The universal von Neumann algebra of smooth four-manifolds revisited

Gábor Etesi

Department of Algebra and Geometry, Institute of Mathematics,

Budapest University of Technology and Economics,

Műegyetem rkp. 3., H-1111 Budapest, Hungary *

August 18, 2023

Abstract

Making use of its smooth structure only, out of a connected oriented smooth 4-manifold a von Neumann algebra is constructed. It is geometric in the sense that is generated by local operators and as a special four dimensional phenomenon it contains all algebraic (i.e., formal or coming from a metric) curvature tensors of the underlying 4-manifold. The von Neumann algebra itself is a hyperfinite factor of II_1-type hence is unique up to abstract isomorphisms of von Neumann algebras. Over a fixed 4-manifold this universal von Neumann algebra admits a particular representation on a Hilbert space such that its unitary equivalence class is preserved by orientation-preserving diffeomorphisms consequently the Murray–von Neumann coupling constant of this representation is well-defined and gives rise to a new and computable real-valued smooth 4-manifold invariant. Its link with Jones' subfactor theory is noticed as well as computations in the simply connected closed case are carried out.

Application to the cosmological constant problem is also discussed. Namely, the aforementioned mathematical construction allows to reformulate the classical vacuum Einstein equation with cosmological constant over a 4-manifold as an operator equation over its tracial universal von Neumann algebra such that the trace of a solution is naturally identified with the cosmological constant. This framework permits to use the observed magnitude of the cosmological constant to estimate by topological means the number of primordial black holes about the Planck era. This number turns out to be negligable which is in agreement with known density estimates based on the Press–Schechter mechanism.

AMS Classification: Primary: 46L10, 83C45, Secondary: 57R55

Keywords: *Smooth* 4*-manifold; Hyperfinite* II_1 *factor; Quantum gravity; Cosmological constant*

1 Introduction and summary

The 20-21$^{\mathrm{st}}$ century has been witness to a great expansion of mathematics and physics bringing a genuinely two-sided interaction between them. The 1980-90's culmination of discoveries in low dimensional differential topology driven by Yang–Mills theory of particle physics has dramatically changed

*E-mail: `etesi@math.bme.hu`

our understanding of four dimensional spaces: nowadays we know that the interplay between topology and smoothness is unexpectedly complicated precisely in four dimensions leading to the existence of a superabundance of smooth four dimensional manifolds. While traditional invariants of differential topology loose power in three and four dimensions, the new quantum invariants provided by various Yang–Mills theories work exactly in these dimensions allowing an at least partial enumeration of manifolds. It is perhaps not just an accidence that quantum invariants are applicable precisely in three and four dimensions, equal to the phenomenological dimensions of physical space and space-time.

It is interesting that unlike Yang–Mills theories, classical general relativity—despite its powerful physical content, too—has not contributed to our understanding of four dimensionality yet. This might follow from the fact that general relativity, unlike Yang–Mills theories with their self-duality phenomena, permits formulations in every dimensions greater than four exhibiting essentially the same properties. This is certainly true when general relativity is considered in its usual fully classical differential-geometric school uniform however four dimensionality enters the game here as well if one tries to link differential geometry with non-commutativity [8].

The purpose of these notes is therefore twofold. On the *mathematical side* they intend to replace in a substantionally new way the overcomplicated, superfluous and, to confess, at many technical steps erroneous construction exhibited in our earlier work [14, Lemmata 2.1, 2.2 and 2.3]. The aim of that construction was to extract from a connected oriented smooth 4-manifold M a II_1 hypefinite factor von Neumann algebra $\mathfrak{R}$. The independent new construction here has exactly the same purpose however, in sharp contrast to our earlier struggles [13, 14], it rests on a standard Clifford algebra construction only hence is both conceptionally and technically very simple; thereby reducing the status of the machinery $M \Rightarrow \mathfrak{R}$ to a sort of so-far-unnoticed triviality. In addition, as a further novelty, a connection with the associated smooth 4-manifold invariant γ introduced and studied in [14, Lemmata 2.4 and 2.5] with Jones' subfactor theory is observed and new computations with this invariant in the simply connected closed case are exhibited.

Our mathematical results are summarized as follows:

Theorem 1.1. *Let M be a connected oriented smooth 4-manifold. Making use of its smooth structure only, a von Neumann algebra $\mathfrak{R}$ can be constructed which is geometric in the sense that it is generated by local operators including all bounded complexified algebraic (i.e., formal or stemming from a metric) curvature tensors of M and $\mathfrak{R}$ itself is a hyperfinite factor of type II_1 hence is unique up to abstract isomorphism of von Neumann algebras.*

Moreover $\mathfrak{R}$ admits a representation on a certain separable Hilbert space over M such that the unitary equivalence class of this representation is invariant under orientation-preserving diffeomorphisms of M. Consequently the Murray–von Neumann coupling constant of this representation gives rise to a smooth invariant $\gamma(M) \in [0,1)$. More precisely it satisfies

$$\gamma(M) = 1 - \frac{1}{x}$$

where $x \in \{4\cos^2\left(\frac{\pi}{n}\right) \mid n \geqq 3\} \cup [4,+\infty)$, the set of Jones' subfactor indices.

This invariant has the following basic properties which are helpful in computations:

Theorem 1.2. *The invariant behaves like $\gamma(\overline{M}) = \gamma(M)$ under reversing orientation, $\gamma(M \setminus Y) = \gamma(M)$ under excision of homologically trivial submanifolds and*

$$\gamma(M\#N) = \frac{\gamma(M)+\gamma(N)}{1+\gamma(M)\gamma(N)}$$

under connected sum.

The invariant in the closed simply connected case can be characterized with the following properties. These show that unfortunately γ is not sensitive enough yet i.e., its construction demands further improvements in order to effectively distinguish smooth structures:

Theorem 1.3. *If M', M'' are connected, closed, simply connected, smooth 4-manifolds which are homeomorphic then $\gamma(M') = \gamma(M'')$.*

More precisely if M is a closed, simply connected, smooth 4-manifold then

$$\gamma(M) = \frac{17^{b_2(M)} - 1}{17^{b_2(M)} + 1}$$

holds.

Thus for example $\gamma(S^4) = 0, \gamma(\mathbb{C}P^2) = \frac{8}{9}$ and $\gamma(\mathbb{C}P^1 \times \mathbb{C}P^1) = \frac{144}{145}$. We also find for instance for the $K3$ surface having $b_2(K3) = 22$ that already $1 - \gamma(K3) \approx 1.70 \times 10^{-27}$ and the general asymptotics of γ in the simply connected case is $0 < 1 - \gamma(M) \approx \mathrm{e}^{-\mathrm{const.}\, b_2(M)}$. This indicates that this invariant maps four dimensional smooth structures into $[0, 1)$ in a logarithmic way in some sense.

Let us add some comments here. The set of Jones indices splits into a discrete and a continuous part. Subfactors belonging to the discrete part $\{4\cos^2\left(\frac{\pi}{n}\right) \,|\, n \geqq 3\}$ have been completely classified [21] and in turn they follow an ADE pattern (with the odditiy that no subfactors corresponding to D_{2k+1} and E_7 exist) [20]. The set of subfactors belonging to the continuous portion $[4, +\infty)$ is however very wild and only partial results are known mainly for the subinterval $[4, 5] \subset [4, +\infty)$ or a bit more (cf. e.g. [18] for a survey and results). The connected sum formula exhibited here in Theorem 1.2 shows that all non-prime 4-manifolds (i.e., which admit non-trivial connected sum decompositions) belong via their invariant to the continuous regime i.e., $\frac{1}{1-\gamma(M)} \in [4, +\infty)$ furthermore we also find by Theorem 1.3 that this is also true for the prime manifold complex projective plane because $\frac{1}{1-\gamma(\mathbb{C}P^2)} = 9$. This observation strongly hints that smooth 4-manifolds might provide a rich reservoir of subfactors in the wild index range moreover poses the question whether or not smooth 4-manifolds corresponding to the tame (i.e., discrete) range exist (e.g. $\gamma(S^4) = 0$ corresponds to the Jones index $4\cos^2\left(\frac{\pi}{3}\right) = 1$ hence the trivial case $n = 3$ in the discrete part is realized by some 4-manifolds).

The other purpose of these notes, on the *physical side*, is to provide further clarifications to the material collected in [13, 14] regarding how to interpret the merely mathematical considerations above within a physical theory as naturally as possible. The outstanding problem of contemporary theoretical physics is how to unify the obviously successful and mathematically consistent *theory of general relativity* with the obviously successful but yet mathematically problematic *relativistic quantum field theory*. It has been generally believed that these two fundamental pillars of modern theoretical physics are in a clash not only by the different mathematical tools they use but even at a deep conceptional level: classical notions of general relativity such as the space-time event, the light cone or the event horizon of a black hole are too “sharp” objects from a quantum theoretic viewpoint meanwhile relativistic quantum field theory is not background independent from the aspect of general relativity.

Our interpretational efforts naturally fit into this context: they rest on the observation made explicit already in Theorem 1.1 that the constructed operator algebra $\mathfrak{R}$ is unique i.e. independent of the particular manifold we have started with and, as a curiosity of four dimensions, it contains among its members algebraic curvature tensors over M. Since curvature plays a central role in classical general relativity it is quite natural to place this interpretation within the broad realm of quantum gravity. More concretely first by recalling the basic dictionary from [13, Section 3] we exhibit a straightforward generalization of the classical vacuum Einstein equation with cosmological constant over a 4-manifold M to an operator equation over its von Neumann algebra $\mathfrak{R}$ (see Definition 3.1 below). One interesting feature of this generalization is that the trace of those operators which satisfy this “quantum vacuum Einstein equation” can be identified with the cosmological constant. Second, we use the previous mathematical results to find solutions to this equation whose traces (hence the cosmological constants they give rise to) are equal to the smooth invariant $1-\gamma$ of Theorem 1.1 hence fall into $(0,1]\subset\mathbb{R}$. Consequently they are always *strictly positive* but *small* numbers (cf. [2] too). Third, the observed small positive value of the cosmological constant [23] allows in this framework to obtain an estimate on the number of primordial black holes in the very early Universe; the magnitude of this number turns out to be $\sim 10^2$ hence their presence is negligable which is in accord with the current conviction in the cosmologist and perhaps particle physicist community [6]. Other recent works treating hyperfinite II factors from a gravitational perspective are [3, 7].

The present paper is heavily based on our earlier attempts [13, 14] but with substantial clarifications and new extensions. It is organized as follows. Section 2 is a self-contained presentation of all the mathematical constructions involved in Theorems 1.1, 1.2 and 1.3 together with their proofs. The language of this part is therefore rather mathematical. Then Section 3 contains the physics by offering a quantum generalization of the classical vacuum Einstein equation (see Definition 3.1) with an application to the cosmological constant problem. The language of this part is therefore rather physical. The physicist-minded reader may skip Section 2 and read Section 3 first hopefully without problem.

2 Mathematical constructions

In this section we shall first exhibit a simple self-contained two-step construction of a von Neumann algebra attached to any oriented smooth 4-manifold. The structure of this algebra will be also explored in some detail. This is then will be followed by introducing a new smooth 4-manifold invariant whose simplest properties will be examined too.

Construction of an algebra. Take the isomorphism class of a connected oriented smooth 4-manifold (without boundary) and from now on let M be a once and for all fixed representative in it carrying the action of its own orientation-preserving group of diffeomorphisms $\mathrm{Diff}^+(M)$. Among all tensor bundles $T^{(p,q)}M$ over M the 2^{nd} exterior power $\wedge^2T^*M\subset T^{(0,2)}M$ is the only one which can be endowed with a pairing in a natural way i.e., with a pairing extracted from the smooth structure (and the orientation) of M alone. Indeed, consider its associated vector space $\Omega^2_c(M):=C^\infty_c(M;\wedge^2T^*M)$ of compactly supported smooth 2-forms on M. Define a pairing $\langle\,\cdot\,,\,\cdot\,\rangle_{L^2(M)}:\Omega^2_c(M)\times\Omega^2_c(M)\to\mathbb{R}$ via integration:

$$\langle\alpha,\beta\rangle_{L^2(M)}:=\int\limits_M\alpha\wedge\beta \tag{1}$$

and observe that this pairing is non-degenerate however is *indefinite* in general thus can be regarded as an indefinite scalar product on $\Omega^2_c(M)$. It therefore induces an indefinite real quadratic form Q on $\Omega^2_c(M)$ given by $Q(\alpha):=\langle\alpha,\alpha\rangle_{L^2(M)}$. Let $C(M)$ denote the complexification of the infinite dimensional real Clifford algebra associated with $(\Omega^2_c(M),Q)$. Because Clifford algebras are usually constructed out of definite quadratic forms, we summarize this construction [19, Section I.§3] to make sure that the resulting object $C(M)$ is well-defined i.e. is not sensitive for the indefiniteness of (1). To begin with, let $V_m\subset\Omega^2_c(M)$ be an m dimensional real subspace and assume that $Q_{r,s}:=Q|_{V_m}$ has signature (r,s) on V_m that is, the maximal positive definite subspace of V_m with respect to $Q_{r,s}$ has dimension r while the dimension of the maximal negative definite subspace is s such that $r+s=m$ by the non-degeneracy of $Q_{r,s}$. Then out of the input data $(V_m,Q_{r,s})$ one constructs in the standard way a finite dimensional real Clifford algebra $C_{r,s}(M)$ with unit $1\in C_{r,s}(M)$ and an embedding $V_m\subset C_{r,s}(M)$ with the property $\alpha^2 = Q_{r,s}(\alpha)1$ for every element $\alpha \in V_m$. This real algebra depends on the signature (r,s) however fortunately its complexification $C_m(M) := C_{r,s}(M) \otimes \mathbb{C}$ is already independent of it. In fact, if $\mathfrak{M}_k(\mathbb{C})$ denotes the algebra of $k \times k$ complex matrices, then it is well-known [19, Section I.§3] that $C_0(M) \cong \mathfrak{M}_1(\mathbb{C})$ while $C_1(M) \cong \mathfrak{M}_1(\mathbb{C}) \oplus \mathfrak{M}_1(\mathbb{C})$ and the higher dimensional cases follow from the complex periodicity $C_{m+2}(M) \cong C_m(M) \otimes \mathfrak{M}_2(\mathbb{C})$. Consequently depending on the parity $C_m(M)$ is isomorphic to either $\mathfrak{M}_{2^{\frac{m}{2}}}(\mathbb{C})$ or $\mathfrak{M}_{2^{\frac{m-1}{2}}}(\mathbb{C}) \oplus \mathfrak{M}_{2^{\frac{m-1}{2}}}(\mathbb{C})$. These imply that 2-step-chains of successive embeddings of real subspaces $V_m \subset V_{m+1} \subset V_{m+2} \subset \Omega^2_c(M)$ starting with $V_0 = \{0\}$ and given by iterating $\omega \mapsto \binom{\omega}{0}$ provide us with injective algebra homomorphisms $\mathfrak{M}_{2^{\frac{m}{2}}}(\mathbb{C}) \hookrightarrow \mathfrak{M}_{2^{\frac{m}{2}+1}}(\mathbb{C})$ having the shape $A \mapsto \left(\begin{smallmatrix} A & 0 \\ 0 & A \end{smallmatrix}\right)$. Therefore $C(M)$ is isomorphic to the injective limit of this directed system, that is there exists a linear-algebraic isomorphism

$$C(M) \cong \bigcup_{n=0}^{+\infty} \mathfrak{M}_{2^n}(\mathbb{C}) \tag{2}$$

or equivalently

$$C(M) \cong \mathfrak{M}_2(\mathbb{C}) \otimes \mathfrak{M}_2(\mathbb{C}) \otimes \dots$$

because this injective limit is also isomorphic to the infinite tensor product of $\mathfrak{M}_2(\mathbb{C})$'s. For clarity note that being (1) a non-local operation, $C(M)$ is a genuine global infinite dimensional object.

It is well-known (cf. [8, Section I.3]) that *any* complexified infinite Clifford algebra like $C(M)$ above generates *the* II_1-type hyperfinite factor von Neumann algebra. Let us summarize this procedure too (cf. [1, Section 1.1.6]). It readily follows that $C(M)$ possesses a unit $1 \in C(M)$ and its center comprises the scalar multiples of the unit only. Moreover $C(M)$ continues to admit a canonical embedding $\Omega^2_c(M;\mathbb{C}) \subset C(M)$ satisfying $\omega^2 = Q(\omega)1$ where now Q denotes the quadratic form induced by the complex-bilinear extension of (1). We also see via (2) already that $C(M)$ is a complex $*$-algebra whose $*$-operation (provided by taking Hermitian matrix transpose, a non-local operation) is written as $A \mapsto A^{*}$. The isomorphism (2) also shows that if $A \in C(M)$ then one can pick the smallest $n \in \mathbb{N}$ such that $A \in \mathfrak{M}_{2^n}(\mathbb{C})$ consequently A has a finite trace defined by $\tau(A) := 2^{-n}\mathrm{Trace}(A)$ i.e., taking the usual normalized trace of the corresponding $2^n \times 2^n$ complex matrix. It is straightforward that $\tau(A) \in \mathbb{C}$ does not depend on n. We can then define a sesquilinear inner product on $C(M)$ by $(A,B) := \tau(AB^{*})$ which is non-degenerate thus the completion of $C(M)$ with respect to the norm $\|\cdot\|$ induced by $(\cdot\,,\cdot)$ renders $C(M)$ a complex Hilbert space what we shall write as $\mathscr{H}$ and its Banach algebra of all bounded linear operators as $\mathfrak{B}(\mathscr{H})$. Multiplication in $C(M)$ from the left on itself is continuous hence gives rise to a representation $\pi : C(M) \to \mathfrak{B}(\mathscr{H})$. Finally our central object effortlessly emerges as the weak closure of the image of $C(M)$ under π within $\mathfrak{B}(\mathscr{H})$ or equivalently, by referring to von Neumann's bicommutant theorem [1, Theorem 2.1.3] we put

$$\mathfrak{R} := (\pi(C(M)))'' \subset \mathfrak{B}(\mathscr{H}) .$$

This von Neumann algebra of course admits a unit $1 \in \mathfrak{R}$ moreover continues to have trivial center i.e., is a factor. Moreover by construction it is hyperfinite. The trace τ as defined extends from $C(M)$ to $\mathfrak{R}$ and satisfies $\tau(1) = 1$. Moreover [1, Proposition 4.1.4] this trace is unique on $\mathfrak{R}$. Likewise we obtain by extension a representation $\pi : \mathfrak{R} \to \mathfrak{B}(\mathscr{H})$. The canonical inclusion $\Omega^2_c(M;\mathbb{C}) \subset C(M)$ recorded above extends to both $\Omega^2_c(M;\mathbb{C}) \subset \mathscr{H}$ and $\Omega^2_c(M;\mathbb{C}) \subset \mathfrak{R}$. Thus in order to carefully distinguish the two different completions $\mathfrak{R}$ and $\mathscr{H}$ of one and the same object $C(M)$ we shall write $A \in \mathfrak{R}$ but $\hat{B} \in \mathscr{H}$ from now on as usual. This is necessary since $\mathfrak{R}$ and $\mathscr{H}$ are very different for example as $\mathrm{U}(\mathscr{H})$-modules: given a unitary operator $V \in \mathrm{U}(\mathscr{H})$ then $A \in \mathfrak{R}$ is acted upon as $A \mapsto VAV^{-1}$ but $\hat{B} \in \mathscr{H}$

transforms as $\hat{B} \mapsto V\hat{B}$. Using this notation and introducing $\hat{A} := \pi(A)\hat{1}$ the trace always can be written as a scalar product with the image of the unit in $\mathscr{H}$ that is, for every $A \in \mathfrak{R}$ we have

$$\tau(A) = (\hat{A}, \hat{1})$$

yielding a general and geometric expression for the trace.

Exploring the algebra $\mathfrak{R}$. Before proceeding further let us make a digression here to gain a better picture. This is desirable because taking the weak limit like $\mathfrak{R}$ of some explicitly known structure like $C(M)$ often involves a sort of loosing control over the latter. Nevertheless we already know promisingly that $\mathfrak{R}$ is a hyperfinite factor von Neumann algebra of II_1-type. Let us now exhibit some of its elements.

1. Our first examples are the 2-forms themselves as it follows from the already mentioned canonical embedding $\Omega^2_c(M;\mathbb{C}) \subset C(M)$ combined with $C(M) \subset \mathfrak{R}$ yielding $\Omega^2_c(M;\mathbb{C}) \subset \mathfrak{R}$. Next, proceeding along the same lines we can construct an interesting natural continuous embedding

$$i_M : M \longrightarrow \mathfrak{R} \tag{3}$$

of any connected oriented smooth 4-manifold M into its $\mathfrak{R}$. To every sufficiently nice closed subset $\emptyset \subsetneqq X \subsetneqq M$ there is an associated linear subspace $\Omega^2_c(M,X;\mathbb{C}) \subset \Omega^2_c(M;\mathbb{C})$ consisting of compactly supported smooth 2-forms vanishing at least along X. Furthermore we have seen that as a by-product of the Clifford algebra construction there exists a canonical embedding $\Omega^2_c(M;\mathbb{C}) \subset \mathscr{H}$ too. In this way to every point $x \in M$ one can attach a closed subspace $V_x \subset \mathscr{H}$ defined by taking the closure of $\Omega^2_c(M,x;\mathbb{C})$ within $\mathscr{H}$. Let $Q_x : \mathscr{H} \to V_x$ be the corresponding orthogonal projection. *A priori* $Q_x \in \mathfrak{B}(\mathscr{H})$ however in fact $Q_x \in \mathfrak{R}$. This is because V_x arises as the image of an operator $A_x \in \mathfrak{R}$ which for instance looks like extending $\omega \mapsto f_x\omega$ with a bounded smooth function f_x vanishing at $x \in M$ from $\Omega^2_c(M;\mathbb{C})$ to $\overline{\Omega^2_c(M;\mathbb{C})} \subset \mathscr{H}$ and defined to be zero on $\overline{\Omega^2_c(M;\mathbb{C})}^{\perp} \subset \mathscr{H}$. It readily follows that the resulting map $x \mapsto Q_x$ is injective and continuous hence gives rise to a continuous embedding. To make a comparison, in *algebraic geometry* points are characterized by maximal ideals of an abstractly given commutative ring. Here the corresponding objects would therefore be the maximal two-sided (weakly closed) ideals of a von Neumann algebra (regarded as a non-commutative ring). However in sharp contrast to the commutative situation a tracial factor von Neumann algebra is always *simple* (cf. e.g. [1, Proposition 4.1.5]) consequently in our case the concept of ideals cannot be used to characterize points hence the reason we used rather special elements of the von Neumann algebra. (3) is a counterpart of embedding Riemannian manifolds into Hilbert spaces via heat kernel techniques [4].

2. To see more examples, let us return to the Clifford algebra in (2) for a moment. We already know that there exists a canonical embedding $\Omega^2_c(M;\mathbb{C}) \subset C(M)$. In addition to this let us find a Clifford module for $C(M)$. Consider again any finite *even* dimensional approximation $C_m(M) = C_{r,s}(M) \otimes \mathbb{C}$ constructed from $(V_m, Q_{r,s})$ where now $V_m \subset \Omega^2_c(M)$ is a real even $m = r+s$ dimensional subspace. Choose any $2^{\frac{m}{2}}$ dimensional complex vector subspace S_m within $\Omega^2_c(M;\mathbb{C})$. If $\mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ denotes the associative algebra of *all* $\mathbb{C}$-linear transformations of $\Omega^2_c(M;\mathbb{C})$ then $S_m \subset \Omega^2_c(M;\mathbb{C})$ induces an embedding $\mathrm{End}\, S_m \subset \mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ moreover we know that $\mathrm{End}\, S_m \cong \mathfrak{M}_{2^{\frac{m}{2}}}(\mathbb{C}) \cong C_m(M)$. Therefore we obtain a non-canonical inclusion $C_m(M) \subset \mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ for every fixed $m \in 2\mathbb{N}$. Furthermore $S_m \subset S_{m+2} \subset \Omega^2_c(M;\mathbb{C})$ given by $\omega \mapsto \binom{\omega}{0}$ induces a sequence $C_m(M) \subset C_{m+2}(M) \subset \mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ for Clifford algebras which is compatible with the previous ascending chain of their matrix algebra realizations. Consequently taking the limit $m \to +\infty$ we come up with a non-canonical injective linear-algebraic homomorphism

$$C(M) \subset \mathrm{End}(\Omega^2_c(M;\mathbb{C}))\ . \tag{4}$$

Of course, unlike $\Omega^2_c(M;\mathbb{C}) \subset C(M)$ above, (4) does not exist in the finite dimensional case.

Although the algebra $\mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ is yet too huge, we can at least exhibit some of its elements. The simplest ones are the 2-forms themselves because $\Omega^2_c(M;\mathbb{C})\subset C(M)\subset \mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ holds as we already know.[1] Moreover $C^\infty(M;\mathrm{End}(\wedge^2T^*M\otimes\mathbb{C}))\subset\mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ i.e., bundle morphisms are also included. These are local (algebraic) operators but are important because they allow to make a contact with local *four dimensional* differential geometry.[2] A peculiarity of four dimensions is that $C^\infty(M;\mathrm{End}(\wedge^2T^*M\otimes\mathbb{C}))$ contains the space of curvature tensors on M. If (M,g) is an oriented Riemannian 4-manifold then its Riemannian curvature tensor R_g is a member of this subalgebra: with respect to the splitting of 2-forms into their (anti)self-dual parts it looks like (cf. [24])

$$R_g=\begin{pmatrix}\frac{1}{12}\mathrm{Scal}+\mathrm{Weyl}^+ & \mathrm{Ric}_0\\ \mathrm{Ric}_0^* & \frac{1}{12}\mathrm{Scal}+\mathrm{Weyl}^-\end{pmatrix}\;:\;\begin{matrix}\Omega^+_c(M;\mathbb{C})\\ \bigoplus\\ \Omega^-_c(M;\mathbb{C})\end{matrix}\longrightarrow\begin{matrix}\Omega^+_c(M;\mathbb{C})\\ \bigoplus\\ \Omega^-_c(M;\mathbb{C})\end{matrix}\qquad(5)$$

and more generally, $\mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ contains all algebraic (i.e. formal only, not stemming from a metric) curvature tensors R over M.

To see non-local examples, let us say that $\omega\in L^1_{loc}(M;\wedge^2T^*M\otimes\mathbb{C})$ if $\langle\omega,\varphi\rangle_{L^2(M)}$ given by (1) exists for all $\varphi\in\Omega^2_c(M;\mathbb{C})$. Then a tensor field K over $M\times M$ is called an *admissible double* 2*-form over* M if the map $x\mapsto K(x,y)$ belongs to $C^\infty_c(M;\wedge^2T^*M\otimes\mathbb{C})=\Omega^2_c(M;\mathbb{C})$ for every $y\in M$ while $y\mapsto K(x,y)$ gives rise to an element of $L^1_{loc}(M;\wedge^2T^*M\otimes\mathbb{C})$ for every $x\in M$. Picking $R\in C^\infty(M;\mathrm{End}(\wedge^2T^*M\otimes\mathbb{C}))$ the map $\omega\mapsto P\omega$ defined by

$$P\omega(x):=\int\limits_{y\in M}K(x,y)\wedge(R(y)\omega(y))$$

is a non-local pseudo-differential operator acting on 2-forms and belongs to $\mathrm{End}(\Omega^2_c(M;\mathbb{C}))$. Furthermore orientation-preserving diffeomorphisms act on 2-forms via non-local pullback hence we conclude that $\mathrm{Diff}^+(M)\subset\mathrm{End}(\Omega^2_c(M;\mathbb{C}))$. Note that the Lie algebra $\mathrm{Lie}(\mathrm{Diff}^+(M))\cong C^\infty_c(M;TM)$ consisting of compactly supported real vector fields acts $\mathbb{C}$-linearly on $\Omega^2_c(M;\mathbb{C})$ through Lie derivatives hence we also find that $\mathrm{Lie}(\mathrm{Diff}^+(M))\subset\mathrm{End}(\Omega^2_c(M;\mathbb{C}))$. However these are local (non-algebraic but differential) operators again.

How to decide whether or not these elements of $\mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ belong to $\mathfrak{R}$? The key concept here is the trace. Compared with the above trace expression generally valid on $\mathfrak{R}$, more specific trace formulata are obtained if M is endowed with a normalized Riemannian metric g i.e., the corresponding volume form $\mu_g=*1$ satisfies $\int_M\mu_g=1$. The unique sesquilinear extension of g induces a positive definite sesquilinear L^2-scalar product

$$(\varphi,\psi)_{L^2(M,g)}:=\int\limits_{x\in M}g(\varphi(x),\psi(x))\mu_g(x)=\int\limits_M\varphi\wedge*\overline{\psi}$$

on $\Omega^2_c(M;\mathbb{C})$. If $\{\varphi_1,\varphi_2,\dots\}$ is a smooth orthonormal frame in $\Omega^2_c(M;\mathbb{C})$ then it readily follows that the trace of any $B\in\mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ formally looks like

$$\tau(B)=\lim_{n\to+\infty}\frac{1}{2^n}\sum_{i=1}^{2^n}(B\varphi_i\,,\,\varphi_i)_{L^2(M,g)}$$

[1] The extension of the Clifford multiplication $(V_m\otimes\mathbb{C})\times S_m\to S_m$ by $m\to+\infty$ thus induces a product structure $\Omega^2_c(M;\mathbb{C})\times\Omega^2_c(M;\mathbb{C})\to\Omega^2_c(M;\mathbb{C})$ rendering the space of 2-forms an infinite dimensional non-unital non-commutative associative algebra carrying a non-degenerate symmetric pairing by (1); hence resembling a Frobenius algebra. It is interesting that the orientation and smooth structure on M alone induces this structure.

[2] In fact all the constructions so far work for an arbitrary oriented and smooth $4k$-manifold with $k=0,1,2,\dots$ (note that in $4k+2$ dimensions the indefinite pairing (1) gives rise to a symplectic structure on $2k+1$-forms).

and, if exists, is independent of the frame used. Obviously $B \in \mathrm{End}(\Omega^2_c(M;\mathbb{C})) \cap C(M)$ if and only if the sum on the right hand side is constant after finitely many terms; and an inspection of this trace expression shows that in general $B \in \mathrm{End}(\Omega^2_c(M;\mathbb{C})) \cap \mathfrak{R}$ if and only if $\tau(B)$ exists. As a consequence note that $\mathrm{End}(\Omega^2_c(M;\mathbb{C})) \cap \mathfrak{R}$ is already independent of the particular inclusion (4). An example: any $\Phi \in \mathrm{Diff}^+(M)$ acts on $\mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ by conjugation hence using (4) extends to a unitary operator on $\mathscr{H}$ yielding $|\tau(\Phi)| = 1$ consequently we obtain $\mathrm{Diff}^+(M) \subset \mathrm{End}(\Omega^2_c(M;\mathbb{C})) \cap \mathfrak{R}$.

When $R \in C^\infty(M;\mathrm{End}(\wedge^2 T^*M \otimes \mathbb{C}))$ the previous trace formula can be further specified because one can compare the global trace $\tau(R)$ and the local trace function $x \mapsto \mathrm{tr}(R(x))$ given by the pointwise traces of the local operators $R(x) : \wedge^2 T^*_x M \otimes \mathbb{C} \to \wedge^2 T^*_x M \otimes \mathbb{C}$ at every $x \in M$. Recall that $\mathfrak{R}$ has been constructed as the weak closure of the Clifford algebra (2). In fact [1, Section 1.1.6] the universality of $\mathfrak{R}$ permits to obtain it from other matrix algebras too, like for instance from $\bigcup\limits_{n=0}^{+\infty} \mathfrak{M}_{6^n}(\mathbb{C})$ whose weak closure therefore is again $\mathfrak{R}$. By the aid of this altered construction we can formally start with

$$\tau(R) = \lim_{n\to+\infty} \frac{1}{6^n} \sum_{i=1}^{6^n} (R\varphi_i\, ,\, \varphi_i)_{L^2(M,g)} \ .$$

Fix $n \in \mathbb{N}$, write $M_n := \bigcap\limits_{i=1}^{6^n} \mathrm{supp}\varphi_i \subseteqq M$ and take a point $x \in M_n$. Since $\dim_{\mathbb{C}}(\wedge^2 T^*_x M \otimes \mathbb{C}) = \binom{4}{2} = 6$ the maximal number of completely disjoint linearly independent sub-6-tuples in $\{\varphi_1(x), \varphi_2(x), \dots, \varphi_{6^n}(x)\}$ is equal to $\frac{6^n}{6} = 6^{n-1}$. Moreover it follows from Sard's lemma that with a generic smooth choice for $\{\varphi_1, \varphi_2, \dots\}$ the subset of those points $y \in M_n$ where this number is less than 6^{n-1} has measure zero in M_n with respect to the measure μ_g. Consequently

$$\sum_{i=1}^{6^n} (R\varphi_i, \varphi_i)_{L^2(M_n,g)} = \int\limits_{x\in M_n} \sum_{i=1}^{6^n} g(R(x)\varphi_i(x), \varphi_i(x))\, \mu_g(x) = 6^{n-1} \int\limits_{x\in M_n} \mathrm{tr}(R(x))\mu_g(x) \ .$$

Since $\{\varphi_1, \varphi_2, \dots\}$ is a basis in $\Omega^2_c(M;\mathbb{C})$ therefore $M \setminus \bigcup\limits_{n=0}^{+\infty} M_n$ has measure zero as well we can let $n \to +\infty$ to end up with

$$\tau(R) = \frac{1}{6} \int\limits_M \mathrm{tr}(R)\mu_g \tag{6}$$

and again $R \in C^\infty(M;\mathrm{End}(\wedge^2 T^*M \otimes \mathbb{C})) \cap \mathfrak{R}$ if and only if (6) exists. Also note that τ is in fact the generalization of the total scalar curvature of a Riemannian manifold. So taking into account (4) we obtain useful criteria for checking whether or not an operator in $\mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ belongs to the completion of $C(M)$ which is $\mathfrak{R}$. For instance if the curvature R_g of (M,g) as an operator in (5) is bounded which means that

$$\sup_{\|\omega\|_{L^2(M,g)}=1} \|R_g\omega\|_{L^2(M,g)} \leqq K < +\infty$$

then

$$0 \leqq |\tau(R_g)| \leqq \lim_{n\to+\infty} \frac{1}{2^n} \sum_{i=1}^{2^n} |(R_g\varphi_i\, ,\, \varphi_i)_{L^2(M,g)}| \leqq \lim_{n\to+\infty} \frac{1}{2^n} \sum_{i=1}^{2^n} \|R_g\varphi_i\|_{L^2(M,g)} \leqq K$$

yielding $R_g \in C^\infty(M;\mathrm{End}(\wedge^2 T^*M \otimes \mathbb{C})) \cap \mathfrak{R}$ and more generally $R \in C^\infty(M;\mathrm{End}(\wedge^2 T^*M \otimes \mathbb{C})) \cap \mathfrak{R}$ if R is bounded. Similar condition stems from (6).

3. We close the partial comprehension of $\mathfrak{R}$ with an observation regarding its overall structure. The canonical inclusion $\Omega^2_c(M;\mathbb{C})\subset\mathfrak{R}$ (and likewise $\Omega^2_c(M;\mathbb{C})\subset\mathscr{H}$) implies that $\mathfrak{R}$ (and likewise $\mathscr{H}$) is generated by all finite products $\omega_1\omega_2\dots\omega_k$ of 2-forms within $C(M)$ however this is not very informative. Rather consider the well-known short exact sequence of groups involving the fiberwise automorphism group (the gauge group) and the global automorphism group of the vector bundle $\wedge^2T^*M\otimes\mathbb{C}$ as well as the orientation-presering diffeomorphism group of the underlying M respectively:

$$1\longrightarrow C^\infty(M;\mathrm{Aut}(\wedge^2T^*M\otimes\mathbb{C}))\longrightarrow\mathrm{Aut}(\Omega^2_c(M;\mathbb{C}))\longrightarrow\mathrm{Diff}^+(M)\longrightarrow 1$$

whose shape at the Lie algebra level looks like

$$0\longrightarrow C^\infty(M;\mathrm{End}(\wedge^2T^*M\otimes\mathbb{C}))\longrightarrow\mathrm{End}(\Omega^2_c(M;\mathbb{C}))\longrightarrow\mathrm{Lie}(\mathrm{Diff}^+(M))\longrightarrow 0\ .$$

We already have an embedding (4). In addition to this there exists an isomorphism of Lie algebras $L: C^\infty_c(M;TM)\to\mathrm{Lie}(\mathrm{Diff}^+(M))$ such that $X\mapsto L_X$ is nothing but taking Lie derivative with respect to a compactly supported real vector field where the first-order $\mathbb{C}$-linear differential operator L_X is supposed to act on 2-forms hence $\mathrm{Lie}(\mathrm{Diff}^+(M))\subset\mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ as we know already too. Therefore the intersection of the latter sequence with $C(M)$ is meaningful and gives

$$0\longrightarrow C^\infty(M;\mathrm{End}(\wedge^2T^*M\otimes\mathbb{C}))\cap C(M)\longrightarrow C(M)\longrightarrow\mathrm{Lie}(\mathrm{Diff}^+(M))\cap C(M)\longrightarrow 0\ .$$

The second term consists of fiberwise algebraic hence local operators having finite trace (6) and likewise the fourth terms consists of finite trace Lie derivatives thus belongs to the class of local operators too. Since the vector spaces underlying $C(M)$ considered either as an associative or a Lie algebra are isomorphic we conclude, as an important structural observation, that the overall construction here is geometric in the sense that *the algebra $\mathfrak{R}$ is generated by (trace class) local operators*.

Summing up all of our findings so far: $\mathfrak{R}$ is a hyperfinite factor von Neumann algebra of II_1-type associated to M such that the solely input in its construction has been the pairing (1). Hence $\mathfrak{R}$ depends only on the orientation and the smooth structure of M in a functorial way. It contains, certainly among many other non-local operators, the space M itself via (3), its orientation-preseving diffeomorphisms as well as its space of bounded algebraic curvature tensors. Nevertheless $\mathfrak{R}$ is geometric in the sense that it is generated by M's finite trace local operators alone. It is remarkable that despite the plethora of smooth 4-manifolds detected since the 1980's their associated von Neumann algebras here are unique up to isomorphisms of von Neumann algebras (cf. e.g. [1, Theorem 11.2.2]) offering a sort of justification terming $\mathfrak{R}$ as "universal". One is therefore tempted to look upon $\mathfrak{R}$ as a natural common non-commutative space generalization of either all the pure oriented smooth 4-manifolds M or rather all the (pseudo-)Riemannian ones (M,g). This universality also justifies the simple notation $\mathfrak{R}$ used throughout the text.

A new smooth 4-manifold invariant. After these preliminary considerations we are in a position to approach four dimensional smoothness from a new operator-algebraic direction.

Lemma 2.1. *Let M be a connected oriented smooth 4-manifold and $\mathfrak{R}$ its von Neumann algebra with trace τ as before. Then there exists a complex separable Hilbert space $\mathscr{I}(M)^\perp$ and a representation $\rho_M:\mathfrak{R}\to\mathfrak{B}(\mathscr{I}(M)^\perp)$ with the following properties. If $\pi:\mathfrak{R}\to\mathfrak{B}(\mathscr{H})$ is the representation constructed above then $\{0\}\subseteqq\mathscr{I}(M)^\perp\subsetneqq\mathscr{H}$ and $\rho_M=\pi|_{\mathscr{I}(M)^\perp}$ holds; therefore, although ρ_M can be the trivial representation, it is surely not unitary equivalent to the standard representation. Moreover the unitary equivalence class of ρ_M is invariant under orientation-preserving diffeomorphisms of M.*

Thus the Murray–von Neumann coupling constant[3] of ρ_M is invariant under orientation-preserving diffeomorphisms. Writing $P_M:\mathscr{H}\to\mathscr{I}(M)^\perp$ for the orthogonal projection the coupling constant is equal to $\tau(P_M)\in[0,1)\subset\mathbb{R}_+$ consequently $\gamma(M):=\tau(P_M)$ is a smooth 4-manifold invariant.

[3] Also called the $\mathfrak{R}$-*dimension* of a left $\mathfrak{R}$-module hence denoted $\dim_{\mathfrak{R}}$, cf. [1, Chapter 8].

Proof. First let us exhibit a representation of $\mathfrak{R}$; this construction is inspired by the general Gelfand–Naimark–Segal technique however exploits the special features of our construction so far as well. Pick a pair (Σ,ω) consisting of an (immersed) closed orientable surface $\Sigma \looparrowright M$ with induced oriantation and $\omega \in \Omega^2_c(M;\mathbb{C})$ which is also closed i.e., $\mathrm{d}\omega = 0$. Consider the differential geometric $\mathbb{C}$-linear functional $F_{\Sigma,\omega} : \mathfrak{R} \to \mathbb{C}$ by continuously extending

$$A \longmapsto \int\limits_\Sigma A\omega$$

from $\mathrm{End}(\Omega^2_c(M;\mathbb{C}))\cap\mathfrak{R}$. This extension is unique because $\mathrm{End}(\Omega^2_c(M;\mathbb{C}))\cap\mathfrak{R}$ is norm-dense in $\mathfrak{R}$. In case of $F_{\Sigma,\omega}(1) \neq 0$ let $\{0\} \subseteqq I(M) \subseteqq \mathfrak{R}$ be the subset of elements $A \in \mathfrak{R}$ satisfying $F_{\Sigma,\omega}(A^*A) = 0$. In fact for all pairs (Σ,ω) we obviously find $\{0\} \subsetneqq I(M)$ and $I(M)\cap\mathbb{C}1 = \{0\}$ hence $I(M) \subsetneqq \mathfrak{R}$ too. We assert that $I(M)$ is a multiplicative left-ideal in $\mathfrak{R}$ which is independent of (Σ,ω). In the case of $F_{\Sigma,\omega}(1) = 0$ we put $I(M) = \mathfrak{R}$ hence it is again independent of (Σ,ω) but trivially in this way.

Consider the case when $F_{\Sigma,\omega}(1) \neq 0$. Then we can assume that $F_{\Sigma,\omega}(1) = 1$ hence $F_{\Sigma,\omega}$ is a positive functional; applications of the standard inequality $|F_{\Sigma,\omega}(A^*B)|^2 \leqq F_{\Sigma,\omega}(A^*A)F_{\Sigma,\omega}(B^*B)$ show that $I_{\Sigma,\omega}(M)$ defined by the elements satisfying $F_{\Sigma,\omega}(A^*A) = 0$ is a multiplicative left-ideal in $\mathfrak{R}$. Concerning its ω-dependence, let $\omega' \in \Omega^2_c(M;\mathbb{C})$ be another closed 2-form having the property $F_{\Sigma,\omega'}(1) = 1$; since neither ω nor ω' are identically zero, we can pick an *invertible* operator $T \in \mathrm{Aut}(\Omega^2_c(M;\mathbb{C}))$ satisfying $\omega'|_\Sigma = T\omega|_\Sigma$. Then by $F_{\Sigma,\omega'}(A^*A) = F_{\Sigma,\omega}(A^*AT)$ and applying the above inequality in the form $|F_{\Sigma,\omega'}(A^*AT)|^2 \leqq F_{\Sigma,\omega}(A^*A)F_{\Sigma,\omega}((AT)^*(AT))$ we find $I_{\Sigma,\omega'}(M) \subseteqq I_{\Sigma,\omega}(M)$. Likewise, making use of $F_{\Sigma,\omega}(A^*A) = F_{\Sigma,\omega'}(A^*AT^{-1})$ and $|F_{\Sigma,\omega}(A^*AT^{-1})|^2 \leqq F_{\Sigma,\omega'}(A^*A)F_{\Sigma,\omega'}((AT^{-1})^*(AT^{-1}))$ imply the converse inequality $I_{\Sigma,\omega'}(M) \supseteqq I_{\Sigma,\omega}(M)$. Consequently $I_{\Sigma,\omega'}(M) = I_{\Sigma,\omega}(M)$.

Concerning the general (Σ,ω)-dependence of $I_{\Sigma,\omega}$ we argue as follows. Let $\eta_\Sigma \in \Omega^2(M;\mathbb{R})$ be a closed real 2-form representing the Poincaré-dual $[\eta_\Sigma] \in H^2(M;\mathbb{R})$ of $\Sigma \looparrowright M$. Equip M with an arbitrary Riemannian metric g; since the corresponding Hodge operator $*$ is an isomorphism on $\wedge^2 T^*M$ we can take a 2-form φ_Σ such that $\eta_\Sigma = *\varphi_\Sigma$. Then via $\int_\Sigma \omega = \int_M \omega\wedge\overline{\eta}_\Sigma = \int_M \omega\wedge*\overline{\varphi}_\Sigma = (\omega,\varphi_\Sigma)_{L^2(M,g)}$ the functional can be re-expressed as $F_{\Sigma,\omega}(A^*A) = (A^*A\omega,\varphi_\Sigma)_{L^2(M,g)}$ in terms of the corresponding definite sesquilinear L^2-scalar product on (M,g). Let $\Sigma' \looparrowright M$ be another closed surface and ω' another closed 2-form such that $F_{\Sigma',\omega'}(1) = 1$. Altering ω and ω' along Σ and Σ' respectively if necessary (which has no effect on $I(M)$ as we have seen) we can pick some compactly supported 2-form Ω on M such that $\Omega|_\Sigma = \omega$ and $\Omega|_{\Sigma'} = \omega'$. Moreover as before take a representative $*\varphi_{\Sigma'} \in \Omega^2(M;\mathbb{R})$ for the Poincaré-dual of Σ'. We can use again $T \in \mathrm{Aut}(\Omega^2_c(M;\mathbb{C}))$ satisfying $\varphi_{\Sigma'} = T\varphi_\Sigma$. Thus

$$F_{\Sigma',\omega'}(A^*A) = (A^*A\Omega,\varphi_{\Sigma'})_{L^2(M,g)} = (A^*A\Omega,T\varphi_\Sigma)_{L^2(M,g)} = (T^*A^*A\Omega,\varphi_\Sigma)_{L^2(M,g)} = F_{\Sigma,\omega}(T^*A^*A)$$

together with $|F_{\Sigma,\omega}(T^*A^*A)|^2 \leqq F_{\Sigma,\omega}\big((AT)^*AT\big)F_{\Sigma,\omega}(A^*A)$ demonstrate that $I_{\Sigma',\omega'}(M) \subseteqq I_{\Sigma,\omega}(M)$. In the same fashion $F_{\Sigma,\omega}(A^*A) = F_{\Sigma',\omega'}((T^{-1})^*A^*A)$ together with the corresponding inequality convinces us that $I_{\Sigma',\omega'}(M) \supseteqq I_{\Sigma,\omega}(M)$. Thus $I_{\Sigma',\omega'}(M) = I_{\Sigma,\omega}(M)$.

Secondly if (Σ,ω) is such that $F_{\Sigma,\omega}(1) = 0$ then by definition $I_{\Sigma,\omega}(M) = \mathfrak{R}$. Therefore if (Σ',ω') is another pair with $F_{\Sigma',\omega'}(1) = 0$ then $I_{\Sigma,\omega}(M) = I_{\Sigma',\omega'}(M)$ (and equal to $\mathfrak{R}$). We are now convinced that it is correct to write $I_{\Sigma,\omega}(M)$ as $I(M)$. In summary it satisfies $\{0\} \subsetneqq I(M) \subseteqq \mathfrak{R}$.

Let us proceed further by exploiting now the observation made during the construction of $\mathfrak{R}$ that it acts on a Hilbert space $\mathscr{H}$ with scalar product $(\cdot,\cdot)$ by the representation π i.e., multiplication from the left. In fact, since $(\hat{A},\hat{B}) = \tau(AB^*)$ we see that π is the standard representation. Consider the space $\{0\} \subsetneqq I(M)I(M)^* \subseteqq \mathfrak{R}$ consisting of all finite sums $A_1B_1 + \cdots + A_kB_k \in \mathfrak{R}$ where $A_i \in I(M)$ and similarly $B_j \in I(M)^*$. It gives rise to a closed linear subspace $\{0\} \subsetneqq \mathscr{I}(M) \subseteqq \mathscr{H}$ by taking the closure of $C(M)\cap I(M)I(M)^{\circledast}$ within $\mathscr{H}\supset C(M)$. Therefore $\mathscr{I}(M)$ is a well-defined closed subspace of $\mathscr{H}$ which is non-trivial if $F_{\Sigma,\omega}(1)\neq 0$ and coincides with $\mathscr{H}$ whenever $F_{\Sigma,\omega}(1)=0$. Take its orthogonal complementum $\{0\}\subseteqq\mathscr{I}(M)^{\perp}\subsetneqq\mathscr{H}$. Note that $\mathscr{I}(M)^{\perp}$ is isomorphic to $\mathscr{H}/\mathscr{I}(M)$. Taking into account that the subset $I(M)$ is a multiplicative left-ideal $\pi:\mathfrak{R}\to\mathfrak{B}(\mathscr{H})$ given by left-multiplication restricts to $\mathscr{I}(M)$ but even more, since the scalar product on $\mathscr{H}$ satisfies the identity $(\widehat{AB},\widehat{C})=(\hat{B},\widehat{A^{\circledast}C})$ the standard representation restricts to $\mathscr{I}(M)^{\perp}$ as well. This latter representation is either a unique non-trivial representation if $\mathscr{I}(M)^{\perp}\neq\{0\}$ (provided by a functional over M with $F_{\Sigma,\omega}(1)\neq 0$ if exists), or the trivial one if $\mathscr{I}(M)^{\perp}=\{0\}$ (provided by a functional with $F_{\Sigma,\omega}(1)=0$ which always exists). Keeping these in mind, for a given M we define

$$\rho_M:\mathfrak{R}\to\mathfrak{B}(\mathscr{I}(M)^{\perp})\ \text{ to be }\ \begin{cases}\pi|_{\mathscr{I}(M)^{\perp}}\text{ on }\mathscr{I}(M)^{\perp}\neq\{0\}\text{ if possible,}\\ \pi|_{\mathscr{I}(M)^{\perp}}\text{ on }\mathscr{I}(M)^{\perp}=\{0\}\text{ otherwise.}\end{cases}$$

The choice is unambigously determined by the topology of M (see the *Remark* below).

From the general theory [1, Chapter 8] we know that if $P_M:\mathscr{H}\to\mathscr{I}(M)^{\perp}$ is the orthogonal projection then $P_M\in\mathfrak{R}$ because $\mathscr{H}$ is the standard $\mathfrak{R}$-module. The Murray–von Neumann coupling constant of ρ_M depends only on the unitary equivalence class of ρ_M and is equal to $\tau(P_M)\in[0,1]$. However observing that ρ_M is surely not isomorphic to π since $\mathscr{I}(M)$ is never trivial the case $\tau(P_M)=1$ is excluded i.e., $\tau(P_M)\in[0,1)$. Let $\Phi:M\to M$ be an orientation-preserving diffeomorphism. It induces an inner automorphism $A\mapsto\Phi^*A(\Phi^*)^{-1}$ of $\mathfrak{R}$. Taking into account that the scalar product on $\mathscr{H}$ is induced by the trace which is invariant under cyclic permutations this inner automorphism is unitary. Moreover it transforms $I_{\Sigma,\omega}(M)$ into $I_{\Sigma',\omega'}(M)=I_{\Phi(\Sigma),\Phi^*\omega}(M)$ hence $F_{\Sigma,\omega}(1)=0$ if and only if $F_{\Sigma',\omega'}(1)=0$ consequently the Hilbert space $\mathscr{I}(M)^{\perp}$ is invariant under Φ. Thus Φ transforms ρ_M into a new representation $\Phi^*\rho_M(\Phi^*)^{-1}$ on $\mathscr{I}(M)^{\perp}$ which is unitary equivalent to ρ_M.

We conclude that $\gamma(M):=\tau(P_M)\in[0,1)$ is a smooth invariant of M as stated. □

Remark. Note that $\gamma(M)=0$ corresponds to the situation when ρ_M is the trivial representation on $\mathscr{I}(M)^{\perp}=\{0\}$. To avoid this we have to demand $F_{\Sigma,\omega}(1)\neq 0$ which by the closedness assumptions on $\Sigma\looparrowright M$ and $\omega\in\Omega^2_c(M;\mathbb{C})$ is in fact a *topological condition*: it is equivalent that

$$\frac{1}{2\pi\sqrt{-1}}F_{\Sigma,\omega}(1)=\frac{1}{2\pi\sqrt{-1}}\int\limits_{\Sigma}\omega=\langle[\Sigma],[\omega]\rangle\in\mathbb{C}$$

as a pairing of $[\Sigma]\in H_2(M;\mathbb{Z})$ and $[\omega]\in H^2(M;\mathbb{C})$ in homology is not trivial. Hence $\gamma(M)=0$ iff $H_2(M;\mathbb{C})=H_2(M;\mathbb{Z})\otimes\mathbb{C}=\{0\}$ (or equivalently, $H^2(M;\mathbb{C})=H^2(M;\mathbb{Z})\otimes\mathbb{C}=\{0\}$). Thus unfortunately $\gamma(M)=0$ for all acyclic or aspherical manifolds (including homology 4-spheres). Examples in the simply connected case are $M=S^4,\mathbb{R}^4,R^4$ (this latter is any exotic or fake $\mathbb{R}^4$) while $M=S^3\times S^1$ is a non-simply connected example.

Jones' subfactor theory (for a summary cf. e.g. [1, Section 9.4] or [8, Chapter V.10]) imposes an interesting restriction on the possible spectrum of γ just constructed.

Lemma 2.2. *Let M be a connected oriented smooth 4-manifold and $\gamma(M)\in[0,1)$ its smooth invariant. Then $\gamma(M)=1-\frac{1}{x}$ where $x\in\{4\cos^2\left(\frac{\pi}{n}\right)\mid n\geqq 3\}\cup[4,+\infty)$ that is, an element from the set of all possible Jones' subfactor indices.*

Proof. Taking into account that $\{0\}\subsetneqq I(M)I(M)^{\circledast}\subseteqq\mathfrak{R}$ is self-adjoint it generates a von Neumann subalgebra; more precisely put

$$\mathfrak{I}(M):=(\pi(C(M)\cap I(M)I(M)^{\circledast}))''\subset\mathfrak{B}(\mathscr{H})\ .$$

Since the set $C(M)\cap I(M)I(M)^{*}$ is already weakly dense in $I(M)I(M)^{*}$ which is a non-trivial left-ideal of $\mathfrak{R}$ its only defect to be a von Neumann algebra is that $C(M)\cap I(M)I(M)^{*}$ is not weakly closed and possesses no unit; therefore its weak closure $\mathfrak{I}(M)$ is a factor. Thus it is a subfactor of $\mathfrak{R}$ therefore it possesses a corresponding Jones index denoted as usual by $[\mathfrak{R}:\mathfrak{I}(M)]\in\{4\cos^2\left(\frac{\pi}{n}\right)\,|\,n\geqq 3\}\cup[4,+\infty)$.

The left-ideal $\{0\}\subsetneqq C(M)\cap I(M)I(M)^{*}\subseteqq C(M)$ acts on itself by letf-multiplications. Since both the von Neumann algebra $\mathbb{C}\cong Z(\mathfrak{R})\subsetneqq\mathfrak{I}(M)\subseteqq\mathfrak{R}$ and the Hilbert space $\{0\}\subsetneqq\mathscr{I}(M)\subseteqq\mathscr{H}$ arise as appropriate closures of this same left-ideal its left action on itself extends to a representation of $\mathfrak{I}(M)$ on $\mathscr{I}(M)$ which, taking into account that $\mathscr{I}(M)\neq\{0\}$, is equivalent to the standard representation of $\mathfrak{I}(M)$. We also know already that $\mathscr{H}=\mathscr{I}(M)\oplus\mathscr{I}(M)^{\perp}$ as $\mathfrak{R}$-modules by Lemma 2.1. Recalling the basic properties of the dimension function of a left von Neumann algebra module over the von Neumann algebra itself (cf. e.g. [1, Chapter 8]) we collect:

$$\begin{aligned}
\dim_{\mathfrak{R}}\mathscr{H} &= \dim_{\mathfrak{R}}\mathscr{I}(M)+\dim_{\mathfrak{R}}\mathscr{I}(M)^{\perp}\quad\text{(additivity)}\\
\dim_{\mathfrak{R}}\mathscr{H} &= 1\quad\text{(the standard representation of }\mathfrak{R})\\
\dim_{\mathfrak{R}}\mathscr{I}(M)^{\perp} &= \gamma(M)\quad\text{(by Lemma 2.1 and Footnote 3)}\\
\dim_{\mathfrak{I}(M)}\mathscr{I}(M) &= [\mathfrak{R}:\mathfrak{I}(M)]\dim_{\mathfrak{R}}\mathscr{I}(M)\quad\text{(dimension comparison)}\\
\dim_{\mathfrak{I}(M)}\mathscr{I}(M) &= 1\quad\text{(the standard representation of }\mathfrak{I}(M))
\end{aligned}$$

from which it follows that $\gamma(M)=1-\frac{1}{[\mathfrak{R}:\mathfrak{I}(M)]}$ yielding the desired result. $\square$

Proof of Theorem 1.1. This theorem follows from the general construction of $\mathfrak{R}$ as we outlined here and in particular from Lemmata 2.1 and 2.2. $\square$

Next we collect some basic useful properties of the invariant.

Lemma 2.3. *(Reversing orientation.) If M is a connected oriented smooth 4-manifold and $\overline{M}$ is its orientation-reversed form then $\gamma(\overline{M})=\gamma(M)$.*

(Excision principle.) Let M be a connected oriented smooth 4-manifold and $\emptyset\subseteqq Y\subset M$ a submanifold so that $M\setminus Y\subseteqq M$ is connected and the embedding $i:M\setminus Y\to M$ induces an isomorphism $i_*:H_2(M\setminus Y;\mathbb{Z})\to H_2(M;\mathbb{Z})$ on the $2^{\rm nd}$ homology. Then $M\setminus Y$ with induced orientation and smooth structure is a connected oriented smooth 4-manifold satisfying $\gamma(M\setminus Y)=\gamma(M)$.

(Gluing principle.) Let M and N be two connected, oriented smooth 4-manifolds and write $M\#N$ for their connected sum. With induced orientation $M\#N$ is a connected, oriented smooth 4-manifold. Its smooth invariant satisfies

$$\gamma(M\#N)=\frac{\gamma(M)+\gamma(N)}{1+\gamma(M)\gamma(N)}\ .$$

Proof. The first assertion is obvious from γ's construction carried out in the proof of Lemma 2.1.

Regarding the second assertion $M\setminus Y$ is a connected oriented smooth 4-manifold by assumption consequently admits an associated von Neumann algebra $\mathfrak{R}(M\setminus Y)$ which is also a hyperfinite II_1 factor. The space of compactly supported smooth forms on $M\setminus Y$ can be identified with the linear subspace consisting of compactly supported smooth forms on M which vanish along $Y\subset M$ that is, $\Omega^2_c(M\setminus Y;\mathbb{C})\cong\Omega^2_c(M,Y;\mathbb{C})\subset\Omega^2_c(M;\mathbb{C})$. By the aid of this embedding it is easy to see that in the L^2-topology induced by any Riemannian metric $\Omega^2_c(M\setminus Y;\mathbb{C})$ is dense in $\Omega^2_c(M;\mathbb{C})$ which immediately implies that $C(M\setminus Y)=C(M)$ for the corresponding complexified Clifford algebras; thus taking closures we come up with $\mathscr{H}(M\setminus Y)=\mathscr{H}(M)$ and then $\mathfrak{R}(M\setminus Y)=\mathfrak{R}(M)$. By definition $\mathfrak{R}(M\setminus Y)$

acts via $\rho_{M\setminus Y}$ on $\mathscr{I}(M\setminus Y)^{\perp}$ and $\mathfrak{R}(M)$ acts via ρ_M on $\mathscr{I}(M)^{\perp}$. By assumption $M\setminus Y \subsetneqq M$ induces an isomorphism on the 2nd homology hence $\mathscr{I}(M\setminus Y)^{\perp}$ and $\mathscr{I}(M)^{\perp}$ are simultaneously trivial or not. In fact $\mathscr{I}(M\setminus Y)^{\perp}=\mathscr{I}(M)^{\perp}$. This is because recall that the condition $A\in I(M\setminus Y)$ can be written as $F_{\Sigma,\varphi_\Sigma}(A^*A)=(A^*A\varphi_\Sigma,\varphi_\Sigma)_{L^2(M\setminus Y,g)}=0$ where φ_Σ is the Hodge-dual of the harmonic representative $\eta_\Sigma$of the Poincaré-dual of $\Sigma \looparrowright M\setminus Y$; since Y has zero 4 dimensional Lebesgue measure $A\in I(M\setminus Y)$ if and only if $A\in I(M)$; therefore $\mathscr{I}(M\setminus Y)=\mathscr{I}(M)$ yielding the same for the complementums in $\mathscr{H}(M\setminus Y)=\mathscr{H}(M)$. Consequently both $\rho_{M\setminus Y}$ and ρ_M are representations of $\mathfrak{R}(M\setminus Y)=\mathfrak{R}(M)$ on $\mathscr{I}(M\setminus Y)^{\perp}=\mathscr{I}(M)^{\perp}$. By definition $\gamma(M\setminus Y)=\dim_{\mathfrak{R}(M\setminus Y)}\mathscr{I}(M\setminus Y)^{\perp}$ and likewise $\gamma(M)=\dim_{\mathfrak{R}(M)}\mathscr{I}(M)^{\perp}$. Therefore $\gamma(M\setminus Y)=\gamma(M)$ as stated.

Concerning the third assertion note that the γ-invariant is a well-defined map from (the category) Man^4 of all orientation-preserving diffeomorphism classes of connected, oriented smooth 4-manifolds into the real interval $[0,1)\subset\mathbb{R}$. But Man^4 forms a commutative semigroup with unit S^4 under the connected sum operation #. That is, if $X,Y,Z\in\mathsf{Man}^4$ and $S^4\in\mathsf{Man}^4$ is the 4-sphere then $X\#Y\cong Y\#X$ and $(X\#Y)\#Z\cong X\#(Y\#Z)$ and $X\#S^4\cong X$. Pick $M,N\in\mathsf{Man}^4$ with their connected sum $M\#N\in\mathsf{Man}^4$ and consider the corresponding invariants $\gamma(M),\gamma(N),\gamma(M\#N)\in[0,1)$. Define $\bullet:[0,1)\times[0,1)\to[0,1)$ by setting $\gamma(M)\bullet\gamma(N):=\gamma(M\#N)$. The $\bullet$-operation is therefore well-defined and has the properties $\gamma(X)\bullet\gamma(Y)=\gamma(Y)\bullet\gamma(X)$ and $(\gamma(X)\bullet\gamma(Y))\bullet\gamma(Z)=\gamma(X)\bullet(\gamma(Y)\bullet\gamma(Z))$ and $\gamma(X)\bullet\gamma(S^4)=\gamma(X)$. These ensure us that $([0,1),\bullet)$ is a unital commutative semigroup and $\gamma:(\mathsf{Man}^4,\#)\to([0,1),\bullet)$ is a unital semigroup homomorphism. Mapping the unique semigroup structure of $[0,+\infty)\subset\mathbb{R}$ to $[0,1)$ by $\alpha\mapsto\tanh\alpha$ gives $x\bullet y=\frac{x+y}{1+xy}$ with 0 being the unit. Thus $\gamma(S^4)=0$ (as we already know) but otherwise γ is not trivial because $\gamma(M)>0$ if $b_2(M)>0$. This yields the shape for $\gamma(M)\bullet\gamma(N)$ as stated. $\square$

Proof of Theorem 1.2. This theorem follows from Lemma 2.3. $\square$

More generally one can demonstrate by techniques from 4-manifold theory and the gluing principle the following non-trivial but quite insensitive behaviour of γ in the closed simply connected case.

Lemma 2.4. *If M' and M'' are connected, simply connected, closed smooth 4-manifolds which are homeomorphic then $\gamma(M')=\gamma(M'')$.*

Take $y\in[0,1)$ and the sequence

$$R_0(y):=0\,,\,R_1(y):=y\,,\,\dots\,,\,R_k(y):=\frac{y+R_{k-1}(y)}{1+yR_{k-1}(y)}\,,\,\dots$$

for all $k=0,1,2,\dots$ representing the semigroup $\{0\}\cup\mathbb{N}$ inside $[0,1)$. Then for any connected, simply connected, closed smooth 4-manifold

$$\gamma(M)=R_{b_2(M)}\left(\tfrac{8}{9}\right)=\frac{17^{b_2(M)}-1}{17^{b_2(M)}+1}\,.$$

Proof. Concerning the first assertion if M' and M'' are as required then there exists an integer $k\geqq 0$ such that $M'\#k(S^2\times S^2)\cong M''\#k(S^2\times S^2)$ (cf. e.g. [15, Theorem 9.1.12]). Thus we know that we have the equality $\gamma(M'\#k(S^2\times S^2))=\gamma(M''\#k(S^2\times S^2))$. Then applying the gluing principle

$$\frac{\gamma(M')+\gamma(k(S^2\times S^2))}{1+\gamma(M')\gamma(k(S^2\times S^2))}=\frac{\gamma(M'')+\gamma(k(S^2\times S^2))}{1+\gamma(M'')\gamma(k(S^2\times S^2))}$$

and observing that the map $x\mapsto\frac{x+x_0}{1+xx_0}$ is invertible we obtain that $\gamma(M')=\gamma(M'')$.

Concerning the second assertion, if M_1 and M_2 are connected, closed, simply connected, smooth then there exist integers $k_1, l_1 \geqq 0$ and $k_2, l_2 \geqq 0$ such that $M_1 \# k_1 \mathbb{C}P^2 \# l_1 \overline{\mathbb{C}P^2} \cong M_2 \# k_2 \mathbb{C}P^2 \# l_2 \overline{\mathbb{C}P^2}$ (cf. e.g. [15, Theorem 9.1.14]) which shows that $\gamma(M_1 \# k_1 \mathbb{C}P^2 \# l_1 \overline{\mathbb{C}P^2}) = \gamma(M_2 \# k_2 \mathbb{C}P^2 \# l_2 \overline{\mathbb{C}P^2})$. Now put $y := \gamma(\mathbb{C}P^2) = \gamma(\overline{\mathbb{C}P}^2)$. Then by the gluing principle

$$\frac{\gamma(M_1) + R_{k_1+l_1}(y)}{1 + \gamma(M_1) R_{k_1+l_1}(y)} = \frac{\gamma(M_2) + R_{k_2+l_2}(y)}{1 + \gamma(M_2) R_{k_2+l_2}(y)} \ .$$

Let $M_1 := M$ be arbitrary and $M_2 := S^4$ hence $\gamma(M_2) = 0$. Then we can suppose that $k_1 + l_1 \leqq k_2 + l_2$ therefore

$$\frac{\gamma(M) + R_{k_1+l_1}(y)}{1 + \gamma(M) R_{k_1+l_1}(y)} = R_{k_2+l_2}(y)$$

from which again by invertability we find $\gamma(M) = R_{k_2+l_2-k_1-l_1}(y)$ hence setting $n := k_2 + l_2 - k_1 - l_1 \geqq 0$ we obtain $\gamma(M) = R_n(y)$. Moreover it is clear from the proof that in fact $n = b_2(M)$.

A closed formula arises if we write $R_k(y) = \frac{a_k(y)}{b_k(y)}$ and observe that

$$\begin{pmatrix} a_k(y) \\ b_k(y) \end{pmatrix} = \begin{pmatrix} 1 & y \\ y & 1 \end{pmatrix}^k \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ .$$

The eigenvalues of $\begin{pmatrix} 1 & y \\ y & 1 \end{pmatrix}$ are equal to $1 \pm y$ with eigenvectors $\begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}$ respectively. Thus inserting the decomposition $\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $k = b_2(M)$ into the above equation we end up with

$$\gamma(M) = R_{b_2(M)}(y) = \frac{(1+y)^{b_2(M)} - (1-y)^{b_2(M)}}{(1+y)^{b_2(M)} + (1-y)^{b_2(M)}} \ .$$

Our last task is to find the precise value of $y \in [0,1)$. Endowing $\mathbb{C}P^2$ with the orientation compatible with its complex structure and putting the Fubini–Study metric the group $\mathrm{U}(3)$ acts by isometries on $\mathbb{C}P^2$. Then the underlying orientation-preserving diffeomorphisms induce an action of the complexification $\mathrm{U}(3)^{\mathbb{C}} = \mathrm{GL}(3;\mathbb{C})$ on $\mathfrak{R}$ by unitary inner automorphisms. Consequently, on the one hand, there exists an invariant subfactor $\mathfrak{J}(\mathbb{C}P^2) \subset \mathfrak{R}$ such that $\mathfrak{R} = \mathfrak{M}_3(\mathfrak{J}(\mathbb{C}P^2))$. In the proof of Lemma 2.1 we have seen that by exploiting the flexibility of $F_{\Sigma,\omega}$ the condition $A \in I(\mathbb{C}P^2)$ can be written as $F_{\Sigma,\varphi_\Sigma}(A^*A) = (A^*A\varphi_\Sigma, \varphi_\Sigma)_{L^2(M,g)} = 0$ where φ_Σ is the Hodge-dual of the harmonic representative η_Σ of the Poincaré-dual of $\Sigma \looparrowright M$ i.e., $I(\mathbb{C}P^2)$ can be defined in terms of harmonic 2-forms only. However the isometries of $\mathbb{C}P^2$ leave the space of harmonic 2-forms unchanged consequently, on the other hand, the subfactor $\mathfrak{I}(\mathbb{C}P^2) \subset \mathfrak{R}$ generated by $I(\mathbb{C}P^2)I(\mathbb{C}P^2)^* \subset \mathfrak{R}$ as in Lemma 2.2 is invariant under $\mathrm{GL}(3;\mathbb{C})$, too. Therefore $\mathfrak{J}(\mathbb{C}P^2) = \mathfrak{I}(\mathbb{C}P^2)$ implying for the Jones index in question that

$$[\mathfrak{R} : \mathfrak{I}(\mathbb{C}P^2)] = [\mathfrak{M}_3(\mathfrak{J}(\mathbb{C}P^2)) : \mathfrak{J}(\mathbb{C}P^2) \otimes 1_{\mathrm{GL}(3;\mathbb{C})}] = 3^2 = 9 \ .$$

Thus putting together everyting we get via Lemma 2.2 that $y = R_1(y) = \gamma(\mathbb{C}P^2) = 1 - \frac{1}{9} = \frac{8}{9}$ and the proof is now complete. $\square$

Remark. It follows that $\log(1 - R_k(y)) \approx k \log\frac{1-y}{1+y}$ for k large enough (here "large" depends on y). This yields an interesting estimate for the 2nd Betti number of a closed simply connected manifold in terms of its γ-invariant. On substituting $\gamma(M) = R_{b_2(M)}(\frac{8}{9})$ we get

$$b_2(M) \approx \frac{\log(1 - \gamma(M))}{\log\frac{1}{17}} \quad \text{if } b_2(M) \text{ is large enough .} \tag{7}$$

Proof of Theorem 1.3. This theorem follows from Lemma 2.4. $\square$

*Remark***.** We have seen that $\mathfrak{R}$ is especially interesting in the four dimensional case since it contains curvature tensors. However actually this has not been used during the construction of γ in the proof of Lemma 2.1. Moreover, as recorded in Lemma 2.4, despite its construction depending on the smooth structure, γ is in fact a topological invariant only at least in the simply connected closed case. Nevertheless γ is computable and non-trivial. In order to appreciate γ as it is we close this section by rapidly introducing another huge class of invariants whose members look very natural and meet the demands above however turn out to be trivial.

By the aid of the canonical inclusions $C(M) \subset \mathscr{H}$ and $C(M) \subset \mathfrak{R} \subset \mathfrak{B}(\mathscr{H})$ and the Riesz representation theorem an immense class of functionals on $\mathfrak{R}$ arises by picking any $T \in \mathfrak{B}(\mathscr{H})$ and continuously extending the map $A \mapsto (\hat{T}, \widehat{A^*}) = \tau(TA)$ from $C(M)$ to $\mathfrak{R}$ giving rise to a continuous $\mathbb{C}$-linear map $F_T : \mathfrak{R} \to \mathbb{C}$ of the form

$$F_T(A) := \tau(TA)$$

which is also positive if $\tau(T) > 0$. Interesting examples emerge if instead of a plain manifold M as in Lemma 2.1 we start now with an oriented Riemannian one (M, g). For instance take the Hodge operator $*$ induced by the metric and orientation and put $T := \frac{1}{2}(1 + *)$ hence $TA = A^+$ is the "self-dual part" of A. Or assume that the curvature is bounded hence $R_g \in \mathfrak{R}$ and satisfies $\tau(R_g) \neq 0$ i.e. the metric itself is sufficiently generic[4] and put $T := \pm R_g$. Let $\{0\} \subseteqq I(M) \subsetneqq \mathfrak{R}$ denote the usual multiplicative left-ideal generated by this functional i.e. $A \in I(M)$ if and only if $F_T(A^*A) = 0$. Now if T' is another element satisfying $\tau(T') \neq 0$ then by assumption both $0 \neq \hat{T}, \hat{T}' \in \mathscr{H}$ hence there exists an invertible element $S \in \mathfrak{B}(\mathscr{H})^\times$ such that $T' = ST$; consequently

$$0 \leqq |F_{T'}(A^*A)|^2 = |\tau(STA^*A)|^2 = |\tau(TA^*AS)|^2 = |F_T(A^*AS)|^2 \leqq F_T(A^*A)F_T((AS)^*AS)$$

and likewise $0 \leqq |F_T(A^*A)|^2 \leqq F_{T'}(A^*A)F_{T'}((AS^{-1})^*AS^{-1})$. This demonstrates the independence of $I(M)$. Take again the self-adjoint subspace $I(M)I(M)^*$ and let $\{0\} \subseteqq \mathscr{I}(M) \subsetneqq \mathscr{H}$ be the closure of $C(M) \cap I(M)I(M)^*$ and consider its orthogonal complementum $\{0\} \subsetneqq \mathscr{I}(M)^\perp \subseteqq \mathscr{H}$ and the corresponding projection $P_M : \mathscr{H} \to \mathscr{I}(M)^\perp$. Also take the subfactor $\mathbb{C} \cong Z(\mathfrak{R}) \subseteqq \mathfrak{I}(M) \subsetneqq \mathfrak{R}$ provided by the weak closure of $C(M) \cap I(M)I(M)^*$. Very similarly to the construction in the proof of Lemma 2.1 we observe that $\mathscr{I}(M)^\perp$ is a left $\mathfrak{R}$-module; hence define a new smooth 4-manifold invariant by its coupling constant more precisely

$$\delta(M) := \tau(P_M)\ .$$

The immediate observation is that $\delta(M) \in (0, 1]$. However the analogue of Lemma 2.2 is not valid because $\mathscr{I}(M)$ is not necessarily the standard $\mathfrak{I}(M)$-module anymore as the case of $\mathscr{I}(M) = \{0\}$ and the corresponding $\mathfrak{I}(M) = Z(\mathfrak{R}) \cong \mathbb{C}$ shows. The computationally useful Lemma 2.3 also no longer holds (except the invariance under reversing orientation). In fact since we can suppose that T is self-adjoint and positive definite therefore $T = S^2$ with a same type hence invertible operator $S \in \mathfrak{B}(\mathscr{H})$; thus using the norm $\|\cdot\|$ on $\mathscr{H}$ we find $0 = F_T(A^*A) = \|\widehat{(AS)^*}\|^2$ implying $AS = 0$ i.e. $A = 0$. Thus $I(M) = 0$ yielding $\mathscr{I}(M)^\perp = \mathscr{H}$. Thus the fact is that always $\delta(M) = 1$.

[4]Note that whenever M is compact it can accommodate a metric g for which either $\tau(R_g) \neq 0$ or $\tau(R_g) = 0$ i.e. both cases can occur. This is obvious for flat manifolds admitting a metric with $R_g = 0$ hence $\tau(R_g) = 0$. But note that (6) says that $\tau(R_g)$ is proportional to the total scalar curvature hence $\tau(R_g) = 0$ for all Ricci-flat manifolds too. But even more if M is any compact oriented manifold then take a smooth function $f : M \to \mathbb{R}$ which is (i) strictly negative in one point and (ii) satisfies $\int_M f\mu_h = 0$ with respect to some auxiliary metric h. Then (i) implies by [5, Theorem 4.35] the existence of a Riemannian metric g having f as its scalar curvature hence via (ii) and (6) satisfying $\tau(R_g) = 0$.

3 Physical interpretation

In this section we cannot resist temptation and introduce the basics of a physical interpretation of the material collected so far. Namely, upon reversing the approach of Section 2, we shall replace the immense class of classical space-times of general relativity with a single universal "quantum space-time" allowing us to lay down the foundations of a manifestly four dimensional, covariant, non-perturbative and genuinely quantum theory of gravity. This construction is natural, simple and self-contained. More precisely here in Section 3 not *one particular* 4*-manifold*—physically regarded as a particular classical space-time—but *the unique hyperfinte* II_1 *factor von Neumann algebra*—physically viewed as the universal quantum space-time—is declared to be the primarily given object. Let us see how it works.

Observables, fields, states and the gauge group. Let $\mathscr{H}$ be an abstractly given infinite dimensional complex separable Hilbert space and $\mathfrak{R}\subset\mathfrak{B}(\mathscr{H})$ be a type II_1 hyperfinite factor hence tracial von Neumann algebra acting on $\mathscr{H}$ by the standard representation. We call $\mathfrak{R}$ the *algebra of (bounded) observables*, its tangent space $T_1\mathfrak{R}\supset\mathfrak{R}$ consisting of the Fréchet derivatives of 1-parameter families of observables at the unit $1\in\mathfrak{R}$ the *algebra of fields*, while $\mathscr{H}$ the *state space* in this quantum theory. The subgroup $\mathrm{U}(\mathscr{H})\cap\mathfrak{R}$ of the unitary group of $\mathscr{H}$ operating as inner automorphisms on $\mathfrak{R}$ is the *gauge group*. Note that the gauge group acts on both $\mathfrak{R}$ and $\mathscr{H}$ but in a different way.

What kind of quantum theory is the one in which $\mathfrak{R}$ plays the role of the algebra of *physical observables*? We have seen in Section 2 that $\mathfrak{R}$, when regarded to operate on a differentiable manifold, contains all bounded (complexified) algebraic curvature tensors along it whenever the real dimension of this manifold is precisely *four*. Moreover $\mathfrak{R}$ is generated by local operators more precisely by trace class algebraic operators R acting like *multiplication* operators and by trace class operators L_X acting like *first order differential* operators. Picking canonically conjugate pairs i.e. necessarily unbounded self-adjoint operators $R,L_X\in T_1\mathfrak{R}$ satisfying $[R,L_X]=1$ (assuming $\hbar=1$) shows that $T_1\mathfrak{R}$ contains a CCR algebra hence revealing the *bosonic* character of $\mathfrak{R}$. Recall that in classical general relativity *local* gravitational phenomena are caused by the curvature of space-time. Hence interpreting $\mathfrak{R}$ physically as an operator algebra of local observables the corresponding quantum theory is expected to be a *four dimensional quantum theory of pure gravity* (hence a bosonic theory). In this way we fulfill the *Heisenberg dictum* that a quantum theory should completely and unambigously be formulated and interpreted in terms of its local physical observables (and not the other way round). In modern understanding by a *physical theory* one means a two-level description of a bundle of phenomenologically interrelated natural phenomena: the theory possesses a *syntax* provided by its mathematical core structure and a *semantics* which is the meaning i.e. interpretation of the bare mathematical model in terms of physical concepts. In this context our quantum theory is not merely a mathematical theory anymore but a physical theory. This is because the bare mathematical structure $\mathfrak{R}$ (together with a representation π on $\mathscr{H}$) is dressed up i.e., interpreted by assigning a physical (in fact, gravitational) meaning to the experiments consistently performable by the aid of this structure (i.e., the usual quantum measurements of operators $A\in\mathfrak{R}$ in pure states $v\in\mathscr{H}$ or in more general ones, see below). In our opinion it is of particular interest that the geometrical dimension—equal to four—is fixed at the semantical level only and it matches the known phenomenological dimension of space-time. This is in sharp contrast to e.g. string theory where the geometrical dimension of the theory is fixed already at its syntactical level i.e., by its mathematical structure (namely, demaninding the theory to be free of conformal anomaly) and it turns out to be much higher than the phenomenological dimension of space-time. As a further clarification it is emphasized that the aim here is *not* to quantize general relativity in some way but rather conversely: being a successful theory at the classical level, general relativity should be derived from this abstractly given quantum theory by taking an appropriate "classical limit" as naturally as possible. Unfortunately, this program yet cannot be carried out completely here.

Observables as the universal space of all space-times. Taking into account the embeddings (3) and the universality of $\mathfrak{R}$ the *unique* algebra of observables $\mathfrak{R}$ can be considered as the collection of *all* classical space-times and we can interpret the appearance of the gauge group as the manifestation of the diffeomorphism gauge symmetry of classical general relativity in this quantum theory. Indeed, orientation-preserving diffeomorphisms of M interchange its points as well as act on the corresponding operators in $\mathfrak{R}$ (projections) by unitary inner automorphisms i.e. $\mathrm{Diff}^+(M)$ embeds into $\mathrm{Aut}\,\mathfrak{R}$. Reformulating this in a more geometric language we can say that classical space-times appear as special orbits within $\mathfrak{R}$ of its gauge group. The operators in $\mathfrak{R}$ representing geometric points via (3) are special operators namely projections. Consequently the full non-commutative algebra $\mathfrak{R}$ is not exhausted by operators representing points of space-time; it certainly contains much more operators—e.g. various projections which are not of geometric origin—therefore this "universal quantum space-time" is more than a bunch of all classical space-times. As a comparison, in *algebraic quantum field theory* [17] one starts with a particular smooth 4-manifold M and considers an assignment $U \mapsto \mathfrak{R}(U)$ describing local algebras of observables along all open subsets $\emptyset \subseteqq U \subseteqq M$. However in our case, quite conversely, space-times are secondary structures only and all of them are injected into the unique observable algebra $\mathfrak{R}$ which is considered to be primary.

More examples of observables, fields and states. Let us take a closer look of the elements of $\mathfrak{R}$ and of $T_1\mathfrak{R} \supset \mathfrak{R}$. Taking into account the items above concerning the interpretation of the mathematical results of Section 2 we have agreed to identify some elements of $\mathfrak{R}$ up to finite accuracy with four dimensional bounded algebraic (i.e., formal) curvature tensors of *all possible* smooth 4-manifolds. In fact one can more accurately reformulate some basic concepts of classical general relativity within this quantum (field-)theoretic framework. The only thing we have to do is to refine or improve the constructions in Section 2 by taking into account not only the non-degeneracy but the indefiniteness of the pairing (1) too.

Remember that given a connected oriented Riemannian 4-manifold (M,g) one can introduce the Hodge operator $*$ which satisfies $*^2 = \mathrm{Id}_{\Omega^2_c(M)}$ hence it induces a splitting

$$\Omega^2_c(M) = \Omega^+_c(M) \oplus \Omega^-_c(M)$$

into (anti)self-dual or ± 1-eigenspaces. Thus the Hodge star as an operator on $\Omega^2_c(M)$ with respect to this splitting has the form $* = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ hence is symmetric; likewise the curvature tensor R_g of (M,g) looks like (5) i.e., is also a pointwise symmetric operator. We summarize these facts by writing that $*, R_g \in C^\infty(M; S^2 \wedge^2 T^*M) \subset C^\infty(M; \mathrm{End}(\wedge^2 T^*M))$. In addition R_g obeys the algebraic Bianchi identity $b(R_g) = 0$. The vacuum Einstein equation with cosmological constant $\Lambda \in \mathbb{R}$ reads as $\mathrm{Ric} = \Lambda g$ therefore it is equivalent to the vanishing of the traceless Ricci part Ric_0 of R_g in (5). From these it readily follows, as noticed in [24], that (M,g) is Einstein i.e., satisfies the vacuum Einstein equation with cosmological constant if and only if $*R_g = R_g*$ or equivalently $*R_g* = R_g$.

Now if R_g is a bounded operator then we have seen in Section 2 that its complex-linear extension gives rise to an element $R_g \in \mathfrak{R}$ which is moreover self-adjoint i.e., $R_g^{\circledast} = R_g$ because R_g is symmetric. Therefore its trace is always real. Likewise for $*$ satisfying $\tau(*) = 0$. The algebraic Bianchi identity has the form $b(R_g) = 0$ where b is a certain fiberwise averaging operator on symmetric bundle morphisms satisfying $b^2 = b$ and $b^* = b$ with respect to the metric. This gives rise to a g-orthogonal decomposition $S^2 \wedge^2 T^*_xM \otimes \mathbb{C} = \ker b_x \oplus \mathrm{im}\, b_x$ at every point $x \in M$. A further subtlety of *four dimensions* is that $\mathrm{im}\, b_x = \mathbb{C} * \mathrm{Id}_{\wedge^2 T^*_xM\otimes\mathbb{C}} \subset S^2 \wedge^2 T^*_xM \otimes \mathbb{C}$ holds [24]. Consequently $b(R_g) = 0$ is equivalent to saying that $g(R_g, *) = 0$ i.e., within $S^2 \wedge^2 T^*M \otimes \mathbb{C}$ the operator R_g is g-orthogonal to $*$ at every $x \in M$. By exploiting that $*$ is a pointwise symmetric operator we can use the pointwise equality $g(R_g, *) = \mathrm{tr}(R_g*)$

and then (6) to obtain

$$\tau(R_g*) = \frac{1}{6}\int\limits_M \mathrm{tr}(R_g*)\mu_g = 0$$

which geometrically means that $(\hat{R}_g, \hat{*}) = 0$ i.e., $\hat{R}_g$ is perpendicular to $\hat{*}$ within $\mathscr{H}$. Hence the Bianchi identity. Concerning the Einstein condition if R_g comes from an Einstein metric with cosmological constant $\Lambda \in \mathbb{R}$ then we already know that $*R_g* = R_g$ which means that $\mathrm{Ric}_0 = 0$ in (5). Moreover taking into account that $\mathrm{Weyl}^{\pm} \in C^\infty(M; S^2\wedge^2 T^*M \otimes \mathbb{C})$ in (5) are traceless their pointwise scalar products with the identity of $C^\infty(M; S^2\wedge^2 T^*M\otimes\mathbb{C})$ look like $g(\mathrm{Weyl}^{\pm}, 1) = \mathrm{tr}(\mathrm{Weyl}^{\pm}) = 0$. Using (6) this shows again that the corresponding vectors $\widehat{\mathrm{Weyl}}^{\pm} \in \mathscr{H}$ are perpendicular to $\hat{1} \in \mathscr{H}$ i.e., $(\widehat{\mathrm{Weyl}}^{\pm}, \hat{1}) = 0$. Moreover $\mathrm{Scal} = 4\Lambda$ consequently via (5) we get

$$\tau(R_g) = (\hat{R}_g\,, \hat{1}) = \frac{1}{12}(\widehat{\mathrm{Scal}}, \hat{1}) = \frac{4\Lambda}{12}(\hat{1}, \hat{1}) = \frac{\Lambda}{3}\;. \tag{8}$$

Finally note that in the previous analysis the signature of the metric plays no significant role because of an involved complexification (we have used Riemannian signature just for convenience).

All of these serve as a motivation for the following operator-algebraic reformulation or generalization of the classical vacuum Einstein equation with cosmological constant:

Definition 3.1. *Let M be a connected oriented smooth* 4*-manifold and $\mathfrak{R}$ its* II_1*-type hyperfinite factor von Neumann algebra as before.*

(i) A refinement of $\mathfrak{R}$ *is a pair $(\mathfrak{R}, *)$ where $1 \neq * \in \mathfrak{R}$ and satisfies $*^2 = 1$;*

(ii) An operator $Q \in \mathfrak{R}$ solves the quantum vacuum Einstein equation *with respect to $(\mathfrak{R}, *)$ if*

$$\begin{cases} Q^{\circledast} & = Q \quad \textit{(self-adjointness)} \\ \tau(Q*) & = 0 \quad \textit{(algebraic Bianchi identity)} \\ *Q* & = Q \quad \textit{(Einstein condition);} \end{cases}$$

(iii) The trace $\tau(Q) =: \frac{\Lambda}{3} \in \mathbb{R}$ is called the corresponding quantum cosmological constant*;*

(iv) $Q \in \mathfrak{R}$ as above is called a vacuum state *if it is positive semi-definite too (hence its corresponding quantum cosmological constant is non-negative).*

Note that the curvature tensor R_g of an Einstein manifold (M, g) if bounded as an operator always solves the quantum vacuum Einstein equation with respect to the canonical refinement provided by metric (anti)self-duality. However in sharp contrast to the classical Einstein equation whose solution is a metric therefore is non-linear, its quantum generalization is linear hence easily solvable. Of course this is beacuse in the generalization we do not demand the solution (which can be any non-local operator) to originate from a metric. Indeed, given $B \in \mathfrak{R}$ then $S = \frac{1}{2}(B + B^{\circledast})$ is self-adjoint and then picking an arbitrary refinement $(\mathfrak{R}, *)$ and taking into account that $*$ is always self-adjoint, the averaged operator $Q := \frac{1}{2}\big(S + *S*\big) - \tau(S*)* \in \mathfrak{R}$ is automatically a solution of the quantum vacuum Einstein equation; moreover using $\tau(*) = 0$ its trace $\frac{\Lambda}{3}$ is equal to $\tau(B)$ hence is independent of the particular refinement. Observe that this linearity permits a BRST-like reformulation of Definition 3.1 too i.e. introducing a linear operator $\Delta : \mathfrak{R} \to \mathfrak{R}$ such that $\ker\Delta$ is precisely equal to the solution space and $\Delta^2 = 0$.

It is important to note that, contrary to smooth solutions, many singular solutions of classical general relativity theory cannot be interpreted as observables because their curvatures lack being bounded operators hence do not belong to $\mathfrak{R}$. As a result, we expect that the classical Schwarzschild or Kerr black hole solutions and more generally gravitatioal fields of isolated bodies (cf. [10, 11]) give rise not to *observables* in $\mathfrak{R}$ but rather *fields* in $T_1\mathfrak{R} \supset \mathfrak{R}$.

Questions and answers. First let us clarify what the answers in this proposed quantum theory are because this is easier. Staying safely within the orthodox framework i.e., the *Copenhagen interpretation* and the standard mathematical formulation of quantum theory (but relaxing this latter somewhat), given an observable represented by $A \in \mathfrak{R}$ and a general (i.e., not necessarily pure) state also represented by an element $B \in \mathfrak{R}$ in the observable algebra (regarded as a "density matrix" operator over the state space $\mathscr{H}$) we declare that an *answer* is formally like

The expectation value of the observable quantity A in the state B is $\tau(AB) \in \mathbb{C}$

where $\tau : \mathfrak{R} \to \mathbb{C}$ is the unique finite trace on the hyperfinite II_1 factor von Neumann algebra $\mathfrak{R}$. Note that in order not to be short sighted, at this level of generality we require neither the observable $A \in \mathfrak{R}$ to be self-adjoint nor the state $B \in \mathfrak{R}$ to be positive semi-definite self-adjoint and normalized (however these can be imposed if they turn out to be necessary, cf. e.g. the characterization of a vacuum state in Definition 3.1 above) hence our answers can be complex numbers in general. Nevertheless $\tau(AB)$ is well-posed i.e., is finite and invariant under the gauge group of this theory namely the unitary automorphisms of $\mathfrak{R}$ thus it is indeed an "answer"—at least syntactically.

Now we come to the most difficult problem namely what are the meaningful questions here? This problem is already fully at the foggy semantical level. The orthodox approach says that a *question* should formally be like

From the collection $\mathrm{Spec}A \subset \mathbb{C}$ *of "all possible values the pointer of the experimental instrument designed to measure A in the laboratory can assume", which does occur in B?*

and the answer (as defined before) to this question is obtained through a *measurement*. Thus let us make a short interlude concerning the measurement. After performing the physical experiment designed to answer the question above should we expect that the state $B \in \mathfrak{R}$ will necessarily "collapse" to an eigenstate $B_\lambda \in \mathfrak{R}$ of A? In our opinion *no* and this is an essential difference between gravity and quantum mechanics. Namely, in quantum mechanics an ideal observer compared to the physical object to be observed is *infinitely large* hence the immense physical interaction accompanying the measurement procedure drastically disturbs the entity leading to the collapse of its state. However, in sharp contrast to this, in gravity an ideal observer is *infinitely small* hence it is reasonable to expect that measurements might *not* alter gravitational states. This is in accordance with our old experience concerning measurements in astronomy.

Concerning the problem of its *meaning*, since $A, B \in \mathfrak{R}$ have something to do with the curvature of local portions of space-time, it is not easy to assign a straightforward meaning to the above question. Therefore instead of offering a general solution to this problem at this initial state of the art, let us rather consider some special cases. For example if (M, g) is a classical non-singular space-time and (N, h) is another one, then their *classical geometrical* habitants may find that $0 \neq \tau(R_g R_h) \in \mathbb{C}$ in general. Consequently "physical contacts" between different classical geometries can already occur (whatever it means). It follows from the construction of the trace in (6) that $\tau(R_g R_h)$ arises by integrating the local trace function $\mathrm{tr}(R_g R_h)$ along $X = M \cap N$. Note that making use of the embeddings of M, N into $\mathfrak{R}$ given by (3) taking intersection (maybe empty) is meaningful. Thus given a nearly flat space-time (M, g) "we live in" i.e., an observer satisfying $|R_g(x)|_g \approx 0$ for all $x \in M$ and a different geometry (N, h) "we observe" i.e., a state R_h then we expect $|\tau(R_g R_h)| \approx 0$ in accord with our physical intuition that frequent encounters with different geometries in quantum gravity should occur rather in the strong gravity regime of space-times (hence the reason we do not experience such strange things).

As another example for measurement, consider the energy. In this universal quantum theory the only distinguished non-trivial self-adjoint gauge invariant operator is the identity $1 \in \mathfrak{R}$. Therefore the only natural candidate for playing the role of a *Hamiltonian* responsible for dynamics in this theory is $1 \in \mathfrak{R}$ (in natural units $c = \hbar = G = 1$). Therefore this dynamics is trivial. Nevertheless, quite interestingly, it coincides with the *modular dynamics* introduced by Connes and Rovelli [9] because it is associated with the *tracial* state τ on $\mathfrak{R}$ having the identity as its modular operator. Therefore the odd-looking equality (8) can be interpreted by saying that the expectation value (in a sequence of measurements) of the energy in the state represented by a vacuum curvature tensor R_g is the number $\tau(R_g)$ and this energy is equal to the cosmological constant constituent $\frac{\Lambda}{3}$ of the corresponding vacuum space-time (M,g). Note that R_g is at least self-adjoint (but perhaps not positive definite neither normalized) hence the energy is a real number but *a priori* can assume negative values too, cf. item (iv) of Definition 3.1.

An application to the cosmological constant problem. The last worry concerning the non-positivity of the energy strongly motivates to test this formalism on the so-called cosmological constant problem. By this we mean the complex problematics introduced into current standard cosmology (based on the cosmological principle and the corresponding Friedmann–Lemaître–Robertson–Walker or FLRW model) and standard particle physics (based on the Standard Model) by the experimental verification [23] of the existence of a strictly *positive* but *small* cosmological constant in 1998. Recall that in cosmology the cosmological constant is defined as $\Lambda = 3\left(\frac{H_0}{c}\right)^2 \Omega_\Lambda$ where H_0 is the Hubble "constant" (it actually varies with time) whose dimenion is s^{-1}, c is the speed of light having dimension ms^{-1} and the dimensionless number $\Omega_\Lambda \approx 0.69$ is the ratio of the dark energy density and the critical material energy density (Ω_Λ also changes in time). Thus the dimension of the physical cosmological constant is m^{-2}. Based on astronomical measurements $\Lambda \approx 2.89 \times 10^{-122}\, \ell_{\text{Planck}}^{-2}$ where $\ell_{\text{Planck}} \approx 1.62 \times 10^{-35}$ m. In order to gain compatibility with the mathematical expressions obtained in Section 2, from now on we shall use the dimensionless number

$$\frac{\Lambda}{3}\ell^2_{\text{Planck}} \approx 0.96 \times 10^{-122}$$

expressing the magnitude of the cosmological constant in natural Planck units. We claim that this sort of (i.e., a strictly positive but small) cosmological constant naturally appears if classical general relativity is replaced with its formal quantum generalization as introduced above. The cosmological constant problem has been investigated from a closely related angle by other authors too, cf. [2].

Before seeing this however let us briefly and roughly clarify what do we mean by this "replacement" from a physical viewpoint. We also want to see how a cosmological context has popped up so suddenly. The modern physical and mathematical basis for thinking about space, time, gravity and the structure and history of the Universe (cosmology) has been provided by classical general relativity in the form of a pseud-Riemannian 4-manifold (M,g) such that g is subject to Einstein's field equation. We have seen in Section 2 especially in Lemma 2.1 that out of this input data (in fact from M alone) an operator algebra $\mathfrak{R}$ and a Hilbert space $\mathscr{I}(M)^\perp$ carrying a representation ρ_M of $\mathfrak{R}$ can be constructed. Moreover the representation ρ_M on $\mathscr{I}(M)^\perp$ is the restriction of the unique standard representation π of $\mathfrak{R}$ on a Hilbert space $\mathscr{H}$ containing $\mathscr{I}(M)^\perp$. The latter objects namely $\mathfrak{R}, \mathscr{H}, \pi$ are all *unique* up to corresponding isomorphisms. The two structures (namely the "classical" (M,g) on the one side and the "quantum" $(\mathfrak{R}, \mathscr{H}, \pi)$ on the other) are naturally connected mathematically by the observation made in Section 2 that $\mathfrak{R}$ contains bounded curvature tensors and by Definition 3.1 offering a re-interpretation of the vacuum Einstein equation in an operator-theoretic language. Remember the plain technicality that a curvature tensor must be *bounded* in order to belong to the operator algebra. However it has been known for a long time that in general relativity the gravitational field of an isolated massive configuration cannot be regular everywhere [10, 11]. Thus we are forced to move towards a global non-singular cosmological context when trying to work out a physical connection between the two structures. Therefore hereby we make a *working hypothesis*: the unique abstract triple $(\mathfrak{R}, \mathscr{H}, \pi)$ as a quantum theory describes the Universe from the Big Bang till the Planck time $t_{\text{Planck}} \approx 5.39 \times 10^{-44}$ s while a highly non-unique pair (M,g) as a space-time in classical general relativity describes its evolution from t_{Planck} onwards. A usual choice for (M,g) is the FLRW solution with or without cosmological constant. The "moment" t_{Planck} would therefore formally or symbolically label the emergence or onset of a space-time structure in the course of the history of the Universe[5] establishing a physical correspondence between the two aforementioned mathematical structures. We symbolically write this correspondence as

$$\begin{array}{rcccc}(\mathfrak{R},\mathscr{H},\pi) & \Longrightarrow & \left(Q_M\in\mathfrak{R},\mathscr{I}(M)^\perp\subset\mathscr{H},\rho_M=\pi|_{\mathscr{I}(M)^\perp}\right) & \Longleftrightarrow & M\\ & \Longrightarrow & \left(Q_M,R_g\in\mathfrak{R},\mathscr{I}(M)^\perp\subset\mathscr{H},\rho_M=\pi|_{\mathscr{I}(M)^\perp}\right) & \Longleftrightarrow & (M,g)\end{array}$$

and having a sort of spontaneous symmetry breaking or phase transition in mind occuring at t_{Planck}. Actually we broke this process into two parts: "first" the emergence of the plain manifold structure M provided by a choice of an (in our convention dimensionless) operator $Q_M \in \mathfrak{R}$ satisfying the quantum vacuum Einstein equation as in Definition 3.1 and "then" the emergence of the geometry g hence the causal structure over M too provided by a further (in our convention dimensionless) curvature operator $R_g = Q_M + 8\pi T \in \mathfrak{R}$ satisfying the classical Einstein equation with already fixed cosmological constant $\frac{\Lambda}{3}\ell^2_{\text{Planck}} = \tau(Q_M)$ and some matter content described by a further (in our convention dimensionless) stress-energy tensor T. This passage from the unique quantum regime $(\mathfrak{R},\mathscr{H},\pi)$ to a particular classical regime (M,g) could probably be more rigorously captured in the framework of algebraic quantum field theory [17] as switching from the unique representation π to a different particular representation ρ_M of $\mathfrak{R}$. In this framework one can also formally label the transition with $T_{\text{Planck}} \approx 1.42 \times 10^{32}$ K; this high temperature is reasonable if we make a further technical observation that π is induced by the unique tracial state τ on $\mathfrak{R}$ hence one can imagine that $(\mathfrak{R},\mathscr{H},\pi)$ describes a quantum statistical mechanical system being in an infinitely high temperature Kubo–Martin–Schwinger or KMS equilibrium state τ (hence having trivial modular dynamics in the sense of [9]).

After this admittedly yet incomplete trial to set up some physical picture we can embark upon the treatment of the cosmological constant problem in this framework. Let M be a connected oriented smooth 4-manifold and consider its associated projector $P_M \in \mathfrak{R}$ from Lemma 2.1. Recall that an operator like P_M has been used to encode in the simplest way a particular connected oriented smooth 4-manifold M within its operator algebra $\mathfrak{R}$ which is however universal. Thus P_M is a distinguished operator in this sense. But even more, it is self-adjoint hence taking an arbitrary refinement $(\mathfrak{R},*)$ its average

$$Q_M := \frac{1}{2}\big(1-P_M+*(1-P_M)*\big)-\tau((1-P_M)*)* = 1-\frac{1}{2}(P_M+*P_M*)+\tau(P_M*)*$$

is an operator which solves the quantum vacuum Einstein equation too in the sense of Definition 3.1 with quantum cosmological constant $\frac{\Lambda}{3}\ell^2_{\text{Planck}} = \tau(Q_M)$. The particular choice for Q_M is verified by Lemma 2.1 making it sure that $\tau(Q_M) = \tau(1-P_M) = 1-\gamma(M)$ thus

$$\frac{\Lambda}{3}\ell^2_{\text{Planck}} = 1-\gamma(M)$$

in terms of the dimensionless γ-invariant of M. The further restriction imposed upon the spectrum of $\gamma(M)$ in Lemma 2.2 guarantees that actually

$$\frac{\Lambda}{3}\ell^2_{\text{Planck}} \in \left(0,\frac{1}{4}\right]\bigcup\left\{\frac{1}{4}\cos^{-2}\left(\frac{\pi}{n}\right)\,\middle|\, n=\dots 5,4,3\right\}\subset(0,1]$$

[5] We are aware of, but cannot do better, how nonsense it is, strictly speaking, to talk about the emergence of space and time at the moment t_{Planck}; this assertion requires the contradictory pre-supposition of time before t_{Planck}.

thus it is permitted to be arbitrary close but never equal to zero. That is, whatever M is, in our model *the (dimensionless) cosmological constant is always a small positive number* (thus, using again the terminology of Definition 3.1, in fact Q_M is a vacuum state). This is consistent with the aforementioned emprical value $\approx 0.96\times 10^{-122}$ of the cosmological constant.

To be more specific we can impose some conditions on M still compatible with cosmological experience. We can suppose that M has the structure $X\setminus\{\text{point}\}$ where X is a connected simply connected closed smooth 4-manifold; this is the case for the FLRW model in the relevant cases $k=0,-1$. Moreover we can suppose that the 2$^{\text{nd}}$ Betti number of X is non-trivial: $b_2(X)>0$ implying $b_2(M)>0$ too. This assumption does not hold in FLRW cosmology but it can be physically interpreted as the necessary topological condition for M i.e., the space-time to contain black holes at t_{Planck}. Indeed, taking e.g. a survey on known solutions [25], we can make an interesting observation that apparently all explicitly known 4 dimensional black hole solutions in vacuum general relativity have the property that because of some general reason their black hole content is even topologically recognizable as a two dimensional "hole" in space-time. This means that their instantaneous event horizons represented by immersed connected oriented surfaces in M give rise to non-trivial elements in $H_2(M;\mathbb{Z})^{\text{free}}\cong\mathbb{Z}^{b_2(M)}$. One can indeed prove this at least for stationary black holes [12]. Using this picture the number $b_2(M)$ can therefore be physically interpreted as the number of primordial black holes in the space-time at its onset. The assumption $0<b_2(M)$ implies that strictly $0<\frac{\Lambda}{3}\ell^2_{\text{Planck}}\lneqq 1$ such that letting $b_2(M)\to+\infty$ we find $0\leftarrow\frac{\Lambda}{3}\ell^2_{\text{Planck}}$. Consequently, the observed magnitude of the cosmological constant allows us to estimate in this model the number $N\approx b_2(M)$ of primordial black holes just around the Planck era. Plugging the empirical value of the cosmological constant into (7) we find

$$N\approx\frac{\log(\frac{\Lambda}{3}\ell^2_{\text{Planck}})}{\log\frac{1}{17}}\sim 10^2$$

yielding a number which is negligable. This qualitative result obtained by topological means is compatible with the agreement in the physicist literature that primordial black holes are essentially absent from the very early Universe, cf. e.g. [6, 16] (and the hundreds of references therein). Note that the usual considerations leading to this conclusion are based on the application of the Press–Schechter mechanism [22] for primordial black hole formation; hence our topological result is an independent confirmation of these considerations.

Acknowledgements. The author is grateful to T. Asselmeyer-Maluga, B. Gyenti and I. Nikolaev for the stimulating discussions. There are no conflicts of interest to declare that are relevant to the content of this article. The work meets all ethical standards applicable here. All the not-referenced results in this work are fully the author's own contribution. No funds, grants, or other financial supports were received. Data sharing is not applicable to this article as no datasets were generated or analysed during the underlying study.

References

[1] Anantharaman, C., Popa, S.T.: *An introduction to* II_1 *factors*, perprint, 331 pp. (no year), available at `http://www.math.ucla.edu/~popa/books.html`;

[2] Asselmeyer-Maluga, T., Król, J.: *How to obtain a cosmological constant from small exotic* $\mathbb{R}^4$, Phys. Dark Universe **19**, 66-77 (2018);

[3] Asselmeyer-Maluga, T., Król, J., Wilms, A.: *Big Bang and topology*, preprint, 17 pp. (2022);

[4] Bérard, P., Besson, G., Gallot, S.: *Embedding Riemannian manifolds by their heat kernel*, Geom. Functional Anal. **4**, 374-398 (1994);

[5] Besse, A.L.: *Einstein manifolds*, Springer-Verlag Berlin (1987);

[6] Carr, B., Kühnel, F.: *Primordial black holes as dark matter: recent developments*, Ann. Rev. Nucl. Part. Sci. **70**, 355-394 (2020);

[7] Chandrasekaran, V., Longo, R., Penington, G., Witten, E.: *An algebra of observables for de Sitter space*, preprint, 56 pp. (2022), arXiv: `2206.10780 [hep-th]`;

[8] Connes, A.: *Noncommutative geometry*, Academic Press, New York (1994);

[9] Connes, A., Rovelli, C.: *Von Neumann algebra automorphisms and time-thermodynamics relation in general covariant quantum theories*, Class. Quant. Grav. **11**, 2899-2918 (1994);

[10] Einstein, A.: *Demonstración de la no existencia de campos gravitacionales sin singularidades de masa total no nula*, Revista de la Universidad Nacional de Tucumán **A2**, 11-15 (1942);

[11] Einstein, A., Pauli, W.: *On the non-existence of regular stationary solutions of the relativistic field equations*, Ann. Math. **44**, 131-137 (1943);

[12] Etesi, G.: *A global uniqueness theorem for stationary black holes*, Comm. Math. Phys. **195**, 691-697 (1998);

[13] Etesi, G.: *Gravity as a four dimensional algebraic quantum field theory*, Adv. Theor. Math. Phys. **20**, 1049-1082 (2016);

[14] Etesi, G.: *The universal von Neumann algebra of smooth four-manifolds*, Adv. Theor. Math. Phys. **25**, 995-1024 (2021);

[15] Gompf, R.E., Stipsicz, A.I.: *4-manifolds and Kirby calculus*, GSM **20**, Amer. Math. Soc. Providence, Rhode Island (1999);

[16] Green, A.M.: *Primordial black holes: sirens of the early universe*, in: *Quantum aspects of black holes* (ed.: X. Calmet) FTPH **178**, 129-149, Springer (2014);

[17] Haag, R.: *Local quantum physics*, Springer–Verlag, Berlin (1993);

[18] Jones, F.R.V., Morrison, S. Snyder, N.: *Classification of subfactors of index at most* 5, Bull. Amer. Math. Soc. **51**, 277-327 (2014);

[19] Lawson Jr., H.B., Michelsohn, M-L.: *Spin geometry*, Princeton Univ. Press, Princeton, New Jersey (1989);

[20] Ocneanu, A. *Quantized groups, string algebras and Galois theory for algebras*, in: *Operator algebras and applications*, Vol. 2, London Math. Soc. Lecture Note Ser. **136**, Cambridge Univ. Press, Cambridge , 119-172 (1988);

[21] Popa, S.T.: *Classification of subfactors: the reduction to commuting squares*, Invent. Math. **101**, 19-43 (1990);

[22] Press, W.H., Schechter, P.: *Formation of galaxies and clusters of galaxies by self similar gravitational condensation*, Astrophys. Journ. **187**, 425-438 (1974);

[23] Riess, A. et al.: *Observational evidence from supernovae for an accelerating universe and a cosmological constant*, Astronomical Journ. **116**, 1009-1038 (1998);

[24] Singer, I.M., Thorpe, J.A.: *The curvature of* 4*-dimensional Einstein spaces*, in: *Global analysis*, Papers in honour of K. Kodaira, 355-365, Princeton Univ. Press, Princeton (1969);

[25] Stephani, H., Kramer, D., MacCallum, M., Hoenselaers, C., Herlt, E.: *Exact solutions of Einstein's field equations*, Cambridge Univ. Press, Cambridge (2003).