bolyai de Bolya F., *Tentamen juventutem studiosam in elementa matheseos puræelementaris ac sublimioris, methodo intuitiva evidentiaque huic propria introducendi, cum appendice triplici*, Vol.1, Collegii Reformatorum, Maros Vásárhelyini. ↑ 89.
[91] Bomze I.M, 1990, *A functional analytic approach to statistical experiments*, Longman, Harlow. ↑ 23.
[92] Bomze I.M, 2003, *On the structure of undominated statistical experiments: sufficiency, invariance, and optimality in unbiased estimation*, Tatra Mt. Math. Publ. **26**, 255. ↑ 23.
[93] Boole G., 1847, *The mathematical analysis of logic, being an essay towards a calculus of deductive reasoning*, Macmillan, Barclay, & Macmillan/Bell, Cambridge/London. libgen.org:CA3B34EDCE8DF0D0413205615F286FE9. ↑ 4.
[94] Boole G., 1854, *An investigation of the laws of thought, on which are founded the mathematical theories of logic and probabilities*, Walton and Maberly/Macmillan & Co, London/Cambridge. libgen.org:19D2B28FCD9E538E41E4D5FA415A8D94. ↑ 4.
[95] Borel É., 1905, *Remarques sur certaines questions de probabilité*, Bull. Soc. Math. France **33**, 123. ↑ 22.
[96] Borel É., 1909, *Les probabilités dénomerables et leurs applications arithmétiques*, Rend. Circ. Math. Palermo **27**, 247. ↑ 22.
[97] Borwein J.M., 1993, *On the failure of maximum entropy reconstruction for Fredholm equations and other infinite systems*, Math. Progr. **61**, 251. ↑ 32.
[98] Borwein J.M., Lewis A.S., 1991, *Convergence of best entropy estimates*, Soc. Industr. Appl. Math. J. Optim. **1**, 191. ↑ 32.
[99] Borwein J.M., Lewis A.S., 1991, *Duality relationships for entropy-like minimization problems*, Soc. Industr. Appl. Math. J. Contr. Optim. **29**, 325. ↑ 32.
[100] Borwein J.M., Lewis A.S., 1991, *Moment-matching and best entropy estimation*, J. Math. Anal. Appl. **185**, 596. ↑ 32.
[101] Borwein J.M., Lewis A.S., 1991, *On the convergence of moment problems*, Trans. Amer. Math. Soc. **325**, 249. ↑ 32.
[102] Borwein J.M., Lewis A.S., 1992, *Decomposition of multivariate functions*, Can. J. Math. **44**, 463. ↑ 32.
[103] Borwein J.M., Lewis A.S., 1992, *Partially-finite convex programming, part I: quasi relative interiors and duality theory*, Math. Prog. **57**, 15. ↑ 32.
[104] Borwein J.M., Lewis A.S., 1992, *Partially-finite convex programming, part II: explicit lattice models*, Math. Prog. **57**, 49. ↑ 32.
[105] Borwein J.M., Lewis A.S., 1993, *Partially-finite programming in L_1 and the existence of maximum entropy estimates*, Soc. Industr. Appl. Math. J. Optim. **3**, 248. ↑ 32, 86.
[106] Borwein J.M., Lewis A.S., 1993, *A survey of convergence results for maximum entropy methods*, in: Mohammad-Djafari A., Demoments G. (eds.), *Maximum entropy and bayesian methods*, Kluwer, Dordrecht, p.39. ↑ 32.
[107] Borwein J.M., Lewis A.S., 1994, *Strong rotundity and optimization*, Soc. Industr. Appl. Math. J. Optim. **4**, 146. ↑ 32.
[108] Borwein J.M., Lewis A.S., 2000, *Convex analysis and nonlinear optimization: theory and examples*, Springer, Berlin (2005, 2nd rev. ed.). ↑ 32.
[109] Borwein J.M., Lewis A.S., Noll D., 1996, *Maximum entropy spectral analysis using derivative information. Part I: Fisher information and convex duality*, Math. Oper. Res. **21**, 442. ↑ 32.
[110] Borwein J.M., Lewis A.S., Nussbaum R.D., 1994, *Entropy minimization, DAD problems and doubly stochastic kernels*, J. Funct. Anal. **123**, 264. ↑ 32.
[111] Borwein J.M., Limber M.A., 1996, *On entropy maximization via convex programming*, preprint, oldweb.cecm.sfu.ca/projects/MomEnt+/ent_max.ps.gz. ↑ 32.
[112] Borwein J.M., Vanderwerff J.D., 2001, *Convex functions of Legendre type on Banach spaces*, J. Conv. Anal. **8**, 569. ↑ 35.
[113] Borwein J.M., Vanderwerff J.D., 2010, *Convex functions: constructions, characterizations and counterexamples*, Cambridge University Press, Cambridge. libgen.org:39732ac85975ed9e47778a4b15799882. ↑ 32, 41, 44.
[114] Borwein J.M., Vanderwerff J.D., 2010, *Fréchet–Legendre functions and reflexive Banach spaces*, J. Conv. Anal. **17**, 915. ↑ 35.
[115] Borwein J.M., Zhu Q.J., 2005, *Techniques of variational analysis*, Springer, Berlin. ↑ 32, 33.
[116] Boyd S., Vandenberghe L., 2004, *Convex optimization*, Cambridge University Press, Cambridge. ↑ 32.
[117] Braustein S.L., Caves C.M., 1994, *Statistical distance and the geometry of quantum states*, Phys. Rev. Lett. **72**, 3439. ↑ 81.
[118] Brègman L.M., 1967, *Relaksacionnyĭ metod nakhozhdeniya obshheĭ tochki vypuklykh mnozhestv i ego primenenie dlya resheniya zadach vypuklogo programmirovaniya*, Zh. vychestel. matem. matem. fiz. **7**, 620. (Engl. transl.: 1967, *The relaxation method for finding common points of convex sets and its application to the solution of problems in convex programming*, USSR Comput. Math. Math. Phys. **7**, 200). ↑ 2, 43.
[119] Broniatowski M., Keziou A., 2006, *Minimization of ϕ-divergences on sets of signed measures*, Studia Sci. Math. Hungar. **43**, 403. arXiv:1003.5457. ↑ 32.
[120] Brønsted A., 1964, *Conjugate convex functions in topological vector spaces*, Mat. Fys. Medd. Dansk. Vid. Selsk. **34**, 1. ↑ 33.
[121] Brønsted A., Rockafellar R.T., 1965, *On the subdifferentiability of convex functions*, Proc. Amer. Math. Soc. **16**, 605. www.math.washington.edu/~rtr/papers/rtr-Subdifferentiability.pdf. ↑ 33.
[122] Brown J.R., 1965, *Doubly stochastic measures and Markov operators*, Michigan Math. J. **12**, 367. ↑ 24.
[123] Brown L.D., 1986, *Fundamentals of statistical exponential families with applications in statistical decision theory*, Institute of Mathematical Statistics, Hayward. ↑ 85.
[124] Burbea J., Rao R.C., 1982, *Entropy differential metric, distance and divergence measures in probability spaces: a unified*

approach, J. Multivar. Anal. **12**, 575. ↑ 74.
[125] Burbea J., Rao R.C., 1984, *Differential metrics in probability spaces*, Prob. Math. Statist. **3**, 241. ↑ 74.
[126] Burdet G., Combe P., Nencka H., 2001, *On real hilbertian info-manifolds*, in: Sollich P., Hughston L.P., Streater R.F., Coolen A.C.C. (eds.), *Disordered and complex systems (King's College, London, UK, 10-14 July 2000)*, AIP Conf. Proc. **553**, AIP, Melville, p.153. ↑ 69.
[127] Bures D.J.C., 1969, *An extension of Kakutani's theorem on infinite product measures to the tensor product of semiinfinite W^* algebras*, Trans. Am. Math. Soc. **135**, 194. ↑ 27.
[128] Burg J., 1975, *Maximum entropy spectral analysis*, Ph.D. thesis, Stanford University, Stanford. ↑ 32.
[129] Burkholder D.L., 1961, *Sufficiency in the undominated case*, Ann. Math. Statist. **32**, 1191. ↑ 23.
[130] Butnariu D., Byrne C., Censor Y., 2003, *Redundant axioms in the definition of Bregman functions*, J. Conv. Anal. **10**, 245. math.haifa.ac.il/dbutnaru/publications/bbc.ps. ↑ 35.
[131] Butnariu D., Censor Y., Reich S., 1997, *Iterative averaging of entropic projections for solving stochastic convex feasibility problems*, Comput. Optim. Appl. **8**, 21. math.haifa.ac.il/yair/BCRcoap97.pdf. ↑ 2, 35, 38.
[132] Butnariu D., Iusem A.N., 2000, *Totally convex functions for fixed point computation and infinite dimensional optimization*, Kluwer, Dordrecht. ↑ 32, 38.
[133] Butnariu D., Iusem A.N., Zălinescu C., 2003, *On uniform convexity, total convexity and convergence of the proximal point and outer Bregman projection algorithms in Banach spaces*, J. Conv. Anal. **10**, 35. ↑ 35.
[134] Butnariu D., Resmerita E., 2006, *Bregman distances, totally convex functions, and a method for solving operator equations in Banach spaces*, Abstr. Appl. Anal. **2006**, 84919. www.emis.de/journals/HOA/AAA/Volume2006/84919.pdf. ↑ 38, 41.
[135] Cai L., Hansen F., 2010, *Metric-adjusted skew information: convexity and restricted forms of superadditivity*, Lett. Math. Phys. **93**, 1. ↑ 82.
[136] Campbell L.L., 1985, *The relation between information theory and the differential geometry approach to statistics*, Inf. Sci. (NY) **35**, 199. ↑ 75.
[137] Campbell L.L., 1986, *An extended Čencov characterization of the information metric*, Proc. Amer. Math. Soc. **98**, 135. ↑ 75, 78.
[138] Canderolo D., Volčič A., 2002, *Radon–Nikodým theorems*, in: Pap E. (ed.), *Handbook of measure theory*, Vol.1, North-Holland, Amsterdam, p.249. libgen.org:5DD785AB146D8478B93ACC810B114377. ↑ 10.
[139] Cartan H., 1922, *Sur une généralisation de la notion de courbure de Riemann et les espaces à torsion*, Compt. Rend. Acad. Sci. Paris **174**, 593. ↑ 61.
[140] Cartan H., 1923, *Sur les variétés à connexion affine et la théorie de la relativité généralisée*, Ann. École Norm. Sup. **40**, 325 (Engl. transl. 1986, *On manifolds with an affine connection and the theory of general relativity*, Bibliopolis, Napoli). ↑ 61.
[141] Cartan H., 1967, *Calcul différentiel*, Hermann, Paris (engl. transl. 1971, *Differential calculus*, Houghton Mifflin, Boston). ↑ 58.
[142] Cena A., 2003, *Geometric structures on the non-parametric statistical manifold*, Ph.D. thesis, Università di Milano, Milano. ↑ 70, 84, 85.
[143] Cena A., Pistone G., 2007, *Exponential statistical manifold*, Ann. Inst. Stat. Math. **59**, 27. ↑ 67, 70, 84, 85.
[144] Censor Y., Lent A., 1981, *An iterative row-action method for interval convex programming*, J. Optim. Theory Appl. **34**, 321. ↑ 27, 38, 39, 41.
[145] Censor Y., Lent A., 1987, *Optimization of "$\log x$" entropy over linear equality constraints*, Soc. Industr. Appl. Math. J. Contr. Optim. **35**, 921. ↑ 38, 39.
[146] Chen G., Teboulle M., 1993, *Convergence analysis of a proximal-like minimization algorithm using Bregman functions*, Soc. Industr. Appl. Math. J. Optim. **3**, 538. ↑ 38.
[147] Chencov N.N., 1964, *Geometriya "mnogoobraziya" raspredeleniĭ veroyatnosteĭ*, Dokl. AN SSSR **158**, 543. ↑ 63, 67, 69, 75.
[148] Chencov N.N., 1965, *Kategorii matematicheskoĭ statistiki*, Dokl. AN SSSR **164**, 511. ↑ 24, 25.
[149] Chencov N.N., 1966, *K sistematicheskoĭ teorii eksponencial'nykh semeĭstv raspredeleniĭ veroyatnosteĭ*, Teor. Ver. Premen. **11**, 483. ↑ 85.
[150] Chencov N.N., 1968, *Nesimmetrichnoe rasstoyanie mezhdu raspredeleniyami veroyatnosteĭ, entropiya i teorema Pifagora*, Mat. Zametki **4**, 323. (Engl. transl. 1968, *Nonsymmetrical distance between probability distributions, entropy and the theorem of Pythagoras*, Math. Notes **4**, 686.). ↑ 32, 63, 76, 84, 85.
[151] Chencov N.N., 1969, *Obshhaya teoriya statisticheskogo vyvoda*, Mat. Zametki **5**, 635 (Engl. transl. 1969, *General theory of statistical decisions*, Math. Notes **5**, 380). ↑ 3, 24, 69, 75.
[152] Chencov N.N., 1972, *Statisticheskie reshayushhie pravila i optimal'nye vyvody*, Nauka, Moskva (Engl. transl.: 1982, *Statistical decision rules and optimal inference*, American Mathematical Society, Providence). ↑ 1, 3, 24, 27, 67, 69, 74, 75, 85.
[153] Chencov N.N., 1978, *Algebraic foundation of mathematical statistics*, Math. Operationsforsch. Statist., Ser. Statist. **9**, 267. ↑ 3, 24.
[154] Chencov N.N., 1987, *Pochemu L_1-podkhod i chto za gorizontom*, appendix to: Devroye L., Györfi L., *Neparametricheskoe ocenivanie plotnosti. L_1-podkhod*, Mir, Moskva, p.348 (Engl. transl. 2011, *Why L_1 view and what is next?*, Kybernetika **47**, 840). ↑ 3.
[155] Cheng S.Y., Yau S.T., 1982, *The real Monge-Ampére equation and affine flat structures*, in: *Proceedings of the 1980 Beijing symposium of differential geometry and differential equations*, Gordon & Breach, New York, p.339. ↑ 64.
[156] Chernoff H., 1952, *A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations*, Ann. Math. Statist. **23**, 493. omega.albany.edu:8008/Classics/Chernoff52.pdf. ↑ 50.
[157] Choi M.-D., 1974, *A Schwarz inequality for positive linear maps on C^*-algebras*, Illinois J. Math. **18**, 565. ↑ 31.
[158] Christoffel E.B., 1869, *Über die Transformation der homogenen Differentialausdrücke zweiten Grades*, J. rein. angew. Math. B **70**, 46. ↑ 61.
[159] Chruściński D., Jamiołkowski A., 2004, *Geometric phases in classical and quantum mechanics*, Birkhäuser, Basel. ↑ 67.
[160] Cichocki A., Amari S.-i., 2010, *Families of alpha- beta- and gamma- divergences: flexible and robust measures of similarity*, Entropy **12**, 1532. www.mdpi.com/1099-4300/12/6/1532/pdf. ↑ 27.
[161] Clarke F.H., 1983, *Optimisation and nonsmooth analysis*, Society for Industrial and Applied Mathematics, Philadelphia (1990, 2nd. ed). ↑ 32.
[162] Clarke F.H., Ledyaev Yu.S., Stern R.J., Wolenski P.R., 1998, *Nonsmooth analysis and control theory*, Springer, Berlin. ↑ 32.
[163] Clarkson J.A., 1936, *Uniformly convex spaces*, Trans. Amer. Math. Soc. **40**, 396. ↑ 34.
[164] Connes A., 1973, *Sur le théorème de Radon–Nikodym pour le poids normaux fidèles semifinies*, Bull. Soc. Math. France 2-ème Sér. **97**, 253. ↑ 16.

[165] Connes A., 1973, *Une classification des facteurs de type III*, Ann. Sci. École Norm. Sup. 4 ème sér. **6**, 133. numdam:ASENS_1973_4_6_2_133_0. ↑ 15, 16.
[166] Connes A., 1974, *Caractérisation des espaces vectoriels ordonnés sous-jacents aux algèbres de von Neumann*, Ann. Inst. Fourier Grenoble **24**, 121. numdam:AIF_1974__24_4_121_0. ↑ 15.
[167] Connes A., 1976, *Classification of injective factors. Cases II_1, II_∞, III_λ, $\lambda \neq 1$*, Ann. Math. **104**, 73. libgen.org:scimag/get.php?doi=10.2307/1971057. ↑ 55.
[168] Connes A., 1980, *On a spatial theory of von Neumann algebras*, J. Func. Anal. **35**, 153. libgen.org:scimag/get.php?doi=10.1016/0022-1236(80)90002-6. ↑ 20.
[169] Connes A., 1982, *Classification des facteurs*, in: Kadison R. (ed.), *Operator algebras and applications (Kingston, Ontario, 1980)*, Proceedings of the Symposia in Pure Mathematics **38**, Part 2, American Mathematical Society, Providence, p.43. ↑ 20.
[170] Connes A., 1994, *Noncommutative geometry*, Academic Press, Boston. www.alainconnes.org/docs/book94bigpdf.pdf. ↑ 20.
[171] Connes A., Størmer E., 1978, *Homogeneity of the state space of factors of type III_1*, J. Funct. Anal. **28**, 187. ↑ 85.
[172] Cover T.M., Thomas J.A., 1991, *Elements of information theory*, Wiley, New York (2nd ed., 2006). ↑ 30.
[173] Cramér H., 1946, *A contribution to the theory of statistical estimation*, Skandinavisk Aktuarietidskrift **29**, 85. ↑ 87.
[174] Cressie N., Read T.R.C., 1984, *Multinomial goodness-of-fit tests*, J. Roy. Statist. Soc. B **46**, 440. ↑ 50.
[175] Crooks G.E., Sivak D.A., 2011, *Measures of trajectory ensemble disparity in nonequilibrium statistical dynamics*, J. Stat. Mech. (to appear). threeplusone.com/pubs/technote/CrooksTechNote009-TrajDiv.pdf. ↑ 50.
[176] Csiszár I., 1963, *Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **8**, 85. ↑ 2, 3, 29, 30.
[177] Csiszár I., 1966, *A note on Jensen's inequality*, Stud. Sci. Math. Hungar. **1**, 227. ↑ 55.
[178] Csiszár I., 1967, *Information-type measures of difference of probability distributions and indirect observation*, Stud. Sci. Math. Hungar. **2**, 229. ↑ 30, 32.
[179] Csiszár I., 1975, *I-divergence geometry of probability distributions and minimization problems*, Ann. Prob. **3**, 146. ↑ 29, 32, 76.
[180] Csiszár I., 1978, *Information measures: a critical survey*, in: Kožešnik J. (ed.), *Transactions of the seventh Prague conference on information theory, statistical decision functions, random processes, and of the 1974 european meeting of statisticians held at Prague, from August 18 to 23, 1974*, Vol.B, Springer, Berlin, p.73. ↑ 30.
[181] Csiszár I., 1984, *Sanov property, generalized I-projection and a conditional limit theorem*, Ann. Prob. **12**, 768. ↑ 32.
[182] Csiszár I., 1991, *Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems*, Ann. Stat. **19**, 2032. ↑ 27, 28, 30, 32, 36, 40, 47.
[183] Csiszár I., 1994, *Maximum entropy and related methods*, in: Lachout P., Víšek J.Á. (eds.), *Transactions of twelfth Prague conference on information theory, statistical decision functions and random processes*, Academy of Sciences of the Czech Republic, Institute of Information Theory and Automation, Prague, p.58. ↑ 36.
[184] Csiszár I., 1995, *Generalized projections for nonnegative functions*, Acta Math. Hung. **68**, 161. ↑ 27, 32, 36.
[185] Csiszár I., 1996, *Maxent, mathematics and information theory*, in: Hanson K.M., Silver R.N. (eds.), *Maximum entropy and bayesian methods: proceedings of the fifteenth international workshop on maximum entropy and bayesian methods, Santa Fe, New Mexico, USA, 1995*, Kluwer, Dordrecht, p.35. ↑ 28.
[186] Csiszár I., 2008, *Axiomatic characterizations of information measures*, Entropy **10**, 261. www.mdpi.com/1099-4300/10/3/261/pdf. ↑ 27, 28.
[187] Csiszár I., Fischer J., 1962, *Informationsentfernungen im Raum der Wahrscheinlichkeitsverteilungen*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **7**, 159. ↑ 29.
[188] Csiszár I., Gamboa F., Gassiat E., 1999, *MEM pixel correlated solutions for generalized moment and interpolation problems*, IEEE Trans. Inf. Theor. **45**, 2253. ↑ 30, 32.
[189] Csiszár I., Matúš F., 2003, *Information projections revisited*, IEEE Trans. Inf. Theor. **49**, 1474. ↑ 32.
[190] Csiszár I., Matúš F., 2008, *On minimization of entropy functionals under moment constraints*, in: *Proceedings of the 2008 IEEE International Symposium on Information Theory*, IEEE Information Theory Society, Washington, p.2101. ↑ 32, 36, 39.
[191] Csiszár I., Matúš F., 2009, *On minimization of multivariate entropy functionals*, in: *Proceedings of 2009 IEEE Information Theory Workshop on Networking and Information Theory*, IEEE Information Theory Society, Washington, p.96. ↑ 32, 36, 37.
[192] Csiszár I., Matúš F., 2012, *Generalized minimizers of convex integral functions, Bregman distance, pythagorean identities*, Kybernetika **48**, 637. arXiv:1202.0666. ↑ 32.
[193] Csiszár I., Matúš F., 2012, *Minimization of entropy functionals revisited*, in: *Proceedings of the 2012 IEEE International Symposium on Information Theory*, IEEE Information Theory Society, Washington, p.150. ↑ 32.
[194] Csiszár I., Shields P.C., 2004, *Information theory and statistics: a tutorial*, Found. Trends Commun. Inf. Theor. **1**, 417. ↑ 30.
[195] Cudia D.F., 1964, *The geometry of Banach spaces. Smoothness*, Trans. Amer. Math. Soc. **110**, 284. ↑ 52.
[196] Dacunha-Castelle D., Gamboa F., 1990, *Maximum entropie et problème des moments*, Ann. Inst. Henri Poincaré B **26**, 567. ↑ 32.
[197] Daniell P.J., 1919, *Integrals in an infinite number of dimensions*, Ann. Math. **21**, 30. libgen.org:scimag/get.php?doi=10.2307/1967122. ↑ 4.
[198] Daniell P.J., 1920, *Stieltjes derivatives*, Bull. Amer. Math. Soc. **26**, 444. euclid:bams/1183425376. ↑ 10.
[199] Daniell P.J., 1921, *Integral products and probability*, Am. J. Math. **43**, 143. ↑ 23.
[200] Darmois G., 1935, *Sur les lois de probabilite a estimation exhaustive*, C.R. Acad. Sci. Paris **260**, 1265. ↑ 85.
[201] Darmois G., 1945, *Sur les limites de la dispersion des certaines estimations*, Rev. Inst. Intern. Statist. **13**, 9. ↑ 87.
[202] Davis C., 1963, *Notions generalizing convexity for functions defined on spaces of matrices*, Proc. Amer. Math. Soc. Symp. **7**, 187. ↑ 31.
[203] Dawid A.P., 1975, *Discussion of paper by B. Efron*, Ann. Statist. **3**, 1231. ↑ 68, 69, 75, 85.
[204] Dawid A.P., 1977, *Further comments on a paper by Bradley Efron*, Ann. Statist. **5**, 1249. ↑ 75.
[205] Day M.M., 1944, *Uniform convexity in factor and conjugate spaces*, Ann. Math. **45**, 375. ↑ 35.
[206] de Finetti B., 1970, *Teoria delle probabilità*, Einaudi, Torino. (Engl. transl.: 1974, *Theory of probability: a critical introductory treatment*, Vol.1-2, Wiley, New York. ↑ 23.
[207] De Pierro A.N., Iusem A.N., 1986, *A relaxed version of Bregman's method for convex programming*, J. Optim. Theor. Appl. **51**, 421. ↑ 38.

[208] Decarreau A., Hilhorst D., Lemaréchal C., Navaza J., 1992, *Dual methods in entropy maximization. Application to some problems in crystallography*, Soc. Industr. Appl. Math. J. Optim. **2**, 173. ↑ 32.
[209] Dedekind R., 1872, *Stetigkeit und die Irrationalzahlen*, Vieweg, Braunschweig (Engl. transl: 1901, *Continuity and irrational numbers*, in: Dedekind R., *Essays on the theory of numbers*, Open Court, Chicago, p.1. www.gutenberg.org/files/21016/21016-pdf.pdf). ↑ 4.
[210] Della Pietra S., Della Pietra V., Lafferty J., 2002, *Duality and auxiliary functions for Bregman distances*, Technical report CMU-CS-01-109R, School of computer science, Carnegie Mellon University, Pittsburgh. ↑ 39.
[211] Dhillon I.S., Tropp J.A., 2007, *Matrix nearness with Bregman divergences*, Soc. Industr. Appl. Math. J. Matrix Anal. Appl. **29**, 1120. ↑ 43, 46.
[212] Dieks D., Veltkamp P., 1983, *Distance between quantum states, statistical inference and the projection postulate*, Phys. Lett. A **97**, 24. ↑ 27.
[213] Diepenbrock F.R., 1971, *Charakterisung einer allgemeineren Bedingung als Dominiertheit mit Hilfe von lokalisierten Maßen*, Ph.D. thesis, Universität Münster, Münster. ↑ 23.
[214] Dieudonné J.A., 1940, *Topologies faibles dans les espaces vectoriels*, Compt. Rend. Acad. Sci. Paris **211**, 94. ↑ 32.
[215] Dieudonné J.A., 1942, *La dualité dans les espaces vectoriels topologiques*, Ann. Sci. École Norm. Sup. **59**, 107. numdam:ASENS_1942_3_59__107_0. ↑ 32.
[216] Dieudonné J.A., 1960, *Foundations of modern analysis*, Academic Press, New York. ↑ 58.
[217] Digernes T., 1975, *Duality for weights on covariant systems and its applications*, Ph.D. thesis, University of California, Los Angeles. ↑ 15.
[218] Dillen F., Nomizu K., Vracken L., 1990, *Conjugate connections and Radon's theorem in affine differential geometry*, Monatsh. Math. **109**, 221. ↑ 62, 66.
[219] Dittmann J., 1993, *On the riemannian geometry of finite dimensional state space*, Sem. Sophus Lie **3**, 73. ↑ 81, 84.
[220] Dittmann J., 1994, *Some properties of the riemannian Bures metric on mixed states*, J. Geom. Phys. **13**, 203. ↑ 81.
[221] Dittmann J., 1995, *On the riemannian metric on the space of density matrices*, Rep. Math. Phys. **36**, 309. ↑ 82.
[222] Dittmann J., 1999, *Note on explicit formulae for the Bures metric*, J. Phys. A **32**, 2663. arXiv:quant-ph/9808044. ↑ 81.
[223] Dittmann J., 1999, *The scalar curvature of the Bures metric on the space of density matrices*, J. Geom. Phys. **31**, 16. arXiv:quant-ph/9810012. ↑ 84.
[224] Dittmann J., 2000, *On the curvature of monotone metrics and a conjecture concerning the Kubo-Mori metric*, Lin. Alg. Appl. **315**, 83. arXiv:quant-ph/9906009. ↑ 84.
[225] Dixmier J., 1950, *Les fonctionnelles linéaires sur l'ensemble des opérateurs bornés d'un espace de Hilbert*, Ann. Math. **51**, 387. libgen.org:scimag/get.php?doi=10.2307/1969331. ↑ 18.
[226] Dixmier J., 1953, *Formes linéaires sur un anneau d'opérateurs*, Bull. Soc. Math. France **81**, 9. numdam:BSMF_1953__81__9_0. ↑ 16, 17, 18.
[227] Dixmier J., 1954, *Sur les anneaux d'opérateurs dans les espaces hilbertiens*, Compt. Rend. Acad. Sci. Paris **238**, 439. ↑ 18.
[228] Dodds P.G., Dodds T.K., de Pagter B., 1989, *Non-commutative Banach function spaces*, Math. Z. **201**, 583. gdz.sub.uni-goettingen.de/dms/load/img/?PPN=GDZPPN002436515&IDDOC=173377. ↑ 73.
[229] Donald M.J., 1986, *On the relative entropy*, Commun. Math. Phys. **105**, 13. euclid:cmp/1104115254. ↑ 47, 50, 54.
[230] Donald M.J., 1987, *Free energy and relative entropy*, J. Stat. Phys. **49**, 81. ↑ 55, 72.
[231] Donald M.J., 1987, *Further results on the relative entropy*, Math. Proc. Cambridge Phil. Soc. **101**, 363. ↑ 54, 55, 57.
[232] Donald M.J., 1990, *Relative hamiltonians which are not bounded from above*, J. Funct. Anal. **91**, 143. ↑ 54, 55, 56, 57, 72.
[233] Donald M.J., 1991, *Continuity of relative hamiltonians*, Commun. Math. Phys. **136**, 625. ↑ 55, 72.
[234] Donoghue W.F. Jr., 1974, *Monotone matrix functions and analytic continuation*, Springer, Berlin. ↑ 31.
[235] Dubrovin B.A., Novikov S.P., Fomenko A.T., 1979, *Sovremennaya geometriya*, Nauka, Moskva. (Engl. transl.: 1984, *Modern geometry: methods and applications*, Springer, Berlin). ↑ 89.
[236] Duistermaat J.J., 2001, *On hessian riemannian structures*, Asian J. Math. **5**, 79. ↑ 65.
[237] Dunford N., 1938, *Uniformity in linear spaces*, Trans. Amer. Math. Soc. **44**, 305. libgen.org:scimag/get.php?doi=10.2307/1989974. ↑ 9.
[238] Dye H.A., 1952, *The Radon–Nikodým theorem for finite rings of operators*, Trans. Amer. Math. Soc. **72**, 243. www.ams.org/journals/tran/1952-072-02/S0002-9947-1952-0045954-5/S0002-9947-1952-0045954-5.pdf. ↑ 16, 17.
[239] Ebanks B., Sahoo P., Sander W., 1998, *Characterizations of information measures*, World Scientific, Singapore. ↑ 27.
[240] Eckstein J., 1993, *Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming*, Math. Oper. Res. **18**, 202. ↑ 38.
[241] Efron B., 1975, *Defining the curvature of a statistical problem (with applications to second order efficiency)*, Ann. Statist. **3**, 1189. ↑ 64.
[242] Efron B., 1978, *Controversion in the foundations of statistics*, Am. Math. Monthly **85**, 231. ↑ 85.
[243] Eguchi S., 1983, *Second order efficiency of minimum contrast estimators in a curved exponential family*, Ann. Statist. **11**, 793. euclid:aos/1176346246. ↑ 27, 62, 75.
[244] Eguchi S., 1985, *A differential geometric approach to statistical inference on the basis of contrast functionals*, Hiroshima Math. J. **15**, 341. euclid:hmj/1206130775. ↑ 62, 75, 84.
[245] Eguchi S., 1992, *Geometry of minimum contrast*, Hiroshima Math. J. **22**, 631. ↑ 62, 84.
[246] Ekeland I., Témam R., 1974, *Analyse convexe et problèmes variationnels*, Gauthier–Villars, Paris. (Engl. transl.: 1976, *Convex analysis and variational problems*, North–Holland, Amsterdam). ↑ 32.
[247] Falcone A.J., Takesaki M., 2001, *The non-commutative flow of weights on a von Neumann algebra*, J. Funct. Anal. **182**, 170. www.math.ucla.edu/∼mt/papers/QFlow-Final.tex.pdf. ↑ 2, 18, 19.
[248] Favretti M., 2005, *Isotropic submanifolds generated by the maximum entropy principle and Onsager reciprocity relations*, J. Funct. Anal. **227**, 227. ↑ 86.
[249] Favretti M., 2005, *Lagrangian submanifolds generated by the maximum entropy principle*, Entropy **7**, 1. ↑ 86.
[250] Fell J.M.G., Kelley J.L, 1952, *An algebra of unbounded operators*, Proc. Nat. Acad. Sci. U.S.A. **38**, 592. www.ncbi.nlm.nih.gov/pmc/articles/PMC1063619/pdf/pnas01580-0040.pdf. ↑ 21.
[251] Fenchel W., 1949, *On conjugate convex functions*, Canadian J. Math. **1**, 73. cms.math.ca/cjm/v1/cjm1949v01.0073-0077.pdf. ↑ 33.
[252] Fick E., Sauermann G., 1990, *The quantum statistics of dynamical processes*, Springer, Berlin. ↑ 79.
[253] Fischer A., 2010, *Quantization and clustering with Bregman divergences*, J. Multivar. Anal. **101**, 2207. ↑ 40.
[254] Fisher R.A., 1922, *On the mathematical foundation of theoretical statistics*, Phil. Trans. Roy. Soc. London Ser. A **222**, 309. ↑ 22.

[255] Fisher R.A., 1925, *Theory of statistical estimation*, Proc. Cambridge Phil. Soc. **22**, 700. ↑ 74.
[256] Fréchet M., 1906, *Sur quelques points du calcul fonctionnel*, Rend. Circ. Matem. Palermo **22**, 1. webpages.ursinus.edu/nscoville/Frechet Thesis French I.pdf, webpages.ursinus.edu/nscoville/Frechet Thesis French II.pdf, webpages.ursinus.edu/nscoville/Frechet Thesis French III.pdf, webpages.ursinus.edu/nscoville/Frechet Thesis French IV.pdf. ↑ 26.
[257] Fréchet M., 1911, *Sur la notion de différentielle*, Compt. Rend. Acad. Sci. Paris **152**, 845, 1050. ↑ 34.
[258] Fréchet M., 1925, *Les espaces abstaits topologiquement affines*, Acta Math. *47*, 25. ↑ 34.
[259] Fréchet M., 1925, *Sur la notion de différentielle dans l'analyse générale*, Compt. Rend. Acad. Sci. Paris *180*, 806. ↑ 34.
[260] Fréchet M., 1926, *Sur la notion de différentielle dans l'analyse générale*, Ann. Sci. École Norm. Sup. **42**, 293. ↑ 34.
[261] Frechét M., 1943, *Sur l'extension des certaines évaluations statistiques au cas de petits énchantillous*, Rev. Inst. Intern. Statist. **11**, 183. ↑ 87.
[262] Fremlin D.H., 2000, 2001, 2002, 2003, 2008, *Measure theory*, Vol.1-5, Torres Fremlin, Colchester. New editions available as LaTeX source at: www.essex.ac.uk/maths/people/fremlin/mt.htm, old editions available as PDF files at: libgen.org:79C0EA89BBA9834E2CB4FF9DB9CFAB97, libgen.org:B526B2F8B1F9BA194BAAD8BDFF100582, libgen.org:15094F863F4B67E680FFE00D1A212F05, libgen.org:D0ECA533BBA2A6C314DF429DA1857C37, libgen.org:2E52DFCA19EF962520A36D8A9BE90946, libgen.org:78273D5EC827F6D0937D6F6F40962E3D. ↑ 9, 21.
[263] Freudenthal H., 1936, *Teilweise geordnete Moduln*, Proc. Roy. Acad. Amsterdam **39**, 641. www.dwc.knaw.nl/DL/publications/PU00016905.pdf. ↑ 4, 21.
[264] Frigyik B.A., Srivastava S., Gupta M.R., 2008, *Functional Bregman divergence and bayesian estimation of distributions*, IEEE Trans. Inf. Theor. **54**, 5130. arXiv:cs/0611123. ↑ 36, 40.
[265] Frigyik B.A., Srivastava S., Gupta M.R., 2008, *An introduction to functional derivatives*, Technical report **2008-0001**, Dept. Electr. Eng. Univ. Washington, Seattle. www.ee.washington.edu/research/guptalab/publications/functionalDerivativesIntroduction.pdf. ↑ 36.
[266] Frölicher A., Kriegl A., 1988, *Linear spaces and differentiation theory*, Wiley, Chisterter. ↑ 32.
[267] Fubini G., 1903, *Suelle metriche definite da una forma hermitiana*, Atti Inst. Veneto **6**, 501. ↑ 82.
[268] Gamboa F., 1989, *Méthode du maximum d'entropie sur la moyenne et applications*, Ph.D. thesis, Université de Paris-Sud, Orsay. ↑ 32.
[269] Gamboa F., Gassiat E., 1991, *Maximum d'entropie et problème des moments cas multidimensionnel*, Prob. Math. Statist. **12**, 67. ↑ 32.
[270] Gamboa F., Gassiat E., 1997, *Bayesian methods and maximum entropy for ill-posed inverse problems*, Ann. Statist. **25**, 328. ↑ 32.
[271] Gâteaux M.R., 1913, *Sur les fonctionnelles continues et les fonctionelles analytiques*, Compt. Rend. Acad. Sci. Paris **157**, 325. ↑ 34.
[272] Gâteaux M.R., 1919, *Fonctions d'une infinité de variables indépendantes*, Bull. Soc. Math. France **47**, 70. ↑ 34.
[273] Gâteaux M.R., 1922, *Sur les fonctionnelles continues et les fonctionelles analytiques*, Bull. Soc. Math. France **50**, 1. ↑ 34.
[274] Gel'fand I.M., Naĭmark M.A., 1943, *On the imbedding of normed rings into the ring of operators in Hilbert space*, Matem. Sbornik (N.S.) **12**, 197. mathnet.ru:msb6155. ↑ 14.
[275] Gibbs J.W., 1902, *Elementary principles in statistical mechanics, developed with especial reference to the rational foundation of thermodynamics*, Yale University Press, New Heaven. ↑ 86.
[276] Gibilisco P., Isola T., 1999, *Connections on statistical manifolds of density operators by geometry of non-commutative L^p spaces*, Infin. Dim. Anal. Quant. Prob. Relat. Topics **2**, 169. www.mat.uniroma2.it/∼isola/research/preprints/GiIs01.pdf. ↑ 51, 85.
[277] Gibilisco P., Isola T., 2001, *Monotone metrics on statistical manifolds of density matrices by geometry of non-commutative L^2-spaces*, in: Sollich P., Hughston L.P., Streater R.F., Coolen A.C.C. (eds.), *Disordered and complex systems (King's College, London, UK, 10-14 July 2000)*, AIP Conf. Proc. **553**, AIP, Melville, p.129. www.mat.uniroma2.it/∼isola/research/preprints/GiIs02.pdf. ↑ 85.
[278] Gibilisco P., Isola T., 2003, *Wigner–Yanase information on quantum state space: the geometric approach*, J. Math. Phys. **44**, 3752. /www.mat.uniroma2.it/∼isola/research/preprints/GiIs05.pdf. ↑ 80, 84.
[279] Gibilisco P., Isola T., 2006, *Some open problems in information geometry*, in: Accardi L., Freudenberg W., Schürmann M. (eds.), *Quantum probability and infinite dimensional analysis. Proceedings of the 26th conference (Levico, Italy, 20-26 February 2005)*, World Scientific, Singapore, p.205. www.mat.uniroma2.it/∼isola/research/preprints/GiIs08.pdf. ↑ 3.
[280] Gibilisco P., Pistone G., 1998, *Connections on non-parametric statistical manifolds by Orlicz space geometry*, Inf. Dim. Anal. Quant. Prob. Relat. Top. **1**, 325. ↑ 70, 84.
[281] Gibilisco P., Pistone G., 1999, *Analytical and geometrical properties of statistical connections in information geometry*, in: Beghi A., Finesso L., Picci G. (eds.), *Mathematical Theory of Networks and Systems (Padova, Italy, July 6-10 1998)*, Il Poligrafo, Padova, p.811. ↑ 84.
[282] Giry M., 1982, *A categorical approach to probability theory*, in: Banaschewski B. (ed.), *Categorical Aspects of Topology and Analysis Proceedings of an International Conference Held at Carleton University, Ottawa, August 11-15, 1981*, LNM **915**, Springer, Berlin, p.68. ↑ 25.
[283] Goldstein D., 1995, *Polar decomposition in Rickart C^*-algebras*, Publ. Mat. **39**, 5. dmle.cindoc.csic.es/pdf/PUBLICACIONSMATEMATIQUES_1995_39_01_01.pdf. ↑ 23.
[284] Good I.J., 1950, *Probability and the weighing of evidence*, Griffin, London. ↑ 27, 47.
[285] Good I.J., 1963, *Maximum entropy for hypothesis formulation, especially for multidimensional contingency tables*, Ann. Math. Statist. **34**, 911. ↑ 47, 87.
[286] Grasselli M.R., 2001, *Classical and quantum information geometry*, Ph.D. thesis, King's College, London. ↑ 67, 71, 72, 83, 84, 85.
[287] Grasselli M.R., 2001, *Dual connections in nonparametric classical information geometry*, arXiv:math-ph/0104031. ↑ 71, 84.
[288] Grasselli M.R., 2004, *Duality, monotonicity and the Wigner–Yanase–Dyson metrics*, Inf. Dim. Anal. Quant. Prob. Relat. Top. **7**, 215. arXiv:math-ph/0212022. ↑ 83.
[289] Grasselli M.R., 2008, *Dual connections in nonparametric classical information geometry*, Ann. Inst. Stat. Math. DOI:**10.1007/s10463-008-0191-3**. ↑ 71, 84.
[290] Grasselli M.R., Streater R.F., 2000, *Quantum information manifolds for ϵ-bounded forms*, Rep. Math. Phys. **46**, 325. ↑ 72.
[291] Grasselli M.R., Streater R.F., 2001, *On the uniqueness of the Chentsov metric in quantum information geometry*, Inf. Dim. Anal. Quant. Prob. Rel. Top. **4**, 173. arXiv:math-ph/0006030. ↑ 83, 84.
[292] Grasselli M.R., Streater R.F., 2001, *The uniqueness of the Chentsov metric*, in: Sollich P., Hughston L.P., Streater R.F.,

Coolen A.C.C. (eds.), *Disordered and complex systems (King's College, London, UK, 10-14 July 2000)*, AIP Conf. Proc. **553**, AIP, Melville, p.165. www.math.mcmaster.ca/∼grasselli/bkmshort.pdf. ↑ 83.
[293] Gudder S.P., 1973, *Convex structures and operational quantum mechanics*, Commun. Math. Phys. **29**, 249. ↑ 27.
[294] Györfi L., Nemetz T., 1978, *f-dissimilarity: a generalization of the affinity of several distribution*, Ann. Inst. Statist. Math. **30**, 105. ↑ 30.
[295] Haag R., Hugenholtz N.M., Winnink M., 1967, *On the equilibrium states in quantum statistical mechanics*, Commun. Math. Phys. **5**, 215. euclid:cmp/1103840050. ↑ 3, 87.
[296] Haagerup U., 1973, *The standard form of von Neumann algebras*, Preprint Ser. 1973 No. **15**, Københavns Universitet Matematisk Institut, København. ↑ 15.
[297] Haagerup U., 1975, *The standard form of von Neumann algebras*, Math. Scand. **37**, 271. www.mscand.dk/article.php?id=2275. ↑ 15.
[298] Haagerup U., 1979, *L^p-spaces associated with an aribitrary von Neumann algebra*, in: *Algèbres d'opérateurs et leurs applications en physique mathématique (Proc. Colloques Internationaux, Marseille 20-24 juin 1977)*, Colloques Internationaux C.N.R.S. **274**, Éditions du C.N.R.S., Paris, p.175. dmitripavlov.org/scans/haagerup.pdf. ↑ 20.
[299] Haagerup U., 1979, *Operator valued weights in von Neumann algebras I*, J. Funct. Anal. **32**, 175. libgen.org:scimag/get.php?doi=10.1016/0022-1236(79)90053-3. ↑ 55.
[300] Hadjisavvas N., 1981, *Distance between states and statistical inference in quantum theory*, Ann. Inst. Henri Poincaré A **38**, 167. numdam:AIHPA_1981__35_4_287_0.pdf. ↑ 27.
[301] Hadjisavvas N., 1986, *Metrics on the set of states of a W^*-algebra*, Lin. Alg. Appl. **84**, 281. ↑ 27.
[302] Hahn H., 1927, *Über lineare Gleichungssysteme in linearen Räumen*, J. Reine Angew. Math. **157**, 214. gdz.sub.uni-goettingen.de/dms/load/img/?PPN=GDZPPN002170183&IDDOC=252029. ↑ 35.
[303] Haldane J.B.S., 1951, *A class of efficient estimators of a parameter*, Bull. Int. Statist. Inst. **33**, 234. ↑ 27.
[304] Halmos P.R., Savage L.J., 1949, *Applications of the Radon–Nikodym theorem to the theory of sufficient statistics*, Ann. Math. Statist. **20**, 225. ↑ 23.
[305] Hansen F., 2006, *Characterization of symmetric monotone metrics on the state space of quantum systems*, Quant. Inf. Comput. **6**, 597. arXiv:math-ph/0601056. ↑ 67, 78.
[306] Hansen F., 2008, *Metric adjusted skew information*, Proc. Nat. Acad. Sci. U.S.A. **105**, 9909. arXiv:math-ph/0607049. ↑ 77, 82.
[307] Hansen F., Pedersen G.K., 1982, *Jensen's inequality for operators and Löwner's theorem*, Math. Ann. **258**, 229. ↑ 31.
[308] Harremoës P., Tishby N., 2007, *The information bottleneck revisited or how to choose a good distortion measure*, in: *Proceedings of the 2007 IEEE International Symposium on Information Theory*, IEEE Information Theory Society, Washington, p.566. ↑ 47.
[309] Hartigan J.A., 1964, *Invariant prior distributions*, Ann. Math. Statist. **35**, 836. ↑ 75.
[310] Hartigan J.A., 1965, *The asymptotically unbiased prior distribution*, Ann. Math. Statist. **36**, 1137. ↑ 75.
[311] Hartigan J.A., 1967, *The likelihood and invariance principles*, J. Roy. Statist. Soc. **29**, 533. ↑ 75.
[312] Hartigan J.A., 1967, *Representation of similarity matrices by trees*, J. Am. Stat. Assoc. **62**, 1140. ↑ 27.
[313] Hasegawa H., 1993, *α-divergence of the non-commutative information geometry*, Rep. Math. Phys. **33**, 87. ↑ 43, 48, 50, 72, 81, 82.
[314] Hasegawa H., 1995, *Noncommutative extension of information geometry*, in: Belavkin V.P., Hirota O., Hudson R.L. (eds.), *Proceedings of an International Workshop on Quantum Communications and Measurement, July 11-16, 1994, in Nottingham, England*, Plenum, New York, p.327. ↑ 72, 82.
[315] Hasegawa H., 1997, *Exponential and mixture families in quantum statistics: dual structure and unbiased parameter estimation*, Rep. Math. Phys. **39**, 49. ↑ 82.
[316] Hasegawa H., 2003, *Dual geometry of the Wigner–Yanase–Dyson information content*, Inf. Dim. Anal. Quant. Prob. Relat. Top. **6**, 413. ↑ 81, 83.
[317] Hasegawa H., Petz D., 1997, *Non-commutative extension of information geometry II*, in: Hirota O. et al. (eds.), *Quantum communication, computing and measurements*, Plenum, New York, p.109. www.renyi.hu/∼petz/pdf/75.pdf. ↑ 81, 83.
[318] Hasegawa M., Perlman M.D., 1974, *On the existence of a minimal sufficient subfield*, Ann. Statist. **2**, 1049 (corr. 1975, Ann. Statist. **3**, 1371). ↑ 23.
[319] Hausdorff F., 1914, *Grundzüge der Mengenlehre*, von Veit, Leipzig. libgen.org:732D04465AF921B6C8CEF79367B97EBB (Engl. transl.: 1962, *Set theory*, Chelsea, New York. libgen.org:8F127D237902B6876C73F76B15B30877). ↑ 4.
[320] Havrda J., Chárvat F., 1967, *Quantification method of classification processes. The concept of structural α-entropy*, Kybernetika **3**, 30. www.kybernetika.cz/content/1967/1/30/paper.pdf. ↑ 50.
[321] Hellinger E., 1909, *Neue Begründung der Theorie quadratischen Formen von unendlich vielen Veränderlichen*, J. Reine Angew. Math. **36**, 210. ↑ 29, 51.
[322] Helstrom C.W., 1967, *Minimum mean-square [error estimation/error of estimates] in quantum statistics*, Phys. Lett. A **25**, 101. ↑ 67, 81.
[323] Helstrom C.W., 1968, *The minimum variance of estimates in quantum signal detection*, IEEE Trans. Inf. Theor. **14**, 234. ↑ 81.
[324] Helstrom C.W., 1969, *Quantum detection and estimation theory*, J. Stat. Phys. **1**, 231. ↑ 81.
[325] Hessenberg G., 1917, *Vektorielle Begründung der Differentialgeometrie*, Math. Ann. **78**, 187. ↑ 60, 61.
[326] Hiai F., 2010, *Matrix analysis: matrix monotone functions, matrix means, and majorization*, Interd. Inf. Sci. **16**, 139. ↑ 31.
[327] Hiai F., Mosonyi M., Petz D., Bény C., 2011, *Quantum f-divergences and error correction*, Rev. Math. Phys. **23**, 691. arXiv:1008.2529. ↑ 31.
[328] Hiai F., Ohya M. (Ōya M.), Tsukada M., 1981, *Sufficiency, KMS condition and relative entropy in von Neumann algebras*, Pacific J. Math. **96**, 99. ↑ 54, 55.
[329] Hiai F., Ohya M. (Ōya M.), Tsukada M., 1983, *Sufficiency and relative entropy in $*$-algebras with applications in quantum systems*, Pacific J. Math. **107**, 117. ↑ 50, 51.
[330] Hörmander L., 1955, *Sur la fonction d'appui des ensembles convexes dans un espace localement convexe*, Ark. Matem. **3**, 181. ↑ 33.
[331] Huber P.J., 2002, *Discussion*, Ann. Statist. **30**, 1289. ↑ 24.
[332] Hübner M., 1992, *Explicit computation of the Bures distance for density matrices*, Phys. Lett. A **163**, 239. ↑ 81.
[333] Huijsmans C.B., de Pagter B., 1982, *Ideal theory in f-algebras*, Trans. Amer. Math. Soc. **269**, 225. www.ams.org/journals/tran/1982-269-01/S0002-9947-1982-0637036-5/S0002-9947-1982-0637036-5.pdf. ↑ 21.
[334] Huijsmans C.B., de Pagter B., 1984, *Subalgebras and Riesz subspaces of an f-algebra*, Proc. London Math. Soc. **48**, 161.

libgen.org:scimag/get.php?doi=10.1112/plms/s3-48.1.161. ↑ 21.
[335] Imparato D., 2008, *Exponential models and Fisher information*, Ph.D. thesis, Politecnico di Torino, Torino. ↑ 85.
[336] Imparato D., Trivellato B., 2010, *Geometry of extended exponential models*, in: Gibilisco P. et al (eds.), *Algebraic and geometric methods in statistics*, Cambridge University Press, Cambridge, p.307. libgen.org:4059765D692D2A0A5583B7E032CE1895. ↑ 85.
[337] Ingarden R.S., 1981, *Information geometry in function spaces of classical and quantum finite statistical systems*, Intern. J. Engineering Sci. **19**, 1609. ↑ 67.
[338] Ingarden R.S., Janyszek H., Kossakowski A., Kawaguchi T., 1982, *Information geometry of quantum statistical systems*, Tensor (N.S.) **37**, 105. ↑ 67, 79.
[339] Ingarden R.S., Kossakowski A., 1975, *On the connection of nonequilibrium information thermodynamics with non-hamiltonian quantum mechanics of open systems*, Ann. Phys. **89**, 451. ↑ 86.
[340] Ingarden R.S., Urbanik K., 1962, *Quantum informational thermodynamics*, Acta Phys. Polon. **21**, 281. ↑ 86.
[341] Ishwar P., Moulin P., 2005, *On the existence and characterization of the maxent distribution under general moment inequality constraints*, IEEE Trans. Inf. Theor. **51**, 3322. arXiv:cs/0506013. ↑ 86.
[342] Janyszek H., 1986, *On the geometrical structure of the generalized quantum Gibbs states*, Rep. Math. Phys. **24**, 11. ↑ 67.
[343] Jauch J.M., Misra B., Gibson A.G., 1968, *On the asymptotic condition of scattering theory*, Helv. Phys. Acta **41**, 513. ↑ 27.
[344] Jaynes E.T., 1957, *Information theory and statistical mechanics*, Phys. Rev. **106**, 620. bayes.wustl.edu/etj/articles/theory.1.pdf. ↑ 86.
[345] Jaynes E.T., 1957, *Information theory and statistical mechanics II*, Phys. Rev. **108**, 171. ↑ 67, 86.
[346] Jaynes E.T., 1967, *Foundations of probability theory and statistical mechanics*, in: Bunge M. (ed.), *Delaware seminars in the foundations of science*, Vol.1, Springer, Berlin, p.77. ↑ 87.
[347] Jaynes E.T., 1968, *Prior probabilities*, IEEE Trans. Sys. Sc. Cyb. **3**, 227. ↑ 87.
[348] Jaynes E.T., 1979, *Where do we stand on maximum entropy?*, in: Levine R.D., Tribus M. (eds.), *The maximum entropy formalism*, MIT Press, Cambridge, p.15. bayes.wustl.edu/etj/articles/stand.on.entropy.pdf. ↑ 86.
[349] Jaynes E.T., 1982, *On the rationale of maximum-entropy methods*, Proc. IEEE **70**, 939. ↑ 87.
[350] Jaynes E.T., 2003, *Probability theory: the logic of science*, Cambridge University Press, Cambridge. ↑ 86.
[351] Jeffreys H., 1946, *An invariant form for the prior probability in estimation problems*, Proc. Roy. Soc. London A **186**, 453. ↑ 27, 67, 74.
[352] Jensen J.L.W.V., 1905, *Om konvekse funktioner og uligheder mellem middelværdier*, Nyt. Tiddskr. Mat. B **16**, 49. ↑ 28.
[353] Jensen J.L.W.V., 1906, *Sur les fonctions convexes et les inégalités entre es valeurs moyennes*, Acta Math. **30**, 175. ↑ 28.
[354] Jensen U., 1995, *Review of the derivation and calculation of Rao distances with an application to portfolio theory*, in: Maddala P., Phillips G.S., Srinivasan T. (eds.), *Advances in econometrics and economics: essays in honor of C.R. Rao*, Blackwell, Cambridge, p.413. ↑ 67.
[355] Jenčová A., 2001, *Dualistic properties of the manifold of quantum states*, in: Sollich P., Hughston L.P., Streater R.F., Coolen A.C.C. (eds.), *Disordered and complex systems (King's College, London, UK, 10-14 July 2000)*, AIP Conf. Proc. **553**, AIP, Melville, p.147. ↑ 83, 84.
[356] Jenčová A., 2001, *Geometry of quantum states: dual connections and divergence functions*, Rep. Math. Phys. **47**, 121. www.mat.savba.sk/∼jencova/geom.pdf. ↑ 83, 84.
[357] Jenčová A., 2003, *Affine connections, duality and divergences for a von Neumann algebra*, arXiv:math-ph/0311004. ↑ 48.
[358] Jenčová A., 2003, *Flat connections and Wigner–Yanase–Dyson metrics*, Rep. Math. Phys. **52**, 331. www.mat.savba.sk/∼jencova/wyd.pdf. ↑ 2, 67, 83, 84.
[359] Jenčová A., 2004, *Generalized relative entropies as contrast functionals on density matrices*, Int. J. Theor. Phys. **43**, 1635. www.mat.savba.sk/∼jencova/ijtp.pdf. ↑ 2, 67, 83, 84.
[360] Jenčová A., 2005, *Quantum information geometry and non-commutative L_p spaces*, Inf. Dim. Anal. Quant. Prob. Relat. Top. **8**, 215. www.mat.savba.sk/∼jencova/lpspaces.pdf. ↑ 2, 40, 41, 48, 51, 54, 57, 58, 85.
[361] Jenčová A., 2006, *A construction of a nonparametric quantum information manifold*, J. Funct. Anal. **239**, 1. arXiv:math-ph/0511065. ↑ 2, 57, 67, 72, 73, 85.
[362] Jenčová A., 2010, *On quantum information manifolds*, in: Gibilisco P. et al (eds.), *Algebraic and geometric methods in statistics*, Cambridge University Press, Cambridge, p.265. libgen.org:4059765D692D2A0A5583B7E032CE1895. ↑ 2, 72, 73.
[363] Jenčová A., Petz D., 2006, *Sufficiency in quantum statistical inference*, Commun. Math. Phys. **263**, 259. arXiv:math-ph/0412093. ↑ 87.
[364] Jenčová A., Petz D., 2006, *Sufficiency in quantum statistical inference. A survey with examples*, Inf. Dim. Anal. Quant. Prob. Relat. Top. **9**, 331. arXiv:quant-ph/0604091. ↑ 87.
[365] Jenčová A., Ruskai M.B., 2010, *A unified treatment of convexity of relative entropy and related trace functions, with conditions for equality*, Rev. Math. Phys. **22**, 1099. arXiv:0903.2895. ↑ 82.
[366] Jones L.K., Byrne C.L., 1990, *General entropy criteria for inverse problems, with applications to ada compression, pattern classification and cluster analysis*, IEEE Trans. Inf. Theory **36**, 23. ↑ 36, 40.
[367] Kadison R.V., 1952, *A generalized Schwarz inequality and algebraic invariants for operator algebras*, Ann. Math. **56**, 494. ↑ 31.
[368] Kagan A.M., 1963, *K teorii informacionnogo kolichestva Fishera*, Dokl. Akad. Nauk SSSR **151**, 277. ↑ 29.
[369] Kakutani S., 1939, *Weak topology and regularity of Banach spaces*, Proc. Imp. Acad. Tōkyō **15**, 169. ↑ 35.
[370] Kakutani S., 1941, *Concrete representation of abstract (L)-spaces and the mean ergodic theorem*, Ann. Math. **42**, 523. ↑ 7.
[371] Kakutani S., 1941, *Concrete representation of abstract (M)-spaces (a characterization of the space of continuous functions)*, Ann. Math. **42**, 994. ↑ 5, 7.
[372] Kakutani S., 1948, *On equivalence of infinite product measures*, Ann. Math. **49**, 214. ↑ 29, 51.
[373] Kantorovich L.V., 1935, *O poluuporyadochennykh lineĭnykh prostranstvakh i ikh primeneniyakh v teorii lineĭnykh operaciĭ*, Dokl. Akad. Nauk SSSR **4**, 11. ↑ 4.
[374] Kantorovich L.V., 1937, *Lineare halbgeordnete Räume*, Matem. sb. N.S. **2**, 121. mathnet.ru:msb5563. ↑ 4.
[375] Kaplansky I., 1951, *Projections in Banach algebras*, Ann. Math. **53**, 235. ↑ 21.
[376] Kaplansky I., 1955, *Rings of operators*, Lecture notes, Mathematics Department of Chicago University, Chigago (rev. ed.: 1968, Benjamin, New York. libgen.org:7CA2C85D0EE64EEF7E1E169CA12A8514). ↑ 21.
[377] Kappos D.A., 1949, *Zur mathematischen Begründung der Wahrscheinlichkeitstheorie*, Sitzungsber. Math.-Naturw. Kl. Bayer. Akad. Wiss. München **1948**, 27. ↑ 23.
[378] Kappos D.A., 1951, *Über die Unabhängigkeit in der Wahrscheinlichkeitstheorie*, Sitzungsber. Math.-Naturw. Kl. Bayer. Akad. Wiss. München **1950**, 157. ↑ 23.

[379] Kappos D.A., 1960, *Strukturtheorie der Wahrscheinlichkeits-Felder und -Räume*, Ergebnisse der Mathematik und ihrer Grenzgebiete. Neue Folge. Heft **24**, Springer, Berlin. ↑ 23.
[380] Kappos D.A., 1969, *Probability algebras and stochastic spaces*, Academic Press, New York. ↑ 23.
[381] Kass R.E., 1980, *The riemannian structure of model spaces: a geometrical approach to inference*, Ph.D. thesis, University of Chicago, Chicago. ↑ 67.
[382] Kass R.E., 1989, *The geometry of asymptotic inference (with comments and rejoinder)*, Statist. Sci. **4**, 188. ↑ 67.
[383] Kass R.E., Vos P.W., 1997, *Geometrical foundations of asymptotic inference*, Wiley, New York. ↑ 66, 67.
[384] Kato T., 1952, *On the perturbation theory of closed linear operators*, J.Math. Soc Japan **4**, 323. ↑ 72.
[385] Kato T., 1966, *Perturbation theory for linear operators*, Springer, Berlin. ↑ 72.
[386] Kelley J.L., 1966, *Decomposition and representation theorems in measure theory*, Math. Ann. **163**, 89. gdz.sub.uni-goettingen.de/dms/load/img/?PPN=GDZPPN00229561X&IDDOC=62555. ↑ 10.
[387] Kelley J.L., Namioka I., 1963, *Linear topological spaces*, Van Nostrand, Princeton. libgen.org:31A7137D85FB6927FE7B07DEC8677C29. ↑ 9.
[388] Khosravifard M., Fooladivanda D., Gulliver T.A., 2007, *Confliction of the convexity and metric properties in f-divergences*, IEICE Trans. Fund. Electr. Comm. Compt. Sci. A **90**, 1848. ↑ 29.
[389] Kiwiel K.C., 1997, *Free-steering relaxation methods for problems with strictly convex costs and linear constraints*, Math. Oper. Res. **22**, 326. ↑ 35, 37.
[390] Kiwiel K.C., 1997, *Proximal minimization methods with generalized Bregman functions*, Soc. Industr. Appl. Math. J. Contr. Optim. **35**, 1142. ↑ 35.
[391] Kiwiel K.C., 1998, *Generalized Bregman projections in convex feasibility problems*, J. Optim. Theor. Appl. **96**, 139. ↑ 35.
[392] Klee V., 1953, *Convex bodies and periodic homeomorphisms in Hilbert space*, Trans. Amer. Math. Soc. **74**, 10. ↑ 35.
[393] Klingenberg W.P.A., 1982, *Riemannian geometry*, de Gruyter, Berlin (2nd rev. ed., 1995). ↑ 60.
[394] Kolmogorov A.N., 1933, *Grundbegriffe der Wahrscheinlichkeitrechnung*, Ergebnisse der Mathematik und Ihrer Grenzgebiete Bd. **2**, Springer, Berlin. (Engl. transl. 1950, *Foundations of the theory of probability*, Chelsea, New York. libgen.org:2BFE8A8E7D7907A81547C003E748B7BD). ↑ 22.
[395] Koopman B.O., 1936, *On distributions admitting a sufficient statistic*, Trans. Amer. Math. Soc. **39**, 399. ↑ 85.
[396] Kosaki H., 1980, *Canonical L^p-spaces associated with an arbitrary abstract von Neumann algebra*, Ph.D. thesis, University of California, Los Angeles. dmitripavlov.org/scans/kosaki-thesis.pdf. ↑ 15, 16, 20.
[397] Kosaki H., 1982, *Interpolation theory and the Wigner–Yanase–Dyson–Lieb concavity*, Commun. Math. Phys. **87**, 315. ↑ 2, 31, 32, 54, 85.
[398] Kosaki H., 1984, *Applications of the complex interpolation method to a von Neumann algebra: non-commutative L^p-spaces*, J. Funct. Anal. **56**, 29. ↑ 20.
[399] Kosaki H., 1984, *On the continuity of the map $\varphi \to |\varphi|$ from the predual of a W^*-algebra*, J. Funct. Anal. **59**, 123. ↑ 19.
[400] Kosaki H., 1986, *Relative entropy for states: a variational expression*, J. Op. Theor. **16**, 335. ↑ 31, 54.
[401] Koshevnik Yu.A., Levit B.Ya., 1976, *On a non-parametric analogue of the information matrix*, Theor. Prob. Appl. **21**, 738. ↑ 69.
[402] Kostecki R.P., 2010, *Quantum theory as inductive inference*, in: Mohammad-Djafari A., Bercher J.-F., Bessière P. (eds.), *Proceedings of the 30th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, AIP Conf. Proc. **1305**, Springer, Berlin, p.24. arXiv:1009.2423. ↑ 48.
[403] Kostecki R.P., 2011, *The general form of γ-family of quantum relative entropies*, Open Sys. Inf. Dyn. **18**, 191. arXiv:1106.2225. ↑ 40, 42, 48.
[404] Kostecki R.P., 2013, *W^*-algebras and noncommutative integration*, arXiv:1307.4818. ↑ 1, 3, 9, 18, 23, 24, 27, 48.
[405] Koszul J.L., 1961, *Domaines bornés homogènes et orbites de groupes de transformations affines*, Bull. Soc. Math. France **89**, 515. ↑ 65.
[406] Kozlov V.P., 1966, *Emkost' mnozhestva v prostranstve signalov i rimanova metrika*, Dokl. Akad. Nauk SSSR **166**, 779. ↑ 74.
[407] Kraft C.H., 1955, *Some conditions for consistency and uniform consistency of statistical procedures*, Univ. Calif. Publ. Statist. **1**, 125. ↑ 50.
[408] Krasnosel'skiĭ M.A., Rutickiĭ Ya.B., 1958, *Vypuklye funkcii i prostranstva Orlicha*, Gosudarstvennoe izdatel'stvo fiziko-matematicheskoĭ literatury, Moskva. libgen.org:CFC3DB6B89C6821457044734272DC224 (engl. trans. 1961, *Convex functions and Orlicz spaces*, Nordhoff, Groningen. libgen.org:f74c452095b800808a4c7e578e83662d). ↑ 3, 69.
[409] Kraus F., 1936, *Über konvexe Matrixfunktionen*, Math. Z. **41**, 18. ↑ 30.
[410] Kreĭn M.G, Kreĭn S.G., 1940, *On an inner characteristic of the set of all continuous functions defined on a bicompact Hausdorff space*, Dokl. Akad. Nauk SSSR **27**, 427. ↑ 5.
[411] Kriegl A., Michor P.W., 1997, *The convenient setting for global analysis*, Mathematical Surveys and Monographs **53**, American Mathematical Society, Providence. www.mat.univie.ac.at/∼michor/apbookh-ams.pdf. ↑ 32, 58.
[412] Kronfli N.S., 1970, *States on generalised logics*, Int. J. Theor. Phys. **3**, 191. ↑ 27.
[413] Kubo F., Andō T., 1980, *Means of positive linear operators*, Math. Ann. **246**, 205. ↑ 79.
[414] Kubo R., 1957, *Statistical-mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems*, J. Phys. Soc. Japan **12**, 570. dx.doi.org/10.1143/JPSJ.12.570. ↑ 79.
[415] Kulhavý R., 1996, *Recursive nonlinear estimation: a geometric approach*, Springer, Berlin. ↑ 67.
[416] Kullback S., 1959, *Information theory and statistics*, Wiley, New York (2nd ed. 1968). libgen.org:66875BFC74B54B9B73E51375EF11FEDB. ↑ 32, 47.
[417] Kullback S., Khairat M.A., 1966, *A note on minimum discrimination information*, Ann. Math. Statist. **37**, 279. ↑ 32.
[418] Kullback S., Leibler R.A., 1951, *On information and sufficiency*, Ann. Math. Statist. **22**, 79. euclid:aoms/1177729694. ↑ 27, 47.
[419] Kumagai W., 2011, *A characterization of extended monotone metrics*, Lin. Alg. Appl. **434**, 224. ↑ 78.
[420] Kunze W., 1990, *Noncommutative Orlicz spaces and generalized Arens algebras*, Math. Nachr. **147**, 123. libgen.org:scimag/get.php?doi=10.1002/mana.19901470114. ↑ 73.
[421] Kurdila A.J., Zabarankin M., 2005, *Convex functional analysis*, Birkhäuser, Basel. ↑ 32.
[422] Kurose T., 1990, *Dual connections and affine geometry*, Math. Z. **203**, 115. ↑ 66.
[423] Kurose T., 1993, *Dual connections and projective geometry*, Hiroshima Math. J. **23**, 327. ↑ 66.
[424] Labuschagne L.E., Majewski W.A., 2008, *Quantum L_p and Orlicz spaces*, in: García J.C., Quezada R., Sontz S.B.(eds.), *Quantum probability and related topics. Proceedings of the 28th Conference (CIMAT-Guanajuato, Mexico, 2-8 September 2007)*, World Scientific, Singapore, p.176. arXiv:0902.4327. ↑ 73.

[425] Lai P.T., 1973, *L'analogue dans $\mathcal{C}^p$ des theorem de convexité de M. Riesz et G.O. Thorin*, Studia Math. **46**, 111. ↑ 18.
[426] Landau L.D., 1927, *Das Dämpfungsproblem in der Wellenmechanik*, Z. Phys. **45**, 430. ↑ 24.
[427] Lanford O.E. III, Robinson D.W., 1968, *Statistical mechanics of quantum spin systems III*, Commun. Math. Phys. **9**, 327. ↑ 87.
[428] Lang S., 1962, *Differentiable manifolds*, Springer, Berlin (nth rev. ed., 1999, *Fundamentals of differential geometry*, Springer, Berlin). ↑ 60.
[429] Laurant F., Oheix M., Raoult J.P., 1969, *Tests d'hypothèses*, Ann. Inst. Henri Poincaré B **5**, 385. ↑ 24.
[430] Lauritzen S.L., 1987, *Conjugate connections in statistical theory*, in: Dodson C.T.J. (ed.), *Geometrization of statistical theory*, Univ. Lancaster Dept. Math. Publ., Lancaster, p.33. ↑ 62, 63.
[431] Lauritzen S.L., 1987, *Statistical manifolds*, in: Gupta S.S. (ed.), *Differential geometry in statistical inference*, Institute of Mathematical Statistics, Hayward, p.163. ↑ 63, 64, 66, 88.
[432] Lawvere F.W., 1962, *The category of probabilistic mappings with applications to stochastic processes, statistics, and pattern recognition*, unpublished preprint. www.fuw.edu.pl/∼kostecki/scans/lawvere1962.pdf. ↑ 25.
[433] Le Besnerais G., 1993, *Méthode du maximum d'entropie sur la moyenne, critères de reconstruction d'image, et synthèse d'ouverture en radio-astronomie*, Ph.D. thesis, Université de Paris-Sud, Orsay. ↑ 32.
[434] Le Besnerais G., Bercher J.-F., Demoment G., 1999, *A new look at entropy for solving linear inverse problems*, IEEE Trans. Inf. Theor. **45**, 1565. ↑ 32.
[435] Le Cam L., 1964, *Sufficiency and approximate sufficiency*, Ann. Math. Statist. **35**, 1419. ↑ 23, 24, 25.
[436] Le Cam L., 1969, *Théorie asymptotique de la décision statistique*, Les Presses de l'Université Montréal, Montréal. ↑ 23.
[437] Le Cam L., 1986, *Asymptotic methods in statistical decision theory*, Springer, Berlin. ↑ 23.
[438] Le Cam L., Yang G.L., 1990, *Asymptotics in statistics: some basic concepts*, Springer, Berlin (2nd ed. 2000). ↑ 51.
[439] Lebesgue H., 1910, *Sur l'intégration des fonctions discontinues*, Ann. Sci. École Norm. Sup. 3^e ser. **27**, 361. numdam:ASENS_1910_3_27__361_0. ↑ 10.
[440] Léonard L., 2001, *Minimization of energy functionals applied to some inverse problems*, Appl. Math. Optim. **44**, 273. ↑ 32.
[441] Léonard L., 2001, *Minimizers of energy functionals*, Acta Math. Hungar. **93**, 281. ↑ 32.
[442] Léonard L., 2003, *Minimizers of energy functionals under not very integrable constraints*, J. Conv. Anal. **10**, 63. ↑ 32.
[443] Léonard L., 2008, *Minimization of entropy functionals*, J. Math. Anal. Appl. **346**, 186. ↑ 32.
[444] Léonard L., 2010, *Convex minimization problems with weak constraint qualifications*, J. Conv. Anal. **17**, 321. ↑ 32.
[445] Léonard L., 2010, *Entropic projections and dominating points*, ESAIM Prob. Statist. **14**, 343. ↑ 32.
[446] Lesniewski A., Ruskai M.B., 1999, *Monotone riemannian metrics and relative entropy on non-commutative probability space*, J. Math. Phys. **40**, 5702. arXiv:math-ph/9808016. ↑ 31, 62, 67, 78, 83.
[447] Levi-Civita T., 1917, *Nozione di parallelismo in una varietà qualunque e consequente specificazione geometrica della curvatura riemanniana*, Rend. Circ. Mat. Palermo **42**, 173. ↑ 60, 61.
[448] Lewin J., Lewin M., 1975, *A reformulation of the Radon-Nikodym theorem*, Proc. Amer. Math. Soc. **47**, 393. www.ams.org/journals/proc/1975-047-02/S0002-9939-1975-0376999-4/S0002-9939-1975-0376999-4.pdf. ↑ 10.
[449] Lewis A.S., 1995, *Consistency of moment systems*, Can. J. Math. **47**, 995. ↑ 32.
[450] Lieb E.H., 1973, *Convex trace functions and the Wigner–Yanase–Dyson conjecture*, Adv. Math. **11**, 267. ↑ 54.
[451] Lieb E.H., Ruskai M.B., 1973, *Proof of the strong subadditivity of quantum mechanical entropy*, J. Math. Phys. **14**, 1938. ↑ 54.
[452] Liese F., 1977, *On the existence of f-projections*, in: Csiszár I., Elias P. (eds.), *Topics in information theory*, North-Holland, Amsterdam, p.431. ↑ 32.
[453] Liese F., Vajda I., 1987, *Convex statistical distances*, Teubner–Texte zur Mathematik **95**, Teubner, Leipzig. ↑ 2, 29, 30, 32, 46, 50.
[454] Liese F., Vajda I., 2006, *On divergences and informations in statistics and information theory*, IEEE Trans. Inf. Theor. **52**, 4394. ↑ 30.
[455] Liese F., Vajda I., 2008, *f-divergences: sufficiency, deficiency and testing of hypotheses*, in: Barnett N.S., Dragomir S.S. (eds.), *Advances in inequalities from probability theory and statistics*, Nova Publishers, New York. ↑ 30.
[456] Lindblad G., 1974, *Expectations and entropy inequalities for finite quantum systems*, Commun. Math. Phys. **39**, 111. euclid:cmp/1103860161. ↑ 54.
[457] Lindblad G., 1975, *Completely positive maps and entropy inequalities*, Commun. Math. Phys. **40**, 147. ↑ 54.
[458] Lindelöf E., 1894, *Sur l'application de la méthode des apporximation successives aux équations differentielles ordinaires du premier orde*, Compt. Rend. Hebd. Séanc. Acad. Sci. **114**, 454. ↑ 62.
[459] Lindenstrauss J., Tzafriri L., 1974, *On the complemented subspaces problem*, Israel J. Math. **2**, 263. ↑ 59.
[460] Lindenstrauss J., Tzafriri L., 1977, 1979, *Classical Banach spaces*, Vol.1-2, Springer, Berlin. libgen.org:6D9FC2E3C27BA845FD629049752944A7. ↑ 69.
[461] Linhard J., Nielsen V., 1971, *Studies in statistical dynamics*, Kong. Danske Vid. Selkab. Mat.-fys. Medd. **38**, 1. ↑ 50.
[462] Lipschitz R., 1869, *Untesuchungen in Betreff der ganzen homogenen Functionen von n Differentialen*, J. reine angew. Math. **70**, 71. ↑ 61.
[463] Littaye-Petit M., Piednoir J.L., Van Cutsem B., 1969, *Exhausivité*, Ann. Inst. Henri Poincaré B. **5**, ?. ↑ 24.
[464] Lobachevskiĭ N.I., 1829, *O nachalakh geometrii*, Kazansk. vestn. **25**, 178, 228; **27**, 227. ↑ 89.
[465] Lobachevskiĭ N.I., 1830, *O nachalakh geometrii*, Kazansk. vestn. **28**, 251, 571. ↑ 89.
[466] Loebl R.I., 1974, *Injective von Neumann algebras*, Proc. Nat. Acad. Sci. U.S.A. **44**, 46. ↑ 55.
[467] Łomnicki A., 1923, *Nouveaux fondements du calcul des probabilités*, Fund. Math. **4**, 34. ↑ 22.
[468] Loomis L.H., 1947, *On the representation of σ-complete boolean algebras*, Bull. Amer. Math. Soc. **53**, 757. euclid:bams/1183510979. ↑ 9.
[469] Lovrić M., Min-Oo M., Ruh E.A., 2000, *Multivariate normal distributions parametrized as a riemannian symmetric space*, J. Multivar. Anal. **74**, 36. ↑ 89.
[470] Löwner K., 1934, *Über monotone Matrixfunctionen*, Math. Z. **38**, 177. ↑ 30.
[471] Luschgy H., Mussmann D., 1985, *Equivalent properties and completion of statistical experiments*, Sankhyā A **47**, 174. ↑ 23.
[472] Luxemburg W.A.J., 1955, *Banach function spaces*, Ph.D. thesis, Technische Hogeschool te Delft, Delft. repository.tudelft.nl/assets/uuid:252868f8-d63f-42e4-934c-20956b86783f/71308.pdf. ↑ 3, 69.
[473] Luxemburg W.A.J., Zaanen A.C., 1971, *Riesz spaces*, Vol.1, North-Holland, Amsterdam. libgen.org:54828591DD8D0A5F5E1C0464E65DB093. ↑ 21.
[474] Lyubovin V.D., 1954, *Poluuporyadochennye kol'ca samospryazhennykh operatorov*, Ph.D. thesis, Leningradskiĭ Gosudarstvennyĭ Pedagogicheskiĭ Institut, Leningrad. ↑ 21.

[475] Lyubovin V.D., 1956, *O K-prostranstve ogranichennykh samospryazhennykh operatorov*, Uch. zap. leningr. pedagog. univ. **125**, 119. ↑ 21.
[476] Mackey G.W., 1945, *On infinite-dimensional spaces*, Trans. Amer. Math. Soc. **57**, 155. www.ams.org/journals/tran/1945-057-02/S0002-9947-1945-0012204-1/S0002-9947-1945-0012204-1.pdf. ↑ 32.
[477] MacNeille H.M., 1937, *Partially ordered sets*, Trans. Amer. Math. Soc. **42**, 416. www.ams.org/journals/tran/1937-042-03/S0002-9947-1937-1501929-X/S0002-9947-1937-1501929-X.pdf. ↑ 4.
[478] Mahalanobis P.C., 1936, *On the generalized distance in statistics*, Proc. Nat. Inst. Sci. India **2**, 49. ↑ 27.
[479] Majewski W.A., Labuschagne L.E., 2014, *On applications of Orlicz spaces to statistical physics*, Ann. H. Poincaré **15**, 1197. arXiv:1302.3460. ↑ 87.
[480] Maligranda L., 1989, *Orlicz spaces and interpolation*, Universidade de Estadual de Campinas, Campinas. ↑ 69.
[481] Mandelbrojt S., 1939, *Sur les fonctions convexes*, Compt. Rend. Acad. Sci. Paris **209**, 977. ↑ 33.
[482] Markov A.A., 1906, *Raspostranenye zakona bol'shikh chisel na velichiny, zavisyashhie drug ot druga*, Izv. Fiz.-mat. Obshh. Kazan. Univ. **15**, 135. (Engl. transl. *The extension of the law of large numbers onto quantities depending on each other*, in: Sheynin O.B. (ed.), 2004, *Probability and statistics. Russian papers*, NG Verlag, Berlin, p.143). ↑ 25.
[483] Martin F., Petit J.-L., Petit-Littaye M., 1971, *Comparison des expériences*, Ann. Inst. Henri Poincaré Sèr. B **7**, 145. ↑ 24.
[484] Martin F., Vaguelsy M.D., 1969, *Propriétés asymptotiques du modèle statistique*, Ann. Inst. Henri Poincaré Sèr. B **5**, 357. ↑ 24.
[485] Masuda T., 1983, *L_p-spaces for von Neumann algebra with reference to a faithful normal semifinite weight*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **19**, 673. doi:10.2977/prims/1195182447. ↑ 19, 20.
[486] Masuda T., 1984, *A note on a theorem of A. Connes on Radon-Nikodym cocycles*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **20**, 131. doi:10.2977/prims/1195181833. ↑ 15.
[487] Matsuzuoe H., 1998, *On a realization of conformally projectively flat statistical manifolds and the divergences*, Hokkaido Math. J. **27**, 409. ↑ 66.
[488] Matumoto T., 1993, *Any statistical manifold has a contrast functions (on the C^3-functions taking the minimum at the diagonal of the product manifold)*, Hiroshima Math. J. **23**, 327. euclid:hmj/1206128255. ↑ 63.
[489] Matusita K., 1951, *On the theory of statistical decision functions*, Ann. Inst. Statist. Math. **3**, 17. ↑ 27.
[490] Matusita K., 1955, *Decision rules based on the distance, for problems of fit, two samples and estimation*, Ann. Math. Statist. **26**, 613. ↑ 29, 51.
[491] Matusita K., 1964, *Distance and decision rules*, Ann. Inst. Statist. Math. **16**, 305. ↑ 51.
[492] Mazur S., 1933, *Über konvexe Mengen in linearen normierten Räumen*, Stud. Math. **4**, 70. ↑ 34.
[493] McCarthy C.A., 1967, c_p, Israel J. Math. **5**, 249. ↑ 18.
[494] McShane E.J., 1962, *Families of measures and representations of algebras of operators*, Trans. Amer. Math. Soc. **102**, 328. www.ams.org/journals/tran/1962-102-02/S0002-9947-1962-0137002-X/S0002-9947-1962-0137002-X.pdf. ↑ 10.
[495] Michor P.W., Petz D., Andai A., 2000, *On the curvature of a certain riemannian space of matrices*, Inf. Dim. Anal. Quant. Prob. Relat. Top. **3**, 199. arXiv:math/9909157. ↑ 84.
[496] Mil'man D.P., 1938, *O nekotorykh priznakakh regulyarnosti prostranstva tipa (B)]*, Dokl. Akad. Nauk. SSSR **20**, 243. ↑ 35.
[497] Mordukhovich B.S., 2006, *Variational analysis and generalized differentiation*, Vol.1-2, Springer, Berlin. ↑ 32.
[498] Moreau J.-J., 1963, *Fonctionelles sous-différentiables*, Compt. Rend. Acad. Sci. Paris **257**, 4117. ↑ 33.
[499] Mori H., 1956, *A quantum-statistical theory of transport processes*, J. Phys. Soc. Japan **11**, 1029. ↑ 79.
[500] Morimoto T., 1963, *Markov processes and the H-theorem*, J. Phys. Soc. Jap. **12**, 328. ↑ 2, 3, 29.
[501] Morozova E.A., Chencov N.N., 1982, *Struktura semeĭstv stacionarnykh raspredeleniĭ dlya nekomutitativnoĭ cepi Markova*, in: *Matematicheskiye modeli statisticheskoĭ fiziki*, Izdatel'stvo Tyumenskogo Gosudarstvennogo Universiteta, Tyumen', p.31. ↑ 3.
[502] Morozova E.A., Chencov N.N., 1985, *Ekvivarianty v kategorii markovskikh otovrazheniĭ*, Usp. mat. nauk **40**, 189. ↑ 3.
[503] Morozova E.A., Chencov N.N., 1986, *Markovskie invarianty par veroyatnostnykh zakonov*, Preprint **1986-121**, Institut Prikladnoĭ Matematiki AN SSSR. ↑ 3.
[504] Morozova E.A., Chencov N.N., 1987, *Markov maps in noncommutative probability theory and mathematical statistics*, in: Prokhorov Yu.V. et al (eds.), *Probability theory and mathematical statistics. Proceedings of the fourth international Vilnius conference held in Vilnius, June 24–29, 1985*, Vol.2, VNU Science Press, Utrecht, p.287. ↑ 3.
[505] Morozova E.A., Chencov N.N., 1988, *Invariantnye lineĭnye svyaznosti na sovokupnostyakh raspredeleniĭ veroyatnosteĭ*, Preprint **1988-167**, Institut Prikladnoĭ Matematiki AN SSSR. ↑ 3, 75.
[506] Morozova E.A., Chencov N.N., 1989, *Markovskaya invaryantnaya geometriya na mnogoobraziyakh sostoyaniĭ*, Itogi Nauk. i Tekh. Ser. Sovrem. Probl. Mat. Nov. Dostizh. **36**, 69 (Engl. transl. 1991, *Markov invariant geometry on state manifolds*, J. Soviet Math. **56**, 2648). ↑ 3, 67, 75, 77.
[507] Morozova E.A., Chencov N.N., 1991, *Estestvennaya geometriya semeĭstv veroyatnostnykh zakonov*, Itogi Nauk. i Tekhn. Ser. Sovr. Probl. Mat. Fundam. Napravlen. **83**, 133. ↑ 1, 3, 67, 75.
[508] Morse M., Transue W., 1950, *Functionals F bilinear over the product $A \times B$ of two pseudo-normed vector spaces*, Ann. Math. **51**, 576. libgen.org:scimag/get.php?doi=10.2307/1969370. ↑ 3, 29, 69.
[509] Morse N., Sacksteder R., 1966, *Statistical isomorphism*, Ann. Math. Statist. **37**, 203. ↑ 24, 25.
[510] Murata N., Takenouchi T., Kanamori T., Eguchi S., 2004, *Information geometry of U-boost and Bregman divergence*, Neural Comput. **16**, 1437. ↑ 67.
[511] Muratov M.A., 1978, *Nekommutativnye prostranstva Orlicha*, Dokl. Akad. Nauk UzSSR **1978:6**, 11. www.fuw.edu.pl/∼kostecki/scans/muratov1978.pdf. ↑ 73.
[512] Muratov M.A., 1979, *Norma Lyuksemburga v prostranstve Orlicha izmerimykh operatorov*, Dokl. Akad. Nauk UzSSR **1979:1**, 5. www.fuw.edu.pl/∼kostecki/scans/muratov1979.pdf. ↑ 73.
[513] Muratov M.A., Chilin V.I., 2007, *Algebry izmerimykh i lokal'no izmerimykh operatorov*, Instytut matematyky NAN Ukraïny, Kyïv. ↑ 17.
[514] Murray F.J., von Neumann J., 1936, *On rings of operators*, Ann. Math. **37**, 116. ↑ 14.
[515] Murray M.K., Rice J.W., 1993, *Differential geometry and statistics*, Chapman and Hall, London. ↑ 65, 67, 89.
[516] Musielak J., 1983, *Orlicz spaces and modular spaces*, LNM **1034**, Springer, Berlin. libgen.org:9D2BBB211D18EA4387AA1D846E73344E. ↑ 69.
[517] Mussmann D., 1972, *Vergleich von Experimenten im schwach dominierten Fall*, Z. Wahrschein. Geb. **24**, 295. ↑ 23.
[518] Mussmann D., 1979, *Sufficiency and f-divergences*, Studia Sci. Math. Hungar. **14**, 37. ↑ 30.
[519] Nagaoka H., 1982, *[Foundations of statistical geometry and applications to robust estimation]*, M.Sc. thesis, Tōkyō daigaku, Tōkyō. ↑ 48.

[520] Nagaoka H., 1987, *?*, in: *Proceedings of 10th Symposium on Information Theory and Its Applications (SITA), Enoshima, Kanagawa, Japan, November 19-21, 1987*, p.241 (Engl. transl. 2005, *On Fisher information of quantum statistical models*, in: Hayashi M. (ed.), *Asymptotic theory of quantum statistical inference*, World Scientific, Singapore). ↑ 67.
[521] Nagaoka H., 1989, *A new approach to Cramér–Rao bounds for quantum state estimation*, IEICE Technical Report No. 228, IT **89-42**, 9. ↑ 67, 81.
[522] Nagaoka H., 1994, *Differential geometrical aspects of quantum state estimation and relative entropy*, Technical report METR **94-14**, University of Tōkyō, Tōkyō. ↑ 2, 82, 84.
[523] Nagaoka H., 1995, *Differential geometrical aspects of quantum state estimation end relative entropy*, in: Belavkin V.P., Hirota O., Hudson R.L. (eds.), *Proceedings of an International Workshop on Quantum Communications and Measurement, July 11-16, 1994, in Nottingham, England*, Plenum, New York, p.449. ↑ 2, 82, 84.
[524] Nagaoka H., Amari S.-i., 1982, *Differential geometry of smooth families of probability distributions*, Technical report METR **82-7**, University of Tōkyō, Tōkyō. ↑ 48, 62, 63, 64, 66, 75, 76.
[525] Nakamura M., Takesaki M., Umegaki H., 1960, *A remark on the expectations of operator algebras*, Kōdai Math. Sem. Rep. **12**, 82. euclid:kmj/1138844264. ↑ 26.
[526] Nakamura M., Umegaki H., 1961, *A note on the entropy on operator algebras*, Proc. Jap. Acad. **37**, 149. ↑ 87.
[527] Nakano H., 1950, *Modulared semi-ordered linear spaces*, Maruzen, Tōkyō. ↑ 3, 5, 69.
[528] Naudts J., 1974, *A generalised entropy function*, Commun. Math. Phys. **37**, 175. ↑ 87.
[529] Naĭmark M.A., 1943, *Ob odnom predstavlenii additivnykh operatornykh funkciĭ mnozhestv*, Dokl. Akad. Nauk. SSSR **41**, 373. ↑ 26.
[530] Navaza J., 1985, *On the maximum entropy estimate of electron density function*, Acta Crystallogr. A **41**, 232. ↑ 32.
[531] Navaza J., 1986, *The use of non-local constraints in maximum-entropy electron density reconstruction*, Acta Crystallogr. A **42**, 212. ↑ 32.
[532] Neirotti J.P., Raggio G.A., 1993, *Maximizing relative entropy at given energies*, Lett. Math. Phys. **29**, 33. ↑ 57, 87.
[533] Nelson E., 1974, *Notes on non-commutative integration*, J. Funct. Anal. **15**, 103. ↑ 16, 17.
[534] Neveu J., 1964, *Bases mathématiques du calcul des probabilités*, Masson et Cie, Paris (1970, 2nd ed. libgen.org:3D5B8FB717180E18B98C3C6488766E83; Engl. transl.: 1965, *Mathematical foundations of the calculus of probability*, Holden–Day, San Francisco). ↑ 10, 24.
[535] Newton N.J., 2012, *An infinite-dimensional statistical manifold modelled on Hilbert space*, J. Funct. Anal. **263**, 1661. ↑ 67.
[536] Neyman J., 1935, *Sur un teorema concernente le considette statistiche suficienti*, Giorn. Ist. Ital. Att. **6**, 320. ↑ 22.
[537] Nikodým O.M., 1930, *Sur uner généralisation des intégrales de M.J.Radon*, Fund. Math. **15**, 131. matwbn.icm.edu.pl/ksiazki/fm/fm15/fm15114.pdf. ↑ 10.
[538] Nikodým O.M., 1931, *Contribution à la théorie des fonctionnelles linéaires en connection avec la théorie de la mesure des ensembles abstraits*, Mathematica, Cluj **5**, 130; Bull Soc. Ştiinţ. Cluj **6**, 79. ↑ 9, 22.
[539] Nomizu K., Sasaki T., 1994, *Affine differential geometry*, Cambridge University Press, Cambridge. ↑ 62.
[540] Nomizu K., Simon U., 1992, *Notes on conjugate connections*, in: Dillen F., Verstraelen L. (eds.), *Geometry and topology of submanifolds IV*, World Scientific, Singapore, p.152. ↑ 62.
[541] Norden A.P., 1937, *O parakh soprazhennykh pereneseniĭ*, Trudy semin. vekt. tenzorn. anal. **4**, 206. ↑ 62.
[542] Norden A.P., 1945, *O parakh soprazhennykh parallel'nykh pereneseniĭ v n-mernykh prostranstvakh*, Dokl. Akad. Nauk SSSR **49**, 9. ↑ 62.
[543] Norden A.P., 1948, *O vnutrennikh geometriyakh poverkhnosteiĭ proektivnogo prostranstva*, Trudy semin. vekt. tenzorn. anal. **6**, 125. ↑ 62.
[544] Norden A.P., 1949, *O vnutrennikh geometriyakh poverkhnosteiĭ proektivnogo prostranstva*, Trudy semin. vekt. tenzorn. anal. **7**, 31. ↑ 62.
[545] Norden A.P., 1950, *Prostranstva affinoĭ svyaznosti*, Gosudarstvennoe Izdatel'stvo Tekhniko-Teoreticheskoĭ Literatury, Moskva/Leningrad (2nd. rev. ed., 1976, Nauka, Moskva. libgen.org:05EC4513F05BA7E7F05195880858AA29). ↑ 62.
[546] Norden A.P., 1959, *O vnutrenneĭ geometrii 2-go roda na giperploskosti affinnogo prostranstva*, in: Shirokov N.A., Shirokov A.N., *Affinnaya differencial'naya geometriya*, Fizmatgiz, Moskva. ↑ 62.
[547] Ochs W., Bayer W., 1973, *Quantum states with maximum information entropy II*, Z. Naturf. **28a**, 1571. ↑ 86.
[548] Ochs W., Spohn H., 1978, *A characterization of the Segal entropy*, Rep. Math. Phys. **14**, 75. ↑ 87.
[549] Odzijewicz A., Ratiu T.S., 2003, *Banach Lie–Poisson spaces and reduction*, Commun. Math. Phys. **243**, 1. arXiv:math/0210207. ↑ 3.
[550] Ohya M. (Ōya M.), Petz D., 1993, *Quantum entropy and its use*, Springer, Berlin (2nd. ed., 2004). ↑ 31, 47, 54, 57, 87.
[551] Ojima I., 2004, *Temperature as order parameter of broken scale invariance*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **40**, 731. arXiv:math-ph/0311025. ↑ 2, 48.
[552] Onicescu O., 1966, *Energie informationelle*, Compt. Rend. Acad. Sci. Paris **263**, 841. ↑ 29.
[553] Ore Ø., 1935, *On the foundations of abstract algebra. I*, Ann. Math. **36**, 406. ↑ 4.
[554] Orlicz W., 1932, *Über eine gewisse Klasse von Räumen vom Typus B*, Bull. Acad. Polon. Sci. Sér. A **1932**:8/9, 207. ↑ 69.
[555] Orlicz W., 1936, *Über Räume* (L^M), Bull. Int. Acad. Polon. Sci. Lett. A **1936**, 93. ↑ 69.
[556] Padmanabhan A.R., 1979, *Probabilistic aspects of von Neumann algebras*, J. Funct. Anal **31**, 139. ↑ 87.
[557] Pardo M.C., Vajda I., 1997, *About distances of discrete distributions satisfying the data processing theorem of information theory*, IEEE Trans. Inf. Theor. **43**, 1288. ↑ 27, 47.
[558] Pearson K., 1900, *On the criterion that a given system of deviations from the probable in the case of correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling*, Phil. Mag. [Ser. 5] **50**, 157. ↑ 29.
[559] Peirce C.S., 1880, *On the algebra of logic*, Amer. J. Math. **3**, 15. www.jstor.org/stable/2369442. ↑ 4.
[560] Peirce C.S., 1885, *On the algebra of logic: a contribution to the philosophy of notation*, Amer. J. Math. **7**, 180. www.jstor.org/stable/2369451, www.jstor.org/stable/2369269. ↑ 4.
[561] Pérez A., 1967, *Information-theoretic risk estimates in statistical decision*, Kybernetika **3**, 1. www.kybernetika.cz/content/1967/1/1/paper.pdf. ↑ 32, 50.
[562] Pérez A., 1967, *Sur l'énergie informationelle de M. Octav Onicescu*, Rev. Roum. Math. Pures Appl. **12**, 1341. ↑ 29.
[563] Pettis B.J., 1939, *A proof that every uniformly convex space is reflexive*, Duke Math. J. **5**, 249. ↑ 35.
[564] Petz D., 1985, *Properties of quantum entropy*, in: Accardi L., von Waldenfels W. (eds.), *Quantum probability and applications II*, LNM **1136**, Springer, Berlin, p.428. ↑ 50.
[565] Petz D., 1985, *Quasi-entropies for states of a von Neumann algebra*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **21**, 787. www.math.bme.hu/∼petz/pdf/26quasi.pdf. ↑ 2, 31, 32, 54.
[566] Petz D., 1986, *Properties of the relative entropy of states of von Neumann algebra*, Acta Math. Hungar. **47**, 65. ↑ 50, 54.

[567] Petz D., 1986, *Quasi-entropies for finite quantum systems*, Rep. Math. Phys. **23**, 57. www.renyi.hu/∼petz/pdf/52.pdf. ↑ 31.
[568] Petz D., 1986, *Sufficient subalgebras and the relative entropy of states of a von Neumann algebra*, Commun. Math. Phys. **105**, 123. ↑ 55.
[569] Petz D., 1988, *A variational expression for the relative entropy*, Comm. Math. Phys. **114**, 345. ↑ 57.
[570] Petz D., 1991, *On certain properties of the relative entropy of states of operator algebras*, Math. Z. **206**, 351. ↑ 54, 57.
[571] Petz D., 1992, *Characterization of the relative entropy of states of matrix algebras*, Acta Math. Hung. **59**, 449. www.renyi.hu/∼petz/pdf/52.pdf. ↑ 47.
[572] Petz D., 1992, *Entropy in quantum probability I*, in: Accardi L. (ed.), *Quantum probability and related topics VII*, World Scientific, Singapore, p.275. ↑ 57.
[573] Petz D., 1994, *Geometry of canonical correlation on the state space of a quantum system*, J. Math. Phys. **35**, 780. www.renyi.hu/∼petz/pdf/61.pdf. ↑ 2, 77, 79, 84.
[574] Petz D., 1994, *On entropy functionals of states of operator algebras*, Acta Math. Hungar. **64**, 333. www.renyi.hu/∼petz/pdf/59.pdf. ↑ 47, 55.
[575] Petz D., 1996, *Monotone metrics on matrix spaces*, Lin. Alg. Appl. **224**, 81. www.renyi.hu/∼petz/pdf/68mon.pdf. ↑ 67, 77, 78.
[576] Petz D., 2007, *Bregman divergence as relative operator entropy*, Acta Math. Hungar. **116**, 127. www.renyi.hu/∼petz/pdf/112bregman.pdf. ↑ 41, 46.
[577] Petz D., 2010, *From f-divergence to quantum quasi-entropies and their use*, Entropy **12**, 304. www.mdpi.com/1099-4300/12/3/304/pdf. ↑ 31.
[578] Petz D., Hasegawa H., 1996, *On the riemannian metric of α-entropies of density matrices*, Lett. Math. Phys. **38**, 221. www.renyi.hu/∼petz/pdf/71.pdf. ↑ 81.
[579] Petz D., Sudár C., 1996, *Geometries of quantum states*, J. Math. Phys. **37**, 2662. ↑ 77.
[580] Petz D., Sudár C., 1999, *Extending the Fisher metric to density matrices*, in: Barndorff-Nielsen O.E., Vedel Jensen E.B. (eds.), *Geometry in present day science*, World Scientific, Singapore, p.21. ↑ 82.
[581] Petz D., Toth G., 1993, *The Bogoliubov inner product on quantum statistics*, Lett. Math. Phys. **27**, 205. ↑ 79.
[582] Pfanzagl J., 1973, *Asymptotic expansions related to minimum contrast estimators*, Ann. Statist. **1**, 993. ↑ 27.
[583] Pfanzagl J., 1981, *The second order optimality of tests and estimators for minimum contrast functionals. I*, Prob. Math. Statist. **2**, 55. ↑ 27.
[584] Phelps R.R., 1989, *Convex functions, monotone operators and differentiablity*, LNM **1364**, Springer, Berlin (1993, 2nd rev. ed.). ↑ 32.
[585] Picard C.E., 1890, *Memoire sur la théorie des équations aux dérivées partielles et la méthode des approximations successives*, J. Math. Pures Appl. **6**, 145. ↑ 62.
[586] Picard D.B., 1992, *Statistical morphisms and related invariance properties*, Ann. Inst. Statist. Math. **44**, 45. ↑ 75.
[587] Pinsker M.S., 1960, *Informaciya i informacionnaya ustoĭchivost' sluchaĭnykh velichin i processov*, Izdatel'stvo Akademii nauk SSSR, Moskva (Engl. transl. 1964, *Information and information stability of random variables and processes*, Holden-Day, San Francisco). ↑ 55.
[588] Pisier G., Xu Q., 2003, *Noncommutative L^p spaces*, in: *Handbook of the geometry of Banach spaces*, vol.2, North-Holland, Amsterdam, p.1459. dmitripavlov.org/scans/pisier-xu.pdf. ↑ 16.
[589] Pistone G., 2001, *New ideas in nonparametric estimation*, in: Sollich P., Hughston L.P., Streater R.F., Coolen A.C.C. (eds.), *Disordered and complex systems (King's College, London, UK, 10-14 July 2000)*, AIP Conf. Proc. **553**, AIP, Melville, p.159. ↑ 71.
[590] Pistone G., 2010, *Algebraic varietes vs. differentiable manifolds in statistical models*, in: Gibilisco P. et al (eds.), *Algebraic and geometric methods in statistics*, Cambridge University Press, Cambridge, p.341. libgen.org:4059765D692D2A0A5583B7E032CE1895. ↑ 67, 70.
[591] Pistone G., 2013, *Examples of the application of nonparametric information geometry to statistical physics*, Entropy **15**, 4042. arXiv:1308.5312. ↑ 70.
[592] Pistone G., 2013, *Nonparametric information geometry*, in: Nielsen F., Barbaresco F. (eds.), *Geometric science of information*, LNCS **8085**, p.5. arXiv:1306.0480. ↑ 70.
[593] Pistone G., Rogantin M.P., 1999, *The exponential statistical manifold, mean parameters, orthogonality and space transformations*, Bernoulli **5**, 721. ↑ 70, 84, 85.
[594] Pistone G., Sempi C., 1995, *An infinite-dimensional geometric structure on the space of all the probability measures equivalent to a given one*, Ann. Statist. **23**, 1543. ↑ 69, 70, 85.
[595] Pitcher T.S., 1957, *Sets of measure not admitting necessary and sufficient statistics on subfields*, Ann. Math. Statist. **28**, 267. ↑ 23.
[596] Pitcher T.S., 1965, *A more general property than domination for sets of probability measures*, Pacific J. Math. **15**, 597. ↑ 23.
[597] Pitman E.J.G., 1936, *Sufficient statistics and intrinsic accuracy*, Proc. Cambridge Phil. Soc. **32**, 567. ↑ 85.
[598] Pusz W., Woronowicz S.L., 1978, *Form convex functions and the WYDL and other inequalities*, Lett. Math. Phys. **2**, 505. ↑ 27, 50, 54, 85.
[599] Radon J., 1913, *Theorie und Anwendungen der absolut additiven Mengenfunktionen*, Sitzungsber. Akad. Wiss. Wien, Math.-Naturwiss. Klasse, Abt. IIa **122**, 1296. ↑ 9, 10.
[600] Raggio G.A., Werner R.F., 1990, *Minimizing the relative entropy in a face*, Lett. Math. Phys. **19**, 7. ↑ 57.
[601] Rao C.R., 1945, *Information and accuracy attainable in the estimation of statistical parameters*, Bull. Calcutta Math. Soc. **37**, 81. ↑ 27, 67, 74, 87.
[602] Rao C.R., 1947, *The problem of classification and distance between two populations*, Nature **159**, 30. ↑ 74.
[603] Rao C.R., 1982, *Diversity and dissimilarity coefficients: a unified approach*, J. Theor. Popul. Biol. **21**, 24. ↑ 27.
[604] Rao M.M., Ren Z.D., 1991, *Theory of Orlicz spaces*, Dekker, New York. libgen.org:377A7D40BDED488A7101383A69A5497F. ↑ 3, 69.
[605] Rao M.M., Ren Z.D., 2002, *Applications of Orlicz spaces*, Dekker, New York. libgen.org:BFADBC4E1FDEACC31EAB6D3E1A72F4C5. ↑ 69.
[606] Read T.R.C., Cressie N., 1988, *Goodness-of-fit statistics for discrete multivariate data*, Springer, Berlin. ↑ 50.
[607] Reid M., Williamson R., 2009, *Generalised Pinsker inequalities*, in: Dasgupta S., Klivans A. (eds.), *Annual conference on computational learning theory (COLT 2009)*, Asociation for Computing Machinery, Montreal, p.125. arXiv:0906.1244. ↑ 55.
[608] Rellich F., 1937, *Störungstheorie der Spektralzerlegung I, II*, Math. Ann. **113**, 600, 677. ↑ 72.

[609] Rellich F., 1939, *Störungstheorie der Spektralzerlegung III*, Math. Ann. **116**, 555. ↑ 72.
[610] Rellich F., 1940, *Störungstheorie der Spektralzerlegung IV*, Math. Ann. **117**, 356. ↑ 72.
[611] Rellich F., 1942, *Störungstheorie der Spektralzerlegung V*, Math. Ann. **118**, 462. ↑ 72.
[612] Rényi A., 1960, *Az információelmélet néhány alapvető kérdése*, MTA III. Oszt. Közl. **10**, 251 (Engl. transl. 1976, *Some fundamental questions of information theory*, in: Rényi A., *Selected papers of Alfred Rényi*, Vol.2, Akadémiai Kiadó, Budapest, p.526). ↑ 50.
[613] Rényi A., 1961, *On measures of entropy and information*, in: Neyman J. (ed.), *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*, Vol.1, University of California Press, Berkeley, p.547. euclid:bsmsp/1200512181. ↑ 50.
[614] Rényi A., 1962, *Wahrscheinlichkeitsrechnung: mit einem Anhang über Informationstheorie*, Deutscher Verlag der Wissenschaften, Berlin (Engl. transl. 1970, *Probability theory*, North-Holland, Amsterdam). ↑ 27.
[615] Resmerita E., 2004, *On total convexity, Bregman projections and stability in Banach spaces*, J. Conv. Anal. **11**, 1. ↑ 38.
[616] Ricci-Curbastro G., 1887, *Sulla derivazione covariante ad una forma quadratica differenziale*, Rend. Acc. Lincei **3**, 15. ↑ 61.
[617] Ricci-Curbastro G., 1888, *Delle derivazioni covarianti e controvarianti e del loro uso nell'analisi applicata*, Studi Ed. Univ. Padova **3**, 3. ↑ 61.
[618] Ricci-Curbastro G., Levi-Civita T., 1901, *Méthodes de calcul différentiel absolu et leurs applications*, Math. Ann. **54**, 125. houchmandzadeh.net/cours/Relativite/Biblio/ricci_levicivita_1900.pdf. ↑ 61.
[619] Riemann B., 1854, *Über die Hypothesen, welche der Geometrie zu Grunde liegen*, Habilitationsschrift, Göttingen, publ. in: 1876, Abhandl. Königl. Gess. Wiss. Göttingen **13**, 87. www.emis.de/classics/Riemann/Geom.pdf (Engl. transl. 1873, *On the hypotheses which lie at the bases of geometry*, Nature **8**, 183:14, 184:36). ↑ 60.
[620] Riemann B., 1861, *Commentatio mathematica, qua respondere tentatur quaestioni ab Ill^{ma} Academia Parisiensi propositiae:*, in: 1876, *Gesammelte mathematische Werke und wissenschaftlicher Nachlass*, Teubner, Leipzig, p.391. www.maths.tcd.ie/pub/HistMath/People/Riemann/Paris/Paris.pdf (Engl. transl. 1990, *A mathematical treatise in which an attempt is made to answer the question prposed by the most illustrious Academy of Paris:*, Hist. Math. **17**, 240). ↑ 61.
[621] Riesz F., 1910, *Untesuchungen über Systeme integrierbarer Funktionen*, Math. Ann. **69**, 449. gdz.sub.uni-goettingen.de/dms/load/img/?PPN=PPN235181684_0069&DMDID=DMDLOG_0040. ↑ 9.
[622] Riesz F., 1917, *Lineáris függvényegyenletekről*, Matematikai és Természettudományi Értesítő **35**, 544 (Germ. transl. 1918, *Über lineare Funktionalgleichungen*, Acta Math. **41**, 71; Engl. transl. www.math.technion.ac.il/hat/fpapers/rie1ineng.pdf). ↑ 17.
[623] Riesz F., 1930, *Sur la décomposition des opérations fonctionelles linéaires*, in: Zanichelli N. (ed.), *Atti Congr. Intern. Mat. Bologna 3-10 Settembre 1928 (VI)*, Vol.3, Società Tipografica già Compositori, Bologna, p.143. www.mathunion.org/ICM/ICM1928.3/Main/icm1928.3.0143.0148.ocr.pdf. ↑ 4.
[624] Riesz F., 1937, *A lineáris operációk általános elméletének néhány alapvető fogalomalkotásáról*, Matematikai és Természettudományi Értesítő **56**, 1. ↑ 4.
[625] Riesz F., 1940, *Sur quelques notions fondamentales dans la théorie générale des opérations linéaires*, Ann. Math. **41**, 174. ↑ 4.
[626] Robinson D.W., 1968, *Statistical mechanics of quantum spin systems II*, Commun. Math. Phys. **7**, 337. ↑ 87.
[627] Rockafellar R.T., 1968, *Integrals which are convex functionals*, Pacific J. Math. **24**, 525. ↑ 32.
[628] Rockafellar R.T., 1970, *Conjugate convex functions in optimal control and the calculus of variations*, J. Math. Anal. Appl. **32**, 174. ↑ 32.
[629] Rockafellar R.T., 1970, *Convex analysis*, Princeton University Press, Princeton. www.math.washington.edu/∼rtr/papers/rtr-ConvexAnalysis.djvu. ↑ 32, 33, 35.
[630] Rockafellar R.T., 1971, *Convex integral functionals and duality*, in: Zarantonello E. (ed.), *Contribution to nonlinear functional analysis*, Academic Press, New York, p.215. ↑ 32.
[631] Rockafellar R.T., 1971, *Integrals which are convex functionals II*, Pacific J. Math. **39**, 439. ↑ 32.
[632] Rockafellar R.T., 1974, *Conjugate duality and optimization*, Regional Conferences Series in Applied Mathematics **16**, Society for Industrial and Applied Mathematics, Philadelphia. ↑ 32.
[633] Rockafellar R.T., Wets R.J.-B., 1998, *Variational analysis*, Springer, Berlin (2006, 2nd ed.). ↑ 32.
[634] Rodríguez C.C., 1991, *Entropic priors*, in: Grandy W.T. Jr., Schick L.H. (eds.), *Maximum entropy and bayesian methods*, Kluwer, Dordrecht. ↑ 89.
[635] Rodríguez C.C., 2003, *A geometric theory of ignorance*, omega.albany.edu:8008/ignorance/ignorance03.pdf. ↑ 2.
[636] Romier G., 1969, *Décision statistique*, Ann. Inst. Henri Poincaré B **5**, ?. ↑ 24.
[637] Romier G., 1969, *Modél d'expérimentation statistique*, Ann. Inst. Henri Poincaré B **5**, 275. ↑ 24.
[638] Rota G.-C., 2001, *Twelve problems in probability no one likes to bring up*, in: Crapo H., Senato D. (eds.), *Algebraic combinatorics and computer science: a tribute to Gian-Carlo Rota*, Springer Italia, Milano, p.57. ↑ 23.
[639] Rüschendorf L., 1984, *On the minimum discrimination information theorem*, Statist. Dec. Suppl. **1**, 263. ↑ 30, 32.
[640] Rüschendorf L., 1987, *Projections of probability measures*, Statistics **18**, 123. ↑ 32.
[641] Ruskai M.B., 1973, *A generalisation of entropy using traces on von Neumann algebras*, Ann. Inst. Henri Poincaré A **19**, 357. ↑ 87.
[642] Ruskai M.B., 1988, *Extremal properties of relative entropy in quantum statistical mechanics*, Rep. Math. Phys. **26**, 143. ↑ 57, 87.
[643] Sacksteder R., 1967, *A note on statistical equivalence*, Ann. Math. Statist. **38**, 787. ↑ 24.
[644] Sacksteder R., 1968, *On products of experiments*, Z. Wahrschein. Gebiete **10**, 203. ↑ 24.
[645] Sadeghi G., 2012, *Non-commutative Orlicz spaces associated to a modular on τ-measurable operators*, J. Math. Anal. App. **395**, 705. libgen.org:scimag/get.php?doi=10.1016/j.jmaa.2012.05.054. ↑ 73.
[646] Saks S., 1933, *Théorie de l'intégrale*, Monografje Matematyczne **2**, Garasiński–Fundusz Kultury Narodowej, Warszawa. libgen.org:3B45744D0B01168F9CEF5EDD2D0477E8 (2nd rev. ed., 1937, *Theory of the integral*, Drukarnia Uniwersytetu Jagiellońskiego/Stechert, Warszawa–Lwów/New York. libgen.org:3417123462FD2580DCDB7B5C8D8B8F3E). ↑ 29.
[647] Sato Y., Sugawa K., Kawaguchi M., 1979, *The geometrical structure of the parametric space of two-dimensional normal distribution*, Rep. Math. Phys. **16**, 111. ↑ 74, 89.
[648] Schatten R., 1946, *The cross-space of linear transformations*, Ann. Math. **47**, 73. ↑ 18.
[649] Schatten R., 1950, *A theory of cross spaces*, Princeton University Press, Princeton. libgen.org:566b60cc48d6cfb9a42c11d0a3521fc0. ↑ 18.
[650] Schatten R., 1960, *Norm ideals of completely continuous operators*, Springer, Berlin. libgen.org:8BA5A444148F4B3722719DB398C9DA85. ↑ 18.

[651] Schatten R., von Neumann J., 1946, *The cross-space of linear transformations II*, Ann. Math. **47**, 608. ↑ 17, 18.
[652] Schatten R., von Neumann J., 1947, *The cross-space of linear transformations III*, Ann. Math. **49**, 557. ↑ 17, 18.
[653] Schauder J., 1927, *Zur Theorie stetiger Abbildungen in Fuktionalräumen*, Math. Z. **26**, 47. ↑ 59.
[654] Schauder J., 1930, *Über lineare, vollstetige Funktionaloperationen*, Studia Math. **2**, 183. matwbn.icm.edu.pl/ksiazki/sm/sm2/sm2116.pdf. ↑ 17.
[655] Schmidt E., 1907, *Zur Theorie der linearen und nichtlinearen Integralgleichungen I. Teil: Entwicklung willkürlicher Funktionen nach Systemen vorgeschriebener*, Math. Ann. **63**, 433. gdz.sub.uni-goettingen.de/dms/load/img/?PPN=GDZPPN002261464&IDDOC=36690. ↑ 18.
[656] Schröder E., 1890, 1891, 1895, *Vorlesungen über die Algebra der Logik*, Vol.1-3, Teubner, Leipzig. libgen.org:8d22045696a5d71bf1e9eaa066c5a262, libgen.org:0f558bcb0f53dd6c6850329a31023b43, libgen.org:1dc5e441edc3fa8abd7ccce43d1e2c9f. ↑ 4.
[657] Segal I.E., 1947, *Irreducible representations of operator algebras*, Bull. Amer. Math. Soc. **61**, 69. www.ams.org/journals/bull/1947-53-02/S0002-9904-1947-08742-5/S0002-9904-1947-08742-5.pdf. ↑ 14.
[658] Segal I.E., 1951, *Equivalences of measure spaces*, Amer. J. Math. **73**, 275. ↑ 7, 8, 9, 10, 22.
[659] Segal I.E., 1953, *A non-commutative extension of abstract integration*, Ann. Math. **57**, 401 (corr. Ann. Math. **58**, 595). ↑ 17.
[660] Segal I.E., 1954, *Abstract probability spaces and a theorem of Kolmogoroff*, Am. J. Math. **76**, 721. ↑ 23.
[661] Segal I.E., 1960, *A note on the concept of entropy*, J. Math. Mech. **9**, 623. www.iumj.indiana.edu/IUMJ/FTDLOAD/1960/9/59036/pdf. ↑ 87.
[662] Segal I.E., 1965, *Algebraic integration theory*, Bull. Amer. Math. Soc. **71**, 571. euclid:bams/1183526903. ↑ 23.
[663] Semadeni Z., 1971, *Banach spaces of continuous functions*, Monografie Matematyczne **55**, Państwowe Wydawnictwo Naukowe, Warszawa. ↑ 23.
[664] Sen H., 1959, *On relations of curvature tensors over Sen's system of affine connections*, Rend. Sem. Mat. Univ. Padova **29**, 242. numdam:RSMUP_1959__29__242_0. ↑ 62.
[665] Sen R.N., 1944, *On parallelism in riemannian space*, Bull. Calcutta Math. Soc. **36**, 102. ↑ 62.
[666] Sen R.N., 1945, *On parallelism in riemannian space-II*, Bull. Calcutta Math. Soc. **37**, 153. ↑ 62.
[667] Sen R.N., 1946, *On parallelism in riemannian space III*, Bull. Calcutta Math. Soc. **38**, 161. ↑ 62.
[668] Sen R.N., 1948, *Parallel displacement and scalar product of vectors*, Proc. Nat. Inst. Sci. India **14**, 45. www.new1.dli.ernet.in/data1/upload/insa/INSA_1/20005b8e_45.pdf. ↑ 62.
[669] Sen R.N., 1949, *Parallel displacement and scalar product of vectors-II, -III*, Bull. Calcutta Math. Soc. **41**, 41, 113 (corr. **42**, 56). ↑ 62.
[670] Sen R.N., 1950, *On an algebraic system generated by a single-element and its application in riemannian geometry*, Bull. Calcutta Math. Soc. **42**, 1, 177. ↑ 62.
[671] Shafer G., Vovk V., 2005, *The origins and legacy of Kolmogorov's Grundbegriffe*, preprint. www.probabilityandfinance.com/articles/04.pdf. ↑ 23.
[672] Shannon C., 1948, *A mathematical theory of communication*, Bell Syst. Tech. J. **27**, 379, 623. ↑ 86.
[673] Shannon C., Weaver W., 1949, *The mathematical theory of communication*, University of Illinois Press, Urbana. ↑ 86.
[674] Sherman D.E., 2001, *The application of modular algebras to relative tensor products and noncommutative L^p modules*, Ph.D. thesis, University of California, Los Angeles. ↑ 20.
[675] Shima H., 1976, *On certain locally flat homogeneus manifolds of solvable Lie groups*, Osaka J. Math. **13**, 213. ↑ 64.
[676] Shima H., 1977, *Symmetric spaces with invariant locally hessian structures*, J. Math. Soc. Japan **29**, 581. ↑ 64.
[677] Shima H., 1986, *Vanishing theorems for compact hessian manifolds*, Ann. Inst. Fourier Grenoble **36**, 183. ↑ 65.
[678] Shima H., 2007, *The geometry of hessian structures*, World Scientific, Singapore. ↑ 65.
[679] Shima H., Yagi K., 1997, *Geometry of hessian manifolds*, Diff. Geom. Appl. **7**, 277. ↑ 64, 65.
[680] Shmul'yan V.L., 1939, *O nekotorykh geometricheskikh svoĭstvakh edinichnoĭ sfery prostranstva tipa* (B), Matem. Sb. N.S. **6**, 77, 90. ↑ 35.
[681] Shmul'yan V.L., 1940, *Sur les topologies différentes dans l'espace de Banach*, Dokl. Akad. Nauk SSSR **23**, 331. ↑ 34, 35.
[682] Siebert E., 1979, *Pairwise sufficiency*, Z. Warschein. Geb. **46**, 237. ↑ 23.
[683] Sikorski R., 1948, *On the representation of boolean algebras as fields of sets*, Fund. Math. **35**, 247. matwbn.icm.edu.pl/ksiazki/fm/fm35/fm35123.pdf. ↑ 9.
[684] Simons S., 1998, *Minimax and monotonicity*, LNM **1693**, Springer, Berlin. ↑ 35.
[685] Skovgaard L.T., 1984, *A riemannian geometry of the multivariate normal model*, Scand. J. Stat. **11**, 211. ↑ 89.
[686] Stacey A., 2007, *Variations on a theme: riemannian geometry in infinite dimensions*, www.math.mtnu.no/∼stacey/documents/geometry.article.pdf. ↑ 60.
[687] Stacey A., 2008, *The co-riemannian structure of smooth loop spaces*, arXiv:0809.3108. ↑ 60.
[688] Steinhaus H., 1919, *Additive und stetige Funktionaloperationen*, Math. Z. **5**, 186. gdz.sub.uni-goettingen.de/dms/load/img/?PPN=GDZPPN002365073&IDDOC=16795. ↑ 9, 22.
[689] Steinhaus H., 1923, *Les probabilités denomerables et leur rapport à théorie de la mesure*, Fund. Math. **4**, 286. ↑ 22.
[690] Stinespring W.F., 1955, *Positive functions on C^*-algebras*, Proc. Amer. Math. Soc. **6**, 211. ↑ 26.
[691] Stirzaker D., 1994, *Probability vicit expectation*, in: Kelly F.P. (ed.), *Probability, statistics and optimization. A tribute to Peter Whittle*, Wiley, Chisterter. ↑ 23.
[692] Stone M.H., 1932, *On one-parameter unitary groups in Hilbert space*, Ann. Math. **33**, 643. ↑ 18.
[693] Stone M.H., 1936, *The theory of representations of boolean algebras*, Trans. Amer. Math. Soc. **40**, 37. www.ams.org/journals/tran/1936-040-01/S0002-9947-1936-1501865-8/S0002-9947-1936-1501865-8.pdf. ↑ 5, 9.
[694] Størmer E., 1963, *Positive maps of operator algebras*, Acta Math. **110**, 233. ↑ 26.
[695] Strasser H., 1985, *Mathematical theory of statistics: statistical experiments and asymptotic decision theory*, de Gruyter, Berlin. ↑ 23.
[696] Streater R.F., 2000, *The analytic quantum information manifold*, in: Gesztesy F., Paycha S., Holden H. (eds.), *Stochastic processes, physics and geometry: new interplays. II: A volume in honor of Sergio Albeverio (Leipzig, 1999)*, Canadian Math. Soc. Conf. Proc. **29**, American Mathematical Society, Providence, p.603. arXiv:math-ph/9910036. ↑ 72.
[697] Streater R.F., 2000, *Classical and quantum info-manifolds*, Sūrikaisekikenkyūsho Kōkyūroku **1196**, 32. arXiv:math-ph/0002050. ↑ 72.
[698] Streater R.F., 2000, *The information manifold for relatively bounded potentials*, in: Vladimirov V.S., Mishhenko E.F. (eds.), *Problems of the modern mathematical physics. Collection of papers dedicated to the 90th anniversary of academician Nikolai Nikolaevich Bogolyubov*, Tr. Mat. Inst. V.A. Steklova **228**, Nauka, Moskva, p.217.

arXiv:math-ph/9910035. ↑ 72, 73.
[699] Streater R.F., 2004, *Duality in quantum information geometry*, Open Sys. Inf. Dyn. **11**, 71. arXiv:math-ph/0308037. ↑ 72, 85.
[700] Streater R.F., 2004, *Quantum Orlicz spaces in information geometry*, Open Sys. Inf. Dyn. **11**, 350. arXiv:math-ph/0407046. ↑ 72, 73, 87.
[701] Streater R.F., 2008, *The set of density operators modelled on a quantum Orlicz space*, in: Belavkin V.P., Guţă M. (eds.), *Quantum stochastics and information: statistics, filtering and control*, World Scientific, Singapore, p.99. libgen.org:06BA05CB2B143C2CB333D764D308BF80. ↑ 72.
[702] Streater R.F., 2009, *Statistical dynamics. A stochastic approach to nonequilibrium thermodynamics*, 2nd rev. ed., Imperial College Press, London. ↑ 67, 72, 87.
[703] Streater R.F., 2010, *The Banach manifold of quantum states*, in: Gibilisco P. et al (eds.), *Algebraic and geometric methods in statistics*, Cambridge University Press, Cambridge, p.257. libgen.org:4059765D692D2A0A5583B7E032CE1895. ↑ 72, 87.
[704] Streater R.F., 2010, *Introduction to non-parametric estimation*, in: Gibilisco P. et al (eds.), *Algebraic and geometric methods in statistics*, Cambridge University Press, Cambridge, p.241. libgen.org:4059765D692D2A0A5583B7E032CE1895. ↑ 67, 72.
[705] Study E., 1905, *Kürzeste Wege in komplexen Gebiet*, Math. Ann. **60**, 321. ↑ 82.
[706] Stummer W., Vajda I., 2010, *On divergences of finite measures and their applicability in statistics and information theory*, Statistics **44**, 169. ↑ 30.
[707] Sudár C., 1996, *Radial extension of monotone riemannian metrics on density matrices*, Publ. Math. Debrecen **49**, 243. ↑ 82.
[708] Takesaki M., 1979, 2003, *Theory of operator algebras*, Vol.1-3, Springer, Berlin. libgen.org:7F0A9F06741272684D62426E348670B1, libgen.org:611936516F867B45282BED443AA0EE51, libgen.org:5052F03EF47C3B9EF73299E4EF29D022. ↑ 16.
[709] Takiyama R., 1974, *[On geometrical structures of parameter spaces of one-dimensional distributions]*, Trans. Inst. Elect. Comm. Eng. Japan A **57**, 67. ↑ 64.
[710] Taneja I., Kumar P., 2004, *Relative information of type s, Csiszár f-divergence, and information inequalities*, Inf. Sci. **166**, 131. ↑ 30.
[711] Teboulle M., 1992, *Entropic proximal mappings with applications to nonlinear programming*, Math. Oper. Res. **17**, 670. ↑ 38.
[712] Teboulle M., Vajda I., 1993, *Convergence of best ϕ-entropy estimates*, IEEE Trans. Inf. Theor. **39**, 297. ↑ 32.
[713] Terp M., 1981, *L^p-spaces associated with von Neumann algebras*, Københavns Univ. Math. Inst. Rapp. No. **3a+3b**, Matematisk Institut, Københavns Universitet, København. www.fuw.edu.pl/∼kostecki/scans/terp1981.pdf. ↑ 20.
[714] Terp M., 1982, *Interpolation spaces between von Neumann algebra and its predual*, J. Oper. Theor. **8**, 327. www.theta.ro/jot/archive/1982-008-002/1982-008-002-008.pdf. ↑ 20.
[715] Terras A., 1979, *Harmonic analysis on symmetric spaces and applications*, Springer, Berlin. ↑ 89.
[716] Tomamichel M., Colbeck R., Renner R., 2009, *A fully quantum asymptotic equipartition property*, IEEE Trans. Inf. Theor. **55**, 5840. arXiv:0811.1221. ↑ 31.
[717] Tomiyama J., 1959, *On the projection of norm one in W^*-algebras, III*, Tōhoku Math. J. **11**, 125. euclid:tmj/1178244633. ↑ 26.
[718] Topsøe F., 2001, *Bounds for entropy and divergence for distributions over a two element set*, J. Ineq. Pure Appl. Math. **2**, 25. ↑ 55.
[719] Torgersen E.N., 1991, *Comparison of statistical experiments*, Cambridge University Press, Cambridge. ↑ 23.
[720] Toussaint G.T., 1975, *Sharper lower bounds for discrimination information in terms of variation*, IEEE Trans. Inf. Theor. **21**, 99. ↑ 55.
[721] Tropp J.A., 2011, *From joint convexity of quantum relative entropy to a concavity theorem of Lieb*, arXiv:1101.1070. ↑ 54.
[722] Tsallis C., 1988, *Possible generalization of Boltzmann–Gibbs statistics*, J. Stat. Phys. **52**, 479. ↑ 50.
[723] Uffink J.B.M., 1995, *Can the maximum entropy principle be explained as a consistency requirement?*, Stud. Hist. Phil. Mod. Phys. B **26**, 223. ↑ 86.
[724] Uhl J.J., 1966, *Orlicz spaces of additive set functions and set-martingales*, Ph.D. thesis, Carnegie–Mellon University, Pittsburgh. ↑ 29.
[725] Uhl J.J., 1967, *Orlicz spaces of finitely additive set functions*, Studia Math. **29**, 19. ↑ 29.
[726] Uhlmann A., 1977, *Relative entropy and the Wigner–Yanase–Dyson–Lieb concavity in an interpolation theory*, Commun. Math. Phys. **54**, 21. www.physik.uni-leipzig.de/∼uhlmann/PDF/Uh77a.pdf. ↑ 27, 31, 50, 54.
[727] Uhlmann A., 1986, *Parallel transport and "quantum holonomy" along density operators*, Rep. Math. Phys. **24**, 229. www.physik.uni-leipzig.de/∼uhlmann/PDF/Uh86a.pdf. ↑ 67.
[728] Uhlmann A., 1987, *Parallel transport and holonomy along density operators*, in: Doebner H.D., Hennig J.D. (eds), *Differential geometric methods in theoretical physics*, World Scientific, Singapore, p.246. www.physik.uni-leipzig.de/∼uhlmann/PDF/Uh87a.pdf. ↑ 67.
[729] Uhlmann A., 1992, *An energy dispersion estimate*, Phys. Lett A **161**, 329. www.physik.uni-leipzig.de/∼uhlmann/PDF/Uh92a.pdf. ↑ 81.
[730] Uhlmann A., 1992, *The metric of Bures and the geometric phase*, in: Gielerak R., Lukierski J., Popowicz Z. (eds.), *Quantum groups and related topics, Proceedings of the First Max Born Symposium, Wrocław 1991*, Kluwer, Dordrecht, p.267. www.physik.uni-leipzig.de/∼uhlmann/PDF/Uh92b.pdf. ↑ 81.
[731] Uhlmann A., 1993, *Density operators as an arena for differential geometry*, Rep. Math. Phys. **33**, 253. www.physik.uni-leipzig.de/∼uhlmann/PDF/Uh93e.pdf. ↑ 81.
[732] Uhlmann A., 1995, *Geometric phases and related structures*, Rep. Math. Phys. **36**, 461. www.physik.uni-leipzig.de/∼uhlmann/PDF/Uh95a.pdf. ↑ 67.
[733] Uhlmann A., Crell B., 2009, *Geometry of state spaces*, in: Buchleitner A. et al. (eds.), *Entanglement and decoherence*, Lecture Notes in Physics **768**, Springer, Berlin, p.1. www.physik.uni-leipzig.de/∼uhlmann/PDF/UC07.pdf. ↑ 67.
[734] Ulam S., 1930, *Zur Masstheorie in der allgemeinen Mengenlehre*, Fund. Math. **16**, 140. matwbn.icm.edu.pl/ksiazki/fm/fm16/fm16114.pdf. ↑ 8.
[735] Umegaki H., 1961, *On information in operator algebras*, Proc. Jap. Acad. **37**, 459. euclid:pja/1195523632. ↑ 27, 50.
[736] Umegaki H., 1962, *Conditional expectation in an operator algebra, IV (entropy and information)*, Kōdai Math. Sem. Rep. **14**, 59. euclid:kmj/1138844604. ↑ 27, 50.
[737] Vajda I., 1970, *Note on discrimination information and variation*, IEEE Trans. Inf. Theor. **16**, 99. ↑ 55.

[738] Vajda I., 1972, *On the f-divergence and singularity of probability measures*, Period. Math. Hung. **2**, 223. ↑ 30.
[739] Vajda I., 1973, *χ^α-divergence and generalized Fisher's information*, in: Koženšnik J. (ed.), *Transactions of the sixth Prague conference on information theory, statistical decision functions and random processes*, Academia, Praha, p.873. ↑ 29.
[740] van der Vaart A., 1998, *Asymptotic statistics*, Cambridge University Press, Cambridge. ↑ 23.
[741] van der Vaart A., 2002, *The statistical work of Lucien Le Cam*, Ann. Statist. **30**, 631. ↑ 23.
[742] van Neerven J.M.A.M., 1997, *The norm of a complex Banach lattice*, Positivity **1**, 381. aw.twi.tudelft.nl/∼neerven/publications/papers/Positivity_98.pdf. ↑ 21.
[743] Vedernikov V.I., 1965, *Simmetricheskie prostranstva i sopryazhennye svyaznosti*, Uchen. zap. kazansk. univ. **125**, 7. ↑ 62.
[744] Vinberg È.B., 1963, *Teoriya odnorodnykh vypuklykh konusov*, Trudy mosk. mat. obshh **12**, 303 (Engl. transl. 1963, *The theory of convex homogeneous cones*, Trans. Moscow Math. Soc. **12**, 340). ↑ 65.
[745] Volle M., Hiriart-Urruty J.-B., 2012, *A characterization of essentially strictly convex functions on reflexive Banach spaces*, Nonlin. Anal. **75**, 1617. ↑ 33, 35, 39.
[746] Volčič A., 1970, *Sul teorema di Radon–Nikodym nel caso non σ-finito*, Rend. Inst. mat. Univ. Trieste **2**, 42. rendiconti.dmi.units.it/volumi/02/04.pdf. ↑ 10.
[747] von Neumann J., 1927, *Mathematische Begründung der Quantenmechanik*, Nachr. Ges. Wiss. Göttingen, Math.-phys. Kl. **1927**, 1. gdz.sub.uni-goettingen.de/dms/load/img/?PPN=GDZPPN002507129&IDDOC=64005. ↑ 18.
[748] von Neumann J., 1927, *Thermodynamik quantenmechanischer Gesamtheiten*, Nachr. Ges. Wiss. Göttingen, Math.-phys. Kl. **1927**, 273. gdz.sub.uni-goettingen.de/dms/load/img/?PPN=GDZPPN002507285&IDDOC=64021. ↑ 24.
[749] von Neumann J., 1930, *Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren*, Math. Ann. **102**, 49. gdz.sub.uni-goettingen.de/index.php?id=11&PPN=PPN235181684_0102&DMDID=DMDLOG_0007&L=0. ↑ 3.
[750] von Neumann J., 1930, *Zur algebra der Funktionaloperatoren und Theorie der normalen Operatoren*, Math. Ann. **102**, 370. gdz.sub.uni-goettingen.de/dms/load/img/?PPN=GDZPPN002273675&IDDOC=38466. ↑ 14.
[751] von Neumann J., 1937, *Some matrix-inequalities and metrization of matrix-spaces*, Tomsk Univ. Rev. **1**, 286. ↑ 18.
[752] Vulikh B.Z., 1957, *Chastichnoe uporyadochenie kolec ogranichennykh samosopryazhennykh operatorov*, Vestnik Leningr. Univ., Mat., Mekh., Astr. **13**, 13. www.fuw.edu.pl/∼kostecki/scans/vulikh1957.pdf. ↑ 21.
[753] Vulikh B.Z., 1961, *Vvedenie v teoriyu poluuporyadochennykh prostranstv*, Gosudarstvennoe Izdatel'stvo Fiziko-Matematicheskoĭ Literatury, Moskva. libgen.org:DC2470A89879A0C69CC4E1C255E12404 (Engl. transl., 1967, *Introduction to the theory of partially ordered spaces*, Woolters–Noordhoff, Groningen. libgen.org:8a75e6848531fa687eb1115bd9b6e2bf). ↑ 21.
[754] Wald A., 1939, *Contributions to the theory of statistical estimation and testing hypothesis*, Math. Statist. **10**, 299. ↑ 24.
[755] Wald A., 1947, *Sequential analysis*, Wiley, New York. ↑ 47.
[756] Wald A., 1950, *Statistical decision functions*, Wiley, New York. ↑ 24.
[757] Wecken F., 1940, *Abstrakte Integrale und fastperiodische Funktionen*, Math. Z. **45**, 377. gdz.sub.uni-goettingen.de/dms/load/img/?PPN=GDZPPN002378647&IDDOC=19755. ↑ 8.
[758] Wehrl A., 1978, *General properties of entropy*, Rev. Mod. Phys. **50**, 221. ↑ 54, 85, 86, 87.
[759] Weil A., 1940, *Calcul des probabilités, méthode axiomatique, intégration*, Revue Sci. (Rev. Rose Illus.) **78**, 201. ↑ 23.
[760] Weyl H.K.H., 1918, *Raum–Zeit–Materie*, Springer, Berlin (5th rev. ed. 1923; Engl. transl. 1952, *Space, time, matter*, Dover, NY). ↑ 60.
[761] Weyl H.K.H., 1918, *Reine Infinitesimalgeometrie*, Math. Z. **2**, 384. ↑ 60.
[762] Whitehead A.N., 1898, *A treatise on universal algebra, with applications*, Cambridge University Press, Cambridge. libgen.org:8c87bdb9f9f0a244dbdca42279b7cae1. ↑ 4.
[763] Whittle P., 1970, *Probability*, Penguin, Harmondsworth. ↑ 23.
[764] Whittle P., 2000, *Probability via expectation*, (4th ed.), Springer, Berlin. ↑ 23.
[765] Wichmann E.H., 1963, *Density matrices arising from incomplete measurements*, J. Math. Phys. **4**, 884. ↑ 86.
[766] Widder D.V., 1946, *The Laplace transform*, Princeton University Press, Princeton. ↑ 70.
[767] Wiener N., 1920, *The mean of a functional of arbitrary elements*, Ann. Math. **22**, 66. ↑ 23.
[768] Wiener N., 1923, *Differential space*, J. Math. Phys. **2**, 131. ↑ 23.
[769] Wiener N., 1948, *Cybernetics or control and communication in the animal and the machine*, MIT Press, Cambridge. libgen.org:f6a1f1fbe59fd6d26e21cdb6b681452f. ↑ 27.
[770] Wigner E.P., Yanase M.M., 1963, *Information contents of distributions*, Proc. Nat. Acad. Sci. U.S.A. **49**, 910. ↑ 80, 81.
[771] Woronowicz S.L., 1979, *Operator systems and their application to the Tomita–Takesaki theory*, LNS **52**, Aarhus Universitet, Matematisk Institut, Aarhus. Reprinted in: J. Op. Theor. **2**, 169. www.theta.ro/jot/archive/1979-002-002/1979-002-002-003.pdf. ↑ 20.
[772] Yamada S., Morimoto H., 1992, *Sufficiency*, in: Ghosh M., Pathak P.K. (eds.), *Current issues in statistical inference: essays in honor of D. Basu*, Institute of Mathematical Statistics, Hayward, p.86. euclid:lnms/1215458840. ↑ 23.
[773] Yamagami S., 1992, *Algebraic aspects in modular theory*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **28**, 1075. dx.doi.org/10.2977/prims/1195167738. ↑ 19, 20.
[774] Yamagami S., 1994, *Modular theory for bimodules*, J. Funct. Anal. **125**, 327. ↑ 20.
[775] Yeadon F.J., 1975, *Non-commutative L_p-spaces*, Math. Proc. Cambridge Phil. Soc. **77**, 91. ↑ 17.
[776] Yoshizawa S., Tanabe K., 1999, *Dual differential geometry associated with the Kullback–Leibler information on the gaussian distributions and its 2-parameter deformations*, SUT J. Math. **1**, 113. ↑ 89.
[777] Yoshizawa T., 1971, *A geometrical interpretation of location and scale parameters*, Memorandom TYH-**3**, Statistics Department, Harvard University, Cambridge. ↑ 64, 88.
[778] Yoshizawa T., 1971, *A geometry of parameter space and its statistical interpretation*, Memorandom TYH-**2**, Statistics Department, Harvard University, Cambridge. ↑ 88.
[779] Young W.H., 1912, *On classes of summable functions and their Fourier series*, Proc. Royal Soc. London Ser. A. **87**, 225. libgen.org:scimag/get.php?doi=10.1098/rspa.1912.0076. ↑ 33.
[780] Yuen H.P.H., 1971, *Communication theory of quantum systems*, Technical report **482**, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge. ↑ 80.
[781] Yuen H.P.H., Lax M., 1973, *Multiple-parameter quantum estimation and measurement of nonselfadjoint observables*, IEEE Trans. Inf. Theor. **19**, 740. ↑ 80.
[782] Zaanen A.C., 1952, *Integral transformations and their resolvents in Orlicz and Lebesgue spaces*, Compositio Math. **10**, 56. ↑ 69.
[783] Zaanen A.C., 1961, *The Radon–Nikodym theorem I, II*, Nederl. Akad. Wetensch. Proc. Ser. A **64**, 157. ↑ 10.
[784] Zhang J., 2004, *Divergence function, duality, and convex analysis*, Neural Comp. **16**, 159.

www.lsa.umich.edu/psych/junz/Neural Comp 2004.pdf. ↑ 40, 54, 64, 67.

[785] Zhang J., 2004, *Dual scaling of comparision and reference stimuli in multi-dimensional psychological space*, J. Math. Psych. **48**, 409. ↑ 44.

[786] Zhang J., 2006, *A note on curvature of α-connections of a statistical manifold*, Inst. Statist. Math. DOI:**10.1007/s10463-006-0105-1**. ↑ 64.

[787] Zhang J., Matsuzoe H., 2009, *Dualistic riemannian manifold structure induced from convex functions*, in: *Advances in applied mathematics and global optimization in honor of Gilbert Strang*, Advances in Mechanics and Mathematics **17**, Springer, Berlin, p.1. www.lsa.umich.edu/psych/junz/Zhang and Matsuzoe hessian.pdf. ↑ 67.

[788] Zhu H., 1998, *Generalized Lebesgue spaces and application to statistics*, Santa Fe Inst. Tech. Rep. **98-06-44**, Santa Fe. www.santafe.edu/media/workingpapers/98-06-044.ps. ↑ 2, 10, 48.

[789] Zhu H., 1998, *Unique extension of metric and affine connections to finite measures on infinite spaces*, preprint. ↑ 75.

[790] Zhu H., Rohwer R., 1995, *Information geometric measurements of generalisation*, Tech. Rep. Neural Comp. Res. Group **4350**, Aston. eprints.aston.ac.uk/507/1/NCRG_95_005.pdf. ↑ 46.

[791] Zhu H., Rohwer R., 1997, *Measurements of generalisation based on information geometry*, in: Ellacott S.W. et al. (eds.), *Mathematics of neural networks: models, algorithms and applications*, Kluwer, Dordrecht, p.394. eprints.aston.ac.uk/514/1/NCRG_95_012.pdf. ↑ 46.

[792] Zhu H., Rohwer R., 1998, *Information geometry, bayesian inference, ideal estimates and error decomposition*, Santa Fe Inst. Tech. Rep. **98-06-45**, Santa Fe. omega.albany.edu:8008/ignorance/zhu98.pdf. ↑ 40, 46, 48.

[793] Zălinescu C., 2002, *Convex analysis in general vector spaces*, World Scientific, Singapore. ↑ 32.

Index