ADV. THEOR. MATH. PHYS.
Volume 25, Number 4, 995–1024, 2021

# The universal von Neumann algebra of smooth four-manifolds

Gábor Etesi

Making use of its smooth structure only, out of a connected oriented smooth 4-manifold a von Neumann algebra is constructed. As a special four dimensional phenomenon this von Neumann algebra is approximated by algebraic (i.e., formal) curvature tensors of the underlying 4-manifold and the von Neumann algebra itself is a hyperfinite factor of $\mathrm{II}_1$ type hence is unique up to abstract isomorphisms of von Neumann algebras. Nevertheless over a fixed 4-manifold this von Neumann algebra admits a 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.

Some consequences of this construction for quantum gravity are also discussed. Namely reversing the construction by starting not with a particular smooth 4-manifold but with the unique hyperfinite $\mathrm{II}_1$ factor, a conceptually simple but manifestly four dimensional, covariant, non-perturbative and genuinely quantum theory is introduced whose classical limit is general relativity in an appropriate sense. Therefore it is reasonable to consider it as a sort of quantum theory of gravity. In this model, among other interesting things, the observed positive but small value of the cosmological constant acquires a natural explanation.

## 1. Introduction

This paper, considered as a substantial technical and conceptual clarification of our earlier work [7], naturally splits up into two parts: a *mathematical* one describing a self-contained and relatively simple way how to attach up to abstract isomorphisms a single von Neumann algebra to every smooth 4-manifold by making use of their smooth structures only; and a *physical* part exhibiting a manifestly covariant, non-perturbative four dimensional quantum theory resembling a quantum theory of four dimensional space-time and gravity. This is achieved by reversing the mathematical construction.

The 20-21st 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 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 context however four dimensionality enters the game here as well if one tries to link differential geometry with non-commutativity [5]. Our main result is the following:

**Theorem 1.1.** *Let $M$ be a connected oriented smooth 4-manifold. Making use of its smooth structure only, a von Neumann algebra $\mathfrak{R}(M)$ can be constructed which is geometric in the sense that it contains a norm-dense subalgebra of algebraic (i.e., formal) curvature tensors on $M$ and $\mathfrak{R}(M)$ itself is a hyperfinite factor of type $\mathrm{II}_1$ (hence is unique up to abstract isomorphism of von Neumann algebras).*

*Moreover $\mathfrak{R}(M)$ 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* $\gamma(M) \in [0,1)$. *It behaves like* $\gamma(M \setminus Y) = \gamma(M)$ *under excision of homologically trivial submanifolds and* $\gamma(M \# N) = (\gamma(M) + \gamma(N))/(1 + \gamma(M)\gamma(N))$ *under connected sum.*

The outstanding problem of modern 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 tension with each other at not only the mathematical apparatus they rest on but even at a deep foundational level (cf. e.g. [10]): classical concepts of general relativity such as the space-time event, the light cone or the event horizon of a black hole are too "sharp" things from a quantum theoretic viewpoint while relativistic quantum field theory is not background independent from the aspect of general relativity. We do not attempt here to survey the vast physical, mathematical and even philosophical literature triggered by the unification problem; we just mention that nowadays the leading candidates expected to be capable for a sort of unification are Hamiltonian or Lagrangian canonical covariant quantization methods of gravity [2, 13] and string theory. But surely there is still a long way ahead; nevertheless most of physcists and mathematicians have the conviction that the language of classical general relativity will sound familiar to quantum theorists one day and *vice versa* i.e., conceptual and technical bridges must exist connecting the two pillars.

In this context it is interesting that reversing the mathematical approach leading to Theorem 1.1, a conceptionally simple, consistent and apparently novel [7] physical theory drops out which is genuinely quantum and is reasonable to say that it describes gravity in four dimensions. More precisely at least mathematically this theory rests on a solid basis provided by the powerful theory of von Neumann algebras follows. The starting setup is to take the unique hyperfinite $\mathrm{II}_1$ factor von Neumann algebra as the primordial structure and to consider the superabundance of oriented smooth 4-manifolds as being embedded into it (roughly speaking as orbits of the inner automorphism group of this von Neumann algebra) as well as view the members of a dense subalgebra of this von Neumann algebra as algebraic (i.e. formal, not stemming from an actual metric) curvature tensors along these 4-manifolds. The next step is to identify the von Neumann algebra itself with the algebra of local observables of a quantum theory; therefore in this quantum theory curvaturelike quantities in four dimensions are measured in a quantum mechanical sense. This measurement procedure is at least mathematically well-defined due to the existence of a unique faithful finite trace on the $II_1$ factor von Neumann algebra.

However the physical interpretation of this theory is more subtle and will not be systematically worked out here; of course the main obstacle is that, lacking direct encounters with strong gravitational fields, yet we cannot raise appropriate experimental physical questions what to measure. Therefore we will only consider sporadic examples, forecasted or suggested by the mathematical structure. For example, the central question in any physical theory is to understand the dynamics it describes. This is related with the concept of energy and time. In our universal quantum theory the only distinguished operator is the unit of the von Neumann algebra consequently it is the only candidate to play the role of a Hamiltonian here. Concerning this choice a convincing mathematical point is that the expectation value of this trivial Hamiltonian i.e., the corresponding energy in certain states coincides with the smooth 4-manifold invariant mentioned in Theorem 1.1; while a physical point is that calculating this energy in a "cosmological state" corresponding to the Friedman–Lemaître–Robertson–Walker model we obtain a simple qualitative explanation of the observed *small positive* value of the cosmological constant. On the other hand the dynamics is trivial in this theory since our choice for the Hamiltonian is trivial leading to the by-now familiar "problem of time" in gravity theories [2, 6] and in more generality [3].

We close the introduction with a comment on this time problem. Apparently our approach here supports the "no time at all" schools like [3]. However in fact, in our opinion, time is intrinsically and deeply present in all quantum theories and is responsible for their very properties like their probabilistic nature. Since the born of phenomenological thermodynamics in the 19th century, physicists have thought that the phenomenon of (macroscopic) time is related with or somehow stems from the thermodynamical properties of matter. A radical idea along these lines is a quite abandoned suggestion of von Weizsäcker [19]. He says that time cannot be completely described in terms of homogeneous geometric extensions like space because the empirically most obvious feature of time is its inhomogeneous character what he calls the *chronologicality of time*: the past consists of *facts* and in principle is subject to unambigous description, while in sharp contrast the future consists of *possibilities* hence allowing a probabilistic description only; the chronologicality of time is equivalent to the empirical validity of the *second law of thermodynamics*. Therefore, since a physical theory makes *predictions* on physical happenings—i.e., says something about the *future*—it must exhibit a probabilistic structure at a sufficiently fundamental level.

In this sense quantum theory is a fundamental theory and its intrinsic probabilistic nature reflects the intrinsic chronological nature of time. A recent proposal of Connes–Rovelli is the *thermodynamical time hypothesis* [6]: they introduce time in quantum theories by interpreting it as the (essentially unique) one-parameter group of modular automorphisms of von Neumann algebras associated with their thermal equilibrium (KMS) states. It turns out that our dynamics coincides with this modular dynamics in the infinitely high temperature (i.e., tracial state) limit.

The paper is organized as follows. Section 2 is self-contained and is devoted to a mathematically rigorous proof of Theorem 1.1 through a chain of lemmata by extending similar results of [7]. The physicist-minded reader can skip this section at first reading by accepting the content of Theorem 1.1. Section 3 contains the introduction of a four dimensional quantum theory of gravity by interpreting Theorem 1.1 from a physical viewpoint. The langauge of this section is therefore quite different from the previous one.

## 2. Mathematical construction

Following [7] consider 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 $2^{\mathrm{nd}}$ 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 $\Omega^2_c(M;\mathbb{C}) := C^\infty_c(M;\wedge^2T^*M\otimes_{\mathbb{R}}\mathbb{C})$ of compactly supported complexified 2-forms on $M$. Define $\langle\,\cdot\,,\,\cdot\,\rangle_{L^2(M)}:\Omega^2_c(M;\mathbb{C})\times\Omega^2_c(M;\mathbb{C})\to\mathbb{C}$ via integration, more precisely put

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

(complex-linear in its first and conjugate-linear in its second variable). This pairing is sesquilinear non-degenerate however is *indefinite* in general hence can be regarded as an indefinite sesquilinear scalar product on $\Omega^2_c(M;\mathbb{C})$. Let $\mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ denote the unital algebra of *all* $\mathbb{C}$-linear operators on $\Omega^2_c(M;\mathbb{C})$; in particular it consists of the unital subalgebra of all (not compactly supported!) bundle-morphisms of $\wedge^2T^*M\otimes_{\mathbb{R}}\mathbb{C}$ i.e.,

$$C^\infty(M;\mathrm{End}(\wedge^2T^*M\otimes_{\mathbb{R}}\mathbb{C}))\subsetneqq\mathrm{End}(\Omega^2_c(M;\mathbb{C})).$$

Likewise, diffeomorphisms are included via pullback i.e.,

$$\mathrm{Diff}^+(M) \subsetneqq \mathrm{End}(\Omega^2_c(M;\mathbb{C}))$$

as well.

In the spirit of *noncommutative geometry* [5] and recalling and extending results of [7] let us now destillate from the *plethora* of four dimensional smooth structures a *single* von Neumann algebra through a sequence of steps as follows. Our overall reference on von Neumann algebras is [1].

**Lemma 2.1.** *Let* $\circledast$ *be the adjoint operation on* $\mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ *with respect to the indefinite sesquilinear scalar product (1) i.e., formally defined by* $\langle A^\circledast\alpha,\beta\rangle_{L^2(M)} := \langle\alpha, A\beta\rangle_{L^2(M)}$ *for all* $\alpha,\beta\in\Omega^2_c(M;\mathbb{C})$. *Consider the* $\circledast$-*closed space*

$$V(M) := \big\{A\in\mathrm{End}(\Omega^2_c(M;\mathbb{C}))\mid A^\circledast\in\mathrm{End}(\Omega^2_c(M;\mathbb{C}))\ \text{exists and } r(A^\circledast A)<+\infty\big\}$$

*defined by the* $\mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ *spectral radius*

$$r(B) := \sup_{\lambda\in\mathbb{C}}\big\{|\lambda|\ \big|\ B-\lambda\mathrm{Id}_{\Omega^2_c(M;\mathbb{C})}\in\mathrm{End}(\Omega^2_c(M;\mathbb{C}))\ \text{is not bijective}\big\}\,.$$

*Then* $\sqrt{r}$ *is a norm and the corresponding completion of* $V(M)$ *renders* $(V(M),\circledast)$ *a* $C^*$-*algebra* $\mathfrak{R}(M)$. *This* $C^*$-*algebra is non-trivial in the sense that* $\mathfrak{R}(M)$ *contains the space of all bounded bundle morphisms i.e.,*

$$C^\infty(M;\mathrm{End}(\wedge^2T^*M\otimes_{\mathbb{R}}\mathbb{C}))\cap V(M)$$

*as well as all orientation preserving diffeomorphisms of* $M$ *i.e.,* $\mathrm{Diff}^+(M)$. *Hence in particular it possesses a unit* $1\in\mathfrak{R}(M)$.

*Proof.* Our strategy to prove the lemma is as follows. Obviously $(V(M),\circledast)$ is a $*$-algebra. Provided it can be equipped with a norm such that the corresponding completion of $V(M)$ improves $(V(M),\circledast)$ to a $C^*$-algebra then, knowing the uniqueness of the $C^*$-algebra norm, this sought norm $[[\,\cdot\,]]$ on all $A\in V(M)$ must look like $[[A]]^2=[[A^\circledast]]^2=[[A^\circledast A]]=r(A^\circledast A)$. Therefore we want to see that the spectral radius gives a norm here by comparing it with known other norms.

We begin with some preparations. First, being $\langle\,\cdot\,,\,\cdot\,\rangle_{L^2(M)}$ given by (1) non-degenerate, there exist non-canonical decompositions

$$\Omega^2_c(M;\mathbb{C}) = \Omega^+_c(M;\mathbb{C}) \oplus \Omega^-_c(M;\mathbb{C}) \tag{2}$$

into maximal definite orthogonal subspaces i.e.,

$$\pm\langle\,\cdot\,,\,\cdot\,\rangle_{L^2(M)}|_{\Omega^\pm_c(M;\mathbb{C})} : \Omega^\pm_c(M;\mathbb{C}) \times \Omega^\pm_c(M;\mathbb{C}) \to \mathbb{C}$$

are both positive definite moreover $\Omega^+_c(M;\mathbb{C}) \perp_{L^2(M)} \Omega^-_c(M;\mathbb{C})$. Therefore these restricted scalar products can be used to complete $\Omega^\pm_c(M;\mathbb{C})$ to separable Hilbert spaces $h^\pm(M)$ respectively yielding non-canonical direct sum Hilbert space completions $h^+(M) \oplus h^-(M) \supset \Omega^2_c(M;\mathbb{C})$ with particular non-degenerate *positive definite* scalar products $(\alpha,\beta)_{L^2(M)} := \langle\alpha^+,\beta^+\rangle_{L^2(M)} - \langle\alpha^-,\beta^-\rangle_{L^2(M)}$ and induced norms $\|\cdot\|_{L^2(M)}$ on these completions. Here $\alpha^\pm := P^\pm\alpha$, etc. where $P^\pm : h^+(M)\oplus h^-(M) \to h^\pm(M)$ are the orthogonal projections with respect to (1). Put $J := P^+ - P^-$ and let $\dagger$ denote the adjoint over $h^+(M)\oplus h^-(M)$. Then $J^2 = \mathrm{Id}_{h^+(M)\oplus h^-(M)}$ and $J^\dagger = J$ hence $J$ is a unitary operator on $h^+(M)\oplus h^-(M)$. It formally satisfies $A^* = JA^\dagger J$ and $A^\dagger = JA^*J$.

Recall that the operator norm on $\mathfrak{B}(h^+(M)\oplus h^-(M))$, the $C^*$-algebra of bounded linear operators on $h^+(M)\oplus h^-(M)$, is

$$\|B\| = \sup_{\|v\|_{L^2(M)}=1} \|Bv\|_{L^2(M)} = \sup_{\|v\|_{L^2(M)}=1,\|w\|_{L^2(M)}=1} \mathrm{Re}(Bv,w)_{L^2(M)}. \tag{3}$$

The adjoint $\dagger$ and this norm $\|\cdot\|$ are actually the $*$-operation and norm on $\mathfrak{B}(h^+(M)\oplus h^-(M))$ therefore by the uniqueness of the $C^*$-algebra norm $\|B\|^2 = \|B^\dagger\|^2 = \|B^\dagger B\| = r'(B^\dagger B)$ where we can define the $\mathfrak{B}(h^+(M)\oplus h^-(M))$ spectral radius as

$$r'(B) := \sup_{\lambda\in\mathbb{C}}\Big\{|\lambda| \;\Big|\; B - \lambda\mathrm{Id}_{h^+(M)\oplus h^-(M)} \in \mathrm{End}(h^+(M)\oplus h^-(M)) \text{ is not bijective}\Big\}$$

by the bounded inverse theorem. It readily follows that if $B \in \mathfrak{B}(h^+(M)\oplus h^-(M)) \cap \mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ then $r(B) = r'(B)$ and if $A \in V(M)$ then $r(A^*A) = r'(A^*A)$. Our last ingredient is Gelfand's spectral radius formula $r'(B) = \lim\limits_{k\to+\infty} \|B^k\|^{\frac{1}{k}} \leqq \|B\|$ (cf. e.g. [16, Sect. XI.149]).

After these preparations we can embark upon the proof. Consider first any bounded linear operator $A \in \mathfrak{B}(h^+(M)\oplus h^-(M)) \cap \mathrm{End}(\Omega^2_c(M;\mathbb{C}))$.

Then $A^\dagger$ hence $A^\circledast = JA^\dagger J$ exists and $\|A\| < +\infty$. Moreover by $\|J\| = 1$ we get on the one hand

$$r(A^\circledast A) = r(JA^\dagger JA) = r'(JA^\dagger JA) \leqq \|JA^\dagger JA\| \leqq \|J\|^2\|A\|^2 = \|A\|^2$$

consequently $A \in V(M)$. Proceeding further, it is straightforward that $\|JA^\circledast JA\| = \|A^\dagger A\|$. Additionally it follows from (3) and the unitarity of $J$ that $\|A^\circledast A\| = \|JA^\dagger JA\| = \|A^\dagger JA\|$. Suppose $A$ is invertible; if $v \in h^+(M) \oplus h^-(M)$ is a unit vector then so is $u = A^{-1}JAv$. Hence $\|A^\dagger Au\|_{L^2(M)} = \|A^\dagger JAv\|_{L^2(M)}$ and we find via (3) that $\|A^\dagger JA\| = \|A^\dagger A\|$. Consequently we obtain for invertible hence by continuity for all operators $A \in \mathfrak{B}(h^+(M) \oplus h^-(M)) \cap \mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ that $\|JA^\circledast JA\| = \|A^\circledast A\|$ (and commonly equal to $\|A\|^2$) implying $\|(JA^\circledast JA)^k\| = \|(A^\circledast A)^k\|$ for all $k \in \mathbb{N}$.

Conversely, consider secondly any $A \in V(M)$. Then $A^\circledast$ hence $A^\dagger = JA^\circledast J$ exists and $r(A^\circledast A) < +\infty$. Therefore, on the other hand, for any $\varepsilon > 0$ one can find a positive integer $k$ such that

$$\begin{aligned} r(A^\circledast A) + \varepsilon = r'(A^\circledast A) + \varepsilon &\geqq \|(A^\circledast A)^k\|^{\frac{1}{k}} = \|(JA^\circledast JA)^k\|^{\frac{1}{k}} \\ &\geqq r'(JA^\circledast JA) - \varepsilon = r'(A^\dagger A) - \varepsilon = \|A\|^2 - \varepsilon \end{aligned}$$

yielding, since $\varepsilon > 0$ was arbitrary, that in fact

$$r(A^\circledast A) \geqq \|A\|^2$$

consequently $A \in \mathfrak{B}(h^+(M) \oplus h^-(M)) \cap \mathrm{End}(\Omega^2_c(M;\mathbb{C}))$. We eventually conclude that

(4)
$$r(A^\circledast A) = \|A\|^2 \quad \text{along} \quad V(M) = \mathfrak{B}(h^+(M) \oplus h^-(M)) \cap \mathrm{End}(\Omega^2_c(M;\mathbb{C}))$$

demonstrating that the spectral radius indeed provides us with a norm on $V(M)$. Therefore putting

(5)
$$[[A]] := \sqrt{r(A^\circledast A)}$$

we can complete $V(M)$ with respect to this norm and enrich the $*$-algebra $(V(M), \circledast)$ to a $C^*$-algebra $\mathfrak{R}(M)$. Additionally we obtain

$$\mathfrak{R}(M) \subset \mathfrak{B}(h^+(M) \oplus h^-(M))$$

however this is not a $*$-inclusion.

Finally, it is obvious that $R \in C^\infty(M; \mathrm{End}(\wedge^2 T^*M \otimes_{\mathbb{R}} \mathbb{C})) \cap V(M)$ if and only if $[[R]] < +\infty$ hence $R \in \mathfrak{R}(M)$ i.e., $\mathfrak{R}(M)$ contains bounded (in this sense) bundle morphisms of $\wedge^2 T^*M \otimes_{\mathbb{R}} \mathbb{C}$. Likewise, a diffeomorphism $\Phi \in \mathrm{Diff}^+(M)$ acts on $\Omega^2_c(M; \mathbb{C})$ via pullback $\omega \mapsto \Phi^*\omega$ and by the orientation-preserving-diffeomorphism invariance of integration obviously

$$\langle \Phi^*\alpha, \Phi^*\beta \rangle_{L^2(M)} = \langle \alpha, \beta \rangle_{L^2(M)}$$

holds; hence we can see that diffeomorphisms are unitary operators i.e., $(\Phi^*)^{\circledast}(\Phi^*) = \mathrm{Id}_{\Omega^2_c(M;\mathbb{C})}$. Consequently $[[\Phi^*]] = 1$ demonstrating $\Phi^* \in \mathfrak{R}(M)$. In particular $\mathfrak{R}(M)$ possesses a unit 1 represented either by the identity bundle morphism $\mathrm{Id}_{\wedge^2 T^*M \otimes_{\mathbb{R}} \mathbb{C}}$ or by $1 \in \mathrm{Diff}^+(M)$ as claimed. ◇

**Remark.** Note that by construction $[[\cdot]] = \|\cdot\|$ where this latter norm is the operator norm (3) for any particular completion $h^+(M) \oplus h^-(M) \supset \Omega^2_c(M; \mathbb{C})$; hence these norms in fact numerically coincide on their common domain $\mathfrak{B}(h^+(M) \oplus h^-(M)) \cap \mathrm{End}(\Omega^2_c(M; \mathbb{C}))$ equal to $V(M)$ by (4).

**Lemma 2.2.** *The norm $[[\cdot]]$ given by (5) on $\mathfrak{R}(M)$ can be improved to a Hermitian scalar product $(\cdot, \cdot) : \mathfrak{R}(M) \times \mathfrak{R}(M) \to \mathbb{C}$ rendering $\mathfrak{R}(M)$ a Hilbert space $\mathscr{H}(M)$ isomorphic to $\mathfrak{R}(M)$ as a complete complex vector space and carrying an action of $\mathfrak{R}(M)$.*

*Moreover $\mathfrak{R}(M) \subset \mathfrak{B}(\mathscr{H}(M))$ inherits the structure of a von Neumann algebra with a finite trace functional $\tau : \mathfrak{R}(M) \to \mathbb{C}$ satisfying $\tau(1) = 1$.*

*Proof.* Let $N : \mathfrak{R}(M) \to \mathbb{R}$ be the squared norm-function $N(T) := [[T]]^2$; it satisfies $N(1) = 1$ where $1 \in \mathfrak{R}(M)$ is the unit. Formally take its derivative at $1 \in \mathfrak{R}(M)$ restricted to $\mathfrak{R}(M)$ i.e., the $\mathbb{C}$-linear map

$$\begin{array}{cccc} N_*(1): & \mathfrak{R}(M) & \longrightarrow & \mathbb{R} \\ & \cap & & \wr\| \\ & T_1\mathfrak{R}(M) & \longrightarrow & T_1\mathbb{R} \end{array}$$

and with two $A, B \in \mathfrak{R}(M)$ consider the pairing

$$(A, B)^{\mathbb{R}} := \frac{1}{2} N_*(1)(A^{\circledast}B + B^{\circledast}A) \tag{6}$$

which, if exists, is clearly a symmetric, $\mathbb{R}$-linear map $\mathfrak{R}(M)^{\mathbb{R}} \times \mathfrak{R}(M)^{\mathbb{R}} \to \mathbb{R}$ where $\mathfrak{R}(M)^{\mathbb{R}}$ denotes the real vector space underlying $\mathfrak{R}(M)$ (i.e., $\mathfrak{R}(M)^{\mathbb{R}}$ coincides with $\mathfrak{R}(M)$ except that the scalar multiplication by $\sqrt{-1}$ is not defined in $\mathfrak{R}(M)^{\mathbb{R}}$). We demonstrate that at least this pairing exists, is in fact non-degenerate and definite hence gives rise to a definite real scalar product on $\mathfrak{R}(M)^{\mathbb{R}}$.

From the proof of Lemma 2.1 consider again $h^+(M)\oplus h^-(M)\supset\Omega^2_c(M;\mathbb{C})$. We can assume that $V(M)$, hence $\mathfrak{R}(M)^{\mathbb{R}}$ acts on this Hilbert space by bounded operators (this is neither a $*$-action nor a direct-sum-preserving action) and by (4) and (5) we know that if $T\in\mathfrak{R}(M)$ then $[[T]]=\|T\|$ with the corresponding operator norm. This implies by (3) that for every $T\in\mathfrak{R}(M)$, $0<\varepsilon$ and $t\in\mathbb{R}$ there is $\omega\in\Omega^2_c(M;\mathbb{C})$ with $\|\omega\|_{L^2(M)}=1$ such that $\|(1+tT)\omega\|^2_{L^2(M)}-\varepsilon\leqq N(1+tT)\leqq\|(1+tT)\omega\|^2_{L^2(M)}+\varepsilon$. If $T$ belongs to a dense subset of $\mathfrak{R}(M)^{\mathbb{R}}$ then $\operatorname{Spec}T$ consists of eigenvalues only (cf. the proof of Lemma 2.3) hence we can assume in this case that $\omega$ is an eigenform of $T$ hence is independent of $t$. Consequently taking the limit $t\to 0$ in the induced inequality

$$\begin{aligned}\operatorname{Re}(T\omega,\omega)_{L^2(M)}-\frac{|t|}{2}\|T\|^2-\varepsilon\leqq\frac{1}{2}\frac{N(1+tT)-N(1)}{t}\\ \leqq\operatorname{Re}(T\omega,\omega)_{L^2(M)}+\frac{|t|}{2}\|T\|^2+\varepsilon\end{aligned}$$

we get

$$\operatorname{Re}(T\omega,\omega)_{L^2(M)}-\varepsilon\leqq\frac{1}{2}N_*(1)(T)\leqq\operatorname{Re}(T\omega,\omega)_{L^2(M)}+\varepsilon.$$

Therefore with $T:=A^*B+B^*A$ we find that $\frac{1}{2}N_*(1)(A^*B+B^*A)$ hence $(A,B)^{\mathbb{R}}$ exists on a dense subspace. Proceeding further if $|t|$ is sufficiently small then we can assume that $\operatorname{Spec}(2tA^*A)$ is contained in a small open ball about the origin hence $0\leqq\frac{\delta}{2}\leqq r(2tA^*A)\leqq\delta$. Recall the spectral mapping theorem [16, Sect. X.151] which states that if $T\in\mathfrak{B}(h^+(M)\oplus h^-(M))$ and $u:\mathbb{C}\to\mathbb{C}$ is a continuous function then there is an equality $u(\operatorname{Spec}(T))=\operatorname{Spec}u(T)$. Thus putting $T:=2tA^*A$ and $u(z):=(1+z)^2$ we know that

$$1+r(2tA^*A)\leqq 1+\delta\leqq(1+\frac{\delta}{2})^2\leqq(1+r(2tA^*A))^2=r((1+2tA^*A)^2).$$

By the aid of (5) we write $r(2tA^*A)=[[\sqrt{2t}A]]^2=2|t|[[A]]^2$ moreover $(1+2tA^*A)^2=(1+2tA^*A)^*(1+2tA^*A)$ thus

$$\begin{aligned}1+2|t|[[A]]^2=1+r(2tA^*A)\leqq r((1+2tA^*A)^2)=N(1+2tA^*A)\\ \leqq 1+4|t|[[A]]^2+4t^2[[A]]^4\end{aligned}$$

consequently we come up again with a two-sided estimate

$$[[A]]^2 \leqq \frac{1}{2}\frac{N(1+2tA^*A)-N(1)}{t} \leqq 2[[A]]^2+2|t|[[A]]^4$$

demonstrating after $t \to 0$ that $N_*(1)(A^*A)$ hence $(A,A)^{\mathbb{R}}$ vanishes if and only if $A=0$. Thus $(\,\cdot\,,\,\cdot\,)^{\mathbb{R}}$ is a densly defined non-degenerate real scalar product on $\mathfrak{R}(M)^{\mathbb{R}}$. Since $N_*(1)((\sqrt{-1}A)^*\sqrt{-1}A) = N_*(1)(A^*A)$ holds we can improve $(\,\cdot\,,\,\cdot\,)^{\mathbb{R}}$ to a densly defined Hermitian scalar product $(\,\cdot\,,\,\cdot\,)$ as usual:

$$\begin{aligned}(A,B) := \frac{1}{2}\big(&N_*(1)((A+B)^*(A+B)) - N_*(1)(A^*A) - N_*(1)(B^*B)\big)\\ &+\frac{\sqrt{-1}}{2}\big(N_*(1)((A+\sqrt{-1}B)^*(A+\sqrt{-1}B))\\ &\quad - N_*(1)(A^*A) - N_*(1)((\sqrt{-1}B)^*\sqrt{-1}B)\big)\end{aligned}$$

(complex-linear in its first and conjugate-linear in its second variable). Finally we continuously extend this to $\mathfrak{R}(M)$; for notational simplicity we continue to denote the induced norm on $\mathfrak{R}(M)$ by $[[\,\cdot\,]]$. Completing $\mathfrak{R}(M)$ with respect to this norm improves $\mathfrak{R}(M)$ to a Hilbert space $\mathscr{H}(M)$. To distinguish the two spaces we will write $A \in \mathfrak{R}(M)$ as so far but $\hat{B} \in \mathscr{H}(M)$. Note that $[[A]] \leqq [[\hat{A}]] \leqq \sqrt{2}[[A]]$ hence $\mathfrak{R}(M)$ and $\mathscr{H}(M)$ are in fact isomorphic as complete complex vector spaces.

Concerning the second statement, since $\mathfrak{R}(M)$ acts on itself via multiplication from the left which is continuous and by (6) satisfies $(\widehat{AB},\hat{C}) = (\hat{B},\widehat{A^*C})$ i.e., is compatible with the scalar product on $\mathscr{H}(M)$, we obtain a unital $*$-inclusion $\pi_M : \mathfrak{R}(M) \to \mathfrak{B}(\mathscr{H}(M))$ into the $C^*$-algebra of bounded linear operators on $\mathscr{H}(M)$ i.e., a faithful representation of $\mathfrak{R}(M)$ on $\mathscr{H}(M)$ given by $\pi_M(A)\hat{B} := \widehat{AB}$. To prove that $\mathfrak{R}(M)$ is a von Neumann algebra over $\mathscr{H}(M)$ it is enough to demonstrate that $\pi_M(\mathfrak{R}(M))$ is closed in any topology on $\mathfrak{B}(\mathscr{H}(M))$ different from its uniform topology [1, Theorem 2.1.3]. So let $\{A_i\}_{i\in\mathbb{N}}$ be a sequence in $\mathfrak{R}(M)$ such that $\{\pi_M(A_i)\}_{i\in\mathbb{N}}$ converges for instance in the strong topology on $\mathfrak{B}(\mathscr{H}(M))$ to an element $a \in \mathfrak{B}(\mathscr{H}(M))$ i.e., for all $\hat{B}_1,\dots,\hat{B}_k \in \mathscr{H}(M)$ with associated seminorms

$$\begin{aligned}&\lim_{i\to+\infty}[[\pi_M(A_i)-a]]_{\hat{B}_1,\dots,\hat{B}_k}\\ &\qquad = \lim_{i\to+\infty}\Big([[(\pi_M(A_i)-a)\hat{B}_1]]+\dots+[[(\pi_M(A_i)-a)\hat{B}_k]]\Big) = 0\end{aligned}$$

holds. Hence taking $k = 1$ and $\hat{B}_1 := \hat{1} \in \mathscr{H}(M)$ we know that

$$\lim_{i \to +\infty} [[\pi_M(A_i) - a]]_{\hat{1}} = 0$$

consequently

$$\begin{aligned} [[A_i - A_j]] \leqq [[\hat{A}_i - \hat{A}_j]] \leqq [[\hat{A}_i - a\hat{1}]] + [[\hat{A}_j - a\hat{1}]] \\ = [[\pi_M(A_i) - a]]_{\hat{1}} + [[\pi_M(A_j) - a]]_{\hat{1}} \to 0 \quad i, j \to +\infty \end{aligned}$$

convincing us that $\{A_i\}_{i \in \mathbb{N}}$ is a Cauchy sequence in the complete space $\mathfrak{R}(M)$ hence it has a unique limit $A \in \mathfrak{R}(M)$. By continuity of $\pi_M$ we find that in fact $\pi_M(A) = a$ therefore $a = A \in \mathfrak{R}(M)$. This demonstrates that $\mathfrak{R}(M)$ is a von Neumann algebra as desired, operating on $\mathscr{H}(M)$.

Finally, the scalar pruduct on $\mathscr{H}(M)$ has the straightforward property $(\hat{A}, \hat{B}) = (\widehat{B^*}, \widehat{A^*})$ consequently putting

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

we obtain a $\mathbb{C}$-linear map $\tau : \mathfrak{R}(M) \to \mathbb{C}$ having the properties $|\tau(A)| \leqq [[\hat{A}]] \leqq \sqrt{2}[[A]] < +\infty$ as well as $\tau(AB) = \tau(BA)$ and $\tau(1) = 1$ i.e., it is a finite trace on $\mathfrak{R}(M)$ as claimed. $\diamond$

**Remark.** Note that although $\mathfrak{R}(M)$ and $\mathscr{H}(M)$ are isomorphic as complete complex vector spaces they are not isomorphic as unitary- more precisely as $\mathrm{U}(\mathscr{H}(M))$-modules: given a unitary operator $U \in \mathrm{U}(\mathscr{H}(M))$ then $A \in \mathfrak{R}(M)$ is acted upon as $A \mapsto UAU^{-1}$ but $\hat{B} \in \mathscr{H}(M)$ transforms as $\hat{B} \mapsto U\hat{B}$.

Let us continue exploring the structure of the operator algebra $\mathfrak{R}(M)$ by making contacts with the local *four dimensional* differential geometry of $M$. In fact all the constructions so far work for an arbitrary oriented and smooth $4k$-manifold with $k = 0, 1, 2, \dots$[1] however the next lemma is the very manifestation of four dimensionality permanently lurking behind these considerations.

[1]In $4k + 2$ dimensions the pairing (1) gives rise to a symplectic structure on $2k + 1$-forms.

**Lemma 2.3.** *The von Neumann algebra $\mathfrak{R}(M)$ is geometric in the sense that for every $A\in\mathfrak{R}(M)$ there exists a sequence*

$$\{R_i(A)\in C^\infty(M;\mathrm{End}(\wedge^2T^*M\otimes_{\mathbb{R}}\mathbb{C}))\cap V(M)\mid i\in\mathbb{N}\}$$

*with the property*

$$\lim_{i\to+\infty}[[A-R_i(A)]]=0 \tag{7}$$

*where $[[\,\cdot\,]]$ is the spectral radius norm (5) for which $\mathfrak{R}(M)$ is complete. In particular $\mathfrak{R}(M)$ contains all bounded complexified algebraic (i.e., formal) curvature tensors on $M$.*

*Moreover $\mathfrak{R}(M)$ is a hyperfinite factor of type $\mathrm{II}_1$ (hence is unique up to abstract isomorphisms of von Neumann algebras).*

*Proof.* A peculiarity of four dimensions is that the $*$-subalgebra $C^\infty(M;\mathrm{End}(\wedge^2T^*M\otimes_{\mathbb{R}}\mathbb{C}))$ of bundle morphisms contains the space of algebraic (i.e., formal) curvature tensors on $M$. For example if $(M,g)$ is an oriented Riemannian 4-manifold then its honest, i.e., not just formal, Riemannian curvature tensor $R_g$ is a member of this subalgebra and with respect to the decomposition of 2-forms into their (anti)self-dual parts it looks like (cf. [17])

$$R_g=\begin{pmatrix}\frac{1}{12}\mathrm{Scal}+\mathrm{Weyl}^+ & \mathrm{Ric}_0\\ \mathrm{Ric}_0^* & \frac{1}{12}\mathrm{Scal}+\mathrm{Weyl}^-\end{pmatrix}:\begin{matrix}\Omega^+_c(M;\mathbb{C})\\ \bigoplus\\ \Omega^-_c(M;\mathbb{C})\end{matrix}\longrightarrow\begin{matrix}\Omega^+_c(M;\mathbb{C})\\ \bigoplus\\ \Omega^-_c(M;\mathbb{C})\end{matrix}. \tag{8}$$

By definition elements of $\mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ map the space of sections $\Omega^2_c(M;\mathbb{C})=C^\infty_c(M;\wedge^2T^*M\otimes_{\mathbb{R}}\mathbb{C})$ into itself. Consequently for all $A\in V(M)\subset\mathrm{End}(\Omega^2_c(M;\mathbb{C}))$ and $\omega\in\Omega^2_c(M;\mathbb{C})$ the image satisfies $A\omega\in\Omega^2_c(M;\mathbb{C})$ hence define

$$R(A,\omega)\in C^\infty(M;\mathrm{End}(\wedge^2T^*M\otimes_{\mathbb{R}}\mathbb{C}))\cap V(M)\subsetneqq\mathrm{End}(\Omega^2_c(M;\mathbb{C}))\cap V(M)$$

by writing the image as $A\omega=R(A,\omega)\omega$ i.e., $(A\omega)_x=R(A,\omega)_x\omega_x$ at every point $x\in M$. The notation indicates that in general this algebraic curvature tensor $R(A,\omega)$ depends even on the 2-form $\omega$, and $A$ itself is an algebraic curvature tensor precisely if $A=R(A)$ i.e., is independent of $\omega$. By construction $\mathfrak{R}(M)$ is generated by $V(M)$ consequently for every $A\in\mathfrak{R}(M)$ and $\omega\in\Omega^2_c(M;\mathbb{C})\subset h^+(M)\oplus h^-(M)$ and real number $\varepsilon>0$ we know $A\omega\in$

$h^+(M) \oplus h^-(M)$ and there exists $R(A,\omega)$ as above such that

$$\|A\omega - R(A,\omega)\omega\|_{L^2(M)} \leqq \frac{\varepsilon}{2}$$

holds. Let $\{\omega_i\}_{i\in\mathbb{N}}$ be a once and for all fixed countable dense subset of $\Omega^2_c(M;\mathbb{C}) \subset h^+(M) \oplus h^-(M)$ and given $R(A,\omega)$, for all $i = 1, 2\ldots$ put $R_i(A) := R(A, \omega_i)$. These algebraic curvature tensors have the property that for any $\omega \in \Omega^2_c(M;\mathbb{C}) \subset h^+(M) \oplus h^-(M)$ and $\varepsilon > 0$ we can find infinitely many indices $i \in \mathbb{N}$ such that

$$\|R(A,\omega)\omega - R_i(A)\omega\|_{L^2(M)} \leqq \frac{\varepsilon}{2} \quad .$$

Then $\{R_i(A) \in C^\infty(M; \mathrm{End}(\wedge^2 T^*M \otimes_{\mathbb{R}} \mathbb{C})) \cap V(M) \mid i \in \mathbb{N}\}$ approximates $A \in \mathfrak{R}(M)$ in the following sense. First note that for all $\omega \in \Omega^2_c(M;\mathbb{C}) \subset h^+(M) \oplus h^-(M)$ and $\varepsilon > 0$ we can find infinitely many indices $i \in \mathbb{N}$ such that

$$\begin{aligned}\|(A - R_i(A))\omega\|_{L^2(M)} &\leqq \|(A - R(A,\omega))\omega\|_{L^2(M)} \\ &\quad + \|(R(A,\omega) - R_i(A))\omega\|_{L^2(M)} \\ &\leqq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\end{aligned}$$

consequently for these indices

$$[[A - R_i(A)]] = \|A - R_i(A)\| = \sup_{\|\omega\|_{L^2(M)}=1} \|(A - R_i(A))\omega\|_{L^2(M)} \leqq \varepsilon$$

yielding $\liminf_{i\to+\infty} [[A - R_i(A)]] = 0$ hence passing to a subsequence we come up with (7).

From this density result we immediately draw two consequences. The first consequence is that $\mathfrak{R}(M)$ is hyperfinite. We demonstrate this by proving that to every finite collection $A_1, \ldots, A_k \in \mathfrak{R}(M)$ of operators and real number $\varepsilon > 0$ there exists a finite dimensional $*$-subalgebra $\mathfrak{S} \subset \mathfrak{R}(M)$ such that $[[A_1 - \mathfrak{S}]] \leqq \varepsilon, \ldots, [[A_k - \mathfrak{S}]] \leqq \varepsilon$. Given an operator $A_j \neq 0$ acting on $h^+(M) \oplus h^-(M)$, take any complex number $\lambda_j \in \mathrm{Spec}\, A_j \subset \mathbb{C}$ from its non-empty spectrum and put $B_{\lambda_j} := A_j - \lambda_j 1 \in \mathfrak{R}(M)$. Then $B_{\lambda_j}$ is not bijective. If $\{0\} \neq \mathrm{Ker}\, B_{\lambda_j}$ for all $j = 1, \ldots, k$ then pick any $0 \neq v_j \in \mathrm{Ker}\, B_{\lambda_j}$. It straightforwardly follows from $A_j v_j = \lambda_j v_j + B_{\lambda_j} v_j = \lambda_j v_j$ that taking the finite dimensional complex subspace $W \subset h^+(M) \oplus h^-(M)$ spanned by $v_1, \ldots, v_k$ and then putting $\mathfrak{S} := \mathrm{End}\, W \cap \mathfrak{R}(M)$ the finite dimensional $*$-subalgebra $\mathfrak{S} \cong \mathfrak{gl}(W)$ satisfies $\mathfrak{S} \subset \mathfrak{R}(M)$ and $A_1, \ldots, A_k \in \mathfrak{S}$. Hence $\mathfrak{S}$

possesses the required property. If it happens that $\mathrm{Ker}\, B_{\lambda_j} = \{0\}$ for some $j = 1, \dots, k$ then replace $A_j$ with any approximating algebraic curvature tensor $R(A_j)$ satisfying $[[A_j - R(A_j)]] \leqq \varepsilon$ and this time take $\lambda_j \in \mathrm{Spec}\, R(A_j) \subset \mathbb{C}$ and $B_{\lambda_j} := R(A_j) - \lambda_j 1 \in \mathfrak{R}(M)$. Then $B_{\lambda_j}$ is again not bijective but surely $\mathrm{Ker}\, B_{\lambda_j} \neq \{0\}$. Indeed, assume the converse is true. Because $B_{\lambda_j} = R(A_j) - \lambda_j \mathrm{Id}_{\wedge^2 T^*M \otimes_{\mathbb{R}} \mathbb{C}}$ is itself an algebraic curvature tensor, being non-bijective means that at least in a neighbourhood $U \subsetneqq M$ of a point the finite dimensional maps $B_{\lambda_j}|_x : \wedge^2 T^*_x M \otimes_{\mathbb{R}} \mathbb{C} \to \wedge^2 T^*_x M \otimes_{\mathbb{R}} \mathbb{C}$ are not invertible for all $x \in U$. But in this case there would exist an element $0 \neq \omega_U \in \Omega^2_c(M; \mathbb{C})$, local in the sense that $\mathrm{supp}\omega_U \subset U$, satisfying $B_{\lambda_j}\omega_U = 0$, a contradiction. Consequently we can suppose $\mathrm{Ker} B_{\lambda_j} \neq \{0\}$ for all $j = 1, \dots, k$. Defining again $W$ as the span of the $v_j$'s with $0 \neq v_j \in \mathrm{Ker} B_{\lambda_j}$ for all $j = 1, \dots, k$ and taking $\mathfrak{S} := \mathrm{End}\, W \cap \mathfrak{R}(M)$ then $A_j \in \mathfrak{S}$ or $R(A_j) \in \mathfrak{S}$. Therefore this $\mathfrak{S} \subset \mathfrak{R}(M)$ is a finite dimensional $*$-subalgebra with the required property for all $\varepsilon > 0$. We conclude that $\mathfrak{R}(M)$ is hyperfinite.

The second consequence is that $\mathfrak{R}(M)$ is a factor. Indeed, although the center of the dense subalgebra of algebraic curvature tensors $C^\infty(M; \mathrm{End}(\wedge^2 T^*M \otimes_{\mathbb{R}} \mathbb{C}))$ is $C^\infty(M; \mathbb{C}) \cdot \mathrm{Id}_{\wedge^2 T^*M \otimes_{\mathbb{R}} \mathbb{C}}$ hence is infinite dimensional, the center of $\mathfrak{R}(M)$ is isomorphic to the one dimensional space $\mathbb{C} \cdot \mathrm{Id}_{\wedge^2 T^*M \otimes_{\mathbb{R}} \mathbb{C}}$ only, taking into account the connectedness of $M$ and the fact that $\mathfrak{R}(M)$ contains all orientation-preserving diffeomorphisms as well. Indeed, if $\Phi : M \to M$ is a diffeomorphism and $R = f \cdot \mathrm{Id}_{\wedge^2 T^*M \otimes_{\mathbb{R}} \mathbb{C}}$ is any diagonal algebraic curvature tensor and $\omega \in \Omega^2_c(M; \mathbb{C})$ then $(\Phi^*(f \cdot \mathrm{Id}_{\wedge^2 T^*M \otimes_{\mathbb{R}} \mathbb{C}})\omega) = \Phi^*(f\omega) = (\Phi^* f)\Phi^*\omega$ but $(f \cdot \mathrm{Id}_{\wedge^2 T^*M \otimes_{\mathbb{R}} \mathbb{C}}\, \Phi^*)\omega = f\Phi^*\omega$ and these coincide if and only if $f : M \to \mathbb{C}$ is a constant by the connectivity of $M$. The center of $\mathfrak{R}(M)$ is therefore one dimensional i.e., $\mathfrak{R}(M)$ is a factor.

Summing up, $\mathfrak{R}(M)$ is a hyperfinite factor possessing a trace. Being trace of a factor, this trace is unique (cf. [1, Proposition 4.1.4]) moreover satisfies $\tau(1) = 1$ as we have seen in Lemma 2.2 consequently $\mathfrak{R}(M)$ is a hyperfinite factor of $\mathrm{II}_1$ type. $\diamond$

**Remark.** Recall that by Lemma 2.1 our von Neumann algebra $\mathfrak{R}(M)$, among its abstract elements, contains simply accessible geometric operators as well namely endomorphisms of the bundle $\wedge^2 T^*M \otimes_{\mathbb{R}} \mathbb{C}$ such that by Lemma 2.3 general elements of $\mathfrak{R}(M)$ can be approximated by these bundle morphisms in the sense of (7). This geometric property allows us to embed $M$ into $\mathfrak{R}(M)$ by the aid of heat kernel techniques.[2] Following [4] take a closed

[2]We are grateful to Ákos Nagy for calling our attention to this as well as to the reference [4] used below.

(i.e. compact without boundary) oriented Riemannian 4-manifold $(M,g)$ and let $\Delta : C^\infty(M;\mathbb{C}) \to C^\infty(M;\mathbb{C})$ be the associated Laplace operator acting on complex-valued functions together with $\{e^{-t\Delta}\}_{t>0}$ the corresponding *heat semigroup*. The heat semigroup is a family of self-adjoint operators possesing a smooth kernel which means that on all $f \in L^2(M;\mathbb{C})$ (constructed by the aid of the metric $g$) the action of the heat semigroup can be written as

$$(e^{-t\Delta}f)(x) = \int\limits_M k_M(t;x,y)f(y)\mathrm{d}y$$

where $k_M(t;x,y)$ is a smooth real function of its variables $t>0$ and $x,y\in M$ moreover $\mathrm{d}y$ denotes the volume form provided by the metric $g$. Therefore the assignment defined by

$$x \longmapsto k_M\left(\tfrac{t}{2};x,\,\cdot\,\right)\mathrm{Id}_{\wedge^2T^*M\otimes_{\mathbb{R}}\mathbb{C}} \quad \text{for all } x\in M \text{ and fixed } t>0 \tag{9}$$

gives rise to a map $i_{M,t} : M \to \mathfrak{R}(M)$. By the aid of [4, Theorem 5] this map is in fact a continuous embedding of $M$ into $\mathfrak{R}(M)$. If $\Phi : M \to M$ is an orientation-preserving diffeomorphism then it acts on functions as $f \mapsto \Phi^*f = f\circ\Phi$ which is unitary with respect to (1) which means that $(\Phi^*)^{-1} = (\Phi^*)^{⁎}$. Hence $\Phi$ acts on the embedded heat kernel regarded as an element of $\mathfrak{R}(M)$ like

$$\begin{aligned} k_M\left(\tfrac{t}{2};x,\,\cdot\,\right)\mathrm{Id}_{\wedge^2T^*M\otimes_{\mathbb{R}}\mathbb{C}} \longmapsto\ & k_M\left(\tfrac{t}{2};\Phi(x),\Phi(\,\cdot\,)\right)\mathrm{Id}_{\wedge^2T^*M\otimes_{\mathbb{R}}\mathbb{C}} \\ &= \Phi^*k_M\left(\tfrac{t}{2};x,\,\cdot\,\right)\Phi^*\mathrm{Id}_{\wedge^2T^*M\otimes_{\mathbb{R}}\mathbb{C}}(\Phi^*)^{⁎} \\ &= \Phi^*\left(k_M\left(\tfrac{t}{2};x,\,\cdot\,\right)\mathrm{Id}_{\wedge^2T^*M\otimes_{\mathbb{R}}\mathbb{C}}\right)(\Phi^*)^{⁎} \end{aligned}$$

thus $M \subset \mathfrak{R}(M)$ can be viewed as the orbit of $\mathrm{Diff}^+(M)$ acting on $\mathfrak{R}(M)$ by unitary automorphisms.

Note that this embedding is not canonical since it depends on a choice of a Riemannian metric $g$ on $M$ and a time parameter $t>0$. Nevertheless this dependence has a benefit: if $\mathscr{H}(M)$ denotes the Hilbert space completion of $\mathfrak{R}(M)$ as in Lemma 2.2 with corresponding scalar product $(\,\cdot\,,\,\cdot\,)$ then applying again [4, Theorem 5] we find that

$$i^*_{M,t}(\,\cdot\,,\,\cdot\,) = g + \tfrac{t}{3}\left(\tfrac{1}{2}\mathrm{Scal} - \mathrm{Ric}\right) + O(t^2) \quad \text{as } t\downarrow 0 \tag{10}$$

that is, the pullback of the canonical metric on $\mathscr{H}(M)$ by this embedding approximates the original metric $g$ on $M$ for very short times. We shall make a good use of this embedding in Section 3.

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. This resembles the reconstruction of space in matrix models, cf. e.g. [12, 14].

We close this section by extracting a smooth 4-manifold invariant out of our efforts so far whose properties will be investigated elsewhere.

**Lemma 2.4.** *Let $M$ be a connected oriented smooth 4-manifold and $\mathfrak{R}(M)$ its von Neumann algebra with trace $\tau$ as before. Then there exists a complex separable Hilbert space $\mathscr{K}(M)$ and a representation $\rho_M : \mathfrak{R}(M) \to \mathfrak{B}(\mathscr{K}(M))$ with the following properties. If $\pi_M : \mathfrak{R}(M) \to \mathfrak{B}(\mathscr{H}(M))$ is the standard representation constructed in Lemma 2.2 then $\{0\} \subseteqq \mathscr{K}(M) \subsetneqq \mathscr{H}(M)$ and $\rho_M = \pi_M|_{\mathscr{K}(M)}$ holds; therefore, although $\rho_M$ can be the trivial representation, it is surely not unitary equivalent to the standard representation. Moreover the unitary equivalence class of $\rho_M$ is invariant under orientation-preserving diffeomorphisms of $M$.*

*Therefore the Murray–von Neumann coupling constant[3] of $\rho_M$, equal to $\tau(P_M) \in [0, 1) \subset \mathbb{R}_+$ where $P_M : \mathscr{H}(M) \to \mathscr{K}(M)$ is the orthogonal projection, is invariant under orientation-preserving diffeomorphisms. Consequently $\gamma(M) := \tau(P_M)$ is a smooth 4-manifold invariant.*

*Proof.* First let us exhibit a representation of $\mathfrak{R}(M)$ by recalling [7, Theorem 3.2]; 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 $(\Sigma, \omega)$ consisting of an (immersed) closed oriented surface $\Sigma \looparrowright M$ and a (not necessarily compactly supported!) 2-form $\omega \in \Omega^2(M; \mathbb{C})$ which is also closed i.e., $\mathrm{d}\omega = 0$. Consider the continuous $\mathbb{C}$-linear functional

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

$F_{\Sigma,\omega} : \mathfrak{R}(M) \to \mathbb{C}$ by continuously extending

$$A \longmapsto \frac{1}{2\pi\sqrt{-1}} \int\limits_{\Sigma} A\omega$$

from $V(M)$ to $\mathfrak{R}(M)$. Let $\{0\} \subseteqq I_{\Sigma,\omega} \subseteqq \mathfrak{R}(M)$ be the closure in the norm $[[\cdot]]$ on $\mathfrak{R}(M)$ of the subset of elements $A \in \mathfrak{R}(M)$ satisfying $F_{\Sigma,\omega}(A^*A) = 0$. In fact for all pairs $(\Sigma, \omega)$ obviously $\{0\} \subsetneqq I_{\Sigma,\omega}$. We assert that $I_{\Sigma,\omega}$ is a norm-closed multiplicative left-ideal in $\mathfrak{R}(M)$ which is non-trivial and independent of $(\Sigma, \omega)$ if $F_{\Sigma,\omega}(1) \neq 0$ and $I_{\Sigma,\omega} = \mathfrak{R}(M)$ hence again independent of $(\Sigma, \omega)$ if $F_{\Sigma,\omega}(1) = 0$.

Consider first 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 $I_{\Sigma,\omega}$ is a multiplicative left-ideal in $\mathfrak{R}(M)$. Concerning its $\omega$-dependence, without loss of generality we can assume that $\omega|_\Sigma$ nowhere vanishes and let $\omega'$ be another closed 2-form with the same property along $\Sigma$ satisfying $F_{\Sigma,\omega'}(1) = 1$; then there always exists an *invertible* and bounded bundle morphism $R \in C^\infty(M; \mathrm{End}(\wedge^2 T^*M \otimes_{\mathbb{R}} \mathbb{C})) \cap \mathfrak{R}(M)^\times$ satisfying $\omega'|_\Sigma = R\omega|_\Sigma$. Then

$$F_{\Sigma,\omega'}(A^*A) = F_{\Sigma,\omega}(A^*AR)$$

hence by the above inequality in the form

$$|F_{\Sigma,\omega'}(A^*A)|^2 \leqq F_{\Sigma,\omega}(A^*A)F_{\Sigma,\omega}((AR)^*(AR))$$

we find $I_{\Sigma,\omega'} \supseteqq I_{\Sigma,\omega}$. Likewise, the equality $F_{\Sigma,\omega}(A^*A) = F_{\Sigma,\omega'}(A^*AR^{-1})$ gives the converse estimate

$$|F_{\Sigma,\omega}(A^*A)|^2 \leqq F_{\Sigma,\omega'}(A^*A)F_{\Sigma,\omega'}((AR^{-1})^*(AR^{-1}))$$

implying $I_{\Sigma,\omega'} \subseteqq I_{\Sigma,\omega}$. Consequently $I_{\Sigma,\omega'} = I_{\Sigma,\omega}$. Concerning the general $(\Sigma, \omega)$-dependence of $I_{\Sigma,\omega}$ we argue as follows. Let $\eta_\Sigma \in \Omega^2(M; \mathbb{R})$ be a nowhere vanishing closed real 2-form representing the Poincaré-dual $[\eta_\Sigma] \in H^2(M; \mathbb{R})$ of $\Sigma \hookrightarrow M$; then referring to the identity $\int_\Sigma \omega = \int_M \omega \wedge \overline{\eta}_\Sigma$ and putting $\omega := \eta_\Sigma$ the functional can be re-expressed as $F_{\Sigma,\eta_\Sigma}(A^*A) = \frac{1}{2\pi\sqrt{-1}}\langle A\eta_\Sigma, A\eta_\Sigma\rangle_{L^2(M)}$ in terms of the indefinite scalar product (1) on $M$. Let $\Sigma' \hookrightarrow M$ be another compact oriented surface and $\omega'$ another closed 2-form such that $F_{\Sigma',\omega'}(1) = 1$. Taking a similar nowhere-vanishing representative $\eta_{\Sigma'} \in \Omega^2(M; \mathbb{R})$ for the Poincaré-dual we can therefore pick again

$R\in C^\infty(M;\mathrm{End}(\wedge^2T^*M\otimes_{\mathbb{R}}\mathbb{C}))\cap\mathfrak{R}(M)^\times$ satisfying $\eta_{\Sigma'}=R\eta_\Sigma$. Then

$$\begin{aligned}F_{\Sigma',\eta_{\Sigma'}}(A^*A)&=\frac{1}{2\pi\sqrt{-1}}\langle A\eta_{\Sigma'},A\eta_{\Sigma'}\rangle_{L^2(M)}\\&=\frac{1}{2\pi\sqrt{-1}}\langle AR\eta_\Sigma,AR\eta_\Sigma\rangle_{L^2(M)}=F_{\Sigma,\eta_\Sigma}((AR)^*(AR))\end{aligned}$$

demonstrating that $I_{\Sigma',\eta_{\Sigma'}}R\subseteqq I_{\Sigma,\eta_\Sigma}$. In the same fashion $F_{\Sigma,\eta_\Sigma}(A^*A)=F_{\Sigma',\eta_{\Sigma'}}((AR^{-1})^*(AR^{-1}))$ convinces us that $I_{\Sigma,\eta_\Sigma}R^{-1}\subseteqq I_{\Sigma',\eta_{\Sigma'}}$. Putting togther all of these we find $I_{\Sigma,\omega}=I_{\Sigma,\eta_\Sigma}=I_{\Sigma',\eta_{\Sigma'}}R=I_{\Sigma',\omega'}R$. Secondly if $(\Sigma,\omega)$ is such that $F_{\Sigma,\omega}(1)=0$ then repeating the previous analysis we obtain that $I_{\Sigma,\omega}$ is a closed multiplicative left-ideal too, however containing $1\in\mathfrak{R}(M)$ as well consequently $I_{\Sigma,\omega}=\mathfrak{R}(M)$. Therefore if $(\Sigma',\omega')$ is another pair with $F_{\Sigma',\omega'}(1)=0$ then obviously $I_{\Sigma,\omega}=I_{\Sigma',\omega'}$ (and equal to $\mathfrak{R}(M)$).

Let us proceed further by exploiting now the observation made in Lemma 2.2 that $\mathfrak{R}(M)$ can be improved to its standard $L^2$ Hilbert space $\mathscr{H}(M)$ with scalar product $(\,\cdot\,,\,\cdot\,)$ therefore $\mathfrak{R}(M)$ acts on $\mathscr{H}(M)$ by the standard representation $\pi_M$ i.e., multiplication from the left. In this way we can regard $\{0\}\subsetneqq I_{\Sigma,\omega}\subseteqq\mathfrak{R}(M)$ as a closed linear subspace $\{0\}\subsetneqq\hat{I}_{\Sigma,\omega}\subseteqq\mathscr{H}(M)$ as well. Consider the orthogonal projection $Q_{\Sigma,\omega}:\mathscr{H}(M)\to\hat{I}_{\Sigma,\omega}$. We can assume $Q_{\Sigma,\omega}\in\mathfrak{R}(M)$ and is acting by $\pi_M(Q_{\Sigma,\omega})\hat{A}=\widehat{Q_{\Sigma,\omega}A}$. Therefore another projection $Q_{\Sigma',\omega'}:\mathscr{H}(M)\to\hat{I}_{\Sigma',\omega'}=\widehat{I_{\Sigma,\omega}R^{-1}}$ acts like

$$\widehat{Q_{\Sigma',\omega'}A}=\big((Q_{\Sigma,\omega}(AR))R^{-1}\big)\widehat{\ }=\widehat{Q_{\Sigma,\omega}A}$$

for all $\hat{A}\in\mathscr{H}(M)$ i.e., $Q_{\Sigma',\omega'}=Q_{\Sigma,\omega}$ hence $I_{\Sigma',\omega'}=I_{\Sigma,\omega}$. Therefore $\{0\}\subsetneqq\hat{I}_{\Sigma,\omega}\subseteqq\mathscr{H}(M)$ is a well-defined closed subspace of $\mathscr{H}(M)$ which is non-trivial if $F_{\Sigma,\omega}(1)\neq0$ and coincides with $\mathscr{H}(M)$ if $F_{\Sigma,\omega}(1)=0$. Take the orthogonal complement $\{0\}\subseteqq\hat{I}^\perp_{\Sigma,\omega}\subsetneqq\mathscr{H}(M)$ with its restricted scalar product $(\,\cdot\,,\,\cdot\,)|_{\hat{I}^\perp_{\Sigma,\omega}}$ and denote the Hilbert space $(\hat{I}^\perp_{\Sigma,\omega},(\,\cdot\,,\,\cdot\,)|_{\hat{I}^\perp_{\Sigma,\omega}})$ simply by $\mathscr{K}(M)$. Note that $\mathscr{K}(M)$ is isomorphic to $\mathscr{H}(M)/\hat{I}_{\Sigma,\omega}$ as a complex vector space however its Hilbert space structure might be different from the one provided by (some) $F_{\Sigma,\omega}$ on the quotient as in the usual GNS construction. Since $\{0\}\subsetneqq I_{\Sigma,\omega}\subseteqq\mathfrak{R}(M)$ is a multiplicative left-ideal and the scalar product on $\mathscr{H}(M)$ satisfies $(\widehat{AB},\widehat{C})=(\hat{B},\widehat{A^*C})$ the standard representation $\pi_M:\mathfrak{R}(M)\to\mathfrak{B}(\mathscr{H}(M))$ restricts to a representation on $\{0\}\subseteqq\mathscr{K}(M)\subsetneqq\mathscr{H}(M)$. This is either a unique non-trivial representation if $\mathscr{K}(M)\neq\{0\}$ (provided by a functional over $M$ with $F_{\Sigma,\omega}(1)\neq0$ if exists), or the trivial one if $\mathscr{K}(M)=\{0\}$ (provided by a functional with $F_{\Sigma,\omega}(1)=0$ which always exists). Keeping these in mind, for a given $M$ we define

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

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

From the general theory [1, Chapter 8] we know that the Murray–von Neumann coupling constant of $\rho_M$ depends only on the unitary equivalence class of $\rho_M$ and if $P_M : \mathscr{H}(M) \to \mathscr{K}(M)$ is the orthogonal projection then $P_M \in \mathfrak{R}(M)$ and the coupling constant is equal to $\tau(P_M) \in [0,1]$. However observing that $\rho_M$ is surely not isomorphic to $\pi_M$ since $\hat{I}_M$ is never trivial the case $\tau(P_M) = 1$ is excluded i.e., in fact $\tau(P_M) \in [0,1)$. Finally, consider an orientation-preserving diffeomorphism $\Phi : M \to M$. It follows from Lemma 2.1 that it induces a unitary inner automorphism $A \mapsto \Phi^* A (\Phi^*)^{*}$ of $\mathfrak{R}(M)$. Moreover it transforms $I_{\Sigma,\omega}$ into $I_{\Sigma',\omega'} = I_{\Phi(\Sigma),\Phi^*\omega}$ hence $F_{\Sigma,\omega}(1) = 0$ if and only if $F_{\Sigma',\omega'}(1) = 0$ consequently the Hilbert space $\mathscr{K}(M)$ is invariant under $\Phi$. We obtain that $\Phi$ transforms $\rho_M$ into a new representation $\Phi^*\rho_M(\Phi^*)^{*}$ on $\mathscr{K}(M)$ which is unitary equivalent to $\rho_M$.

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

**Remark.** Note that $\gamma(M) = 0$ corresponds to the situation when $\rho_M$ is the trivial representation on $\mathscr{K}(M) = \{0\}$. To avoid this we have to demand $F_{\Sigma,\omega}(1) \neq 0$ which by the closedness assumptions on $\Sigma$ and $\omega$ is in fact a *topological condition*: it is equivalent that $F_{\Sigma,\omega}(1) = \frac{1}{2\pi\sqrt{-1}} \int_\Sigma \omega = \langle [\Sigma], [\omega] \rangle \in \mathbb{C}$ 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 $H_2(M;\mathbb{C}) = H_2(M;\mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{C} = \{0\}$ (or equivalently, $H^2(M;\mathbb{C}) = \{0\}$). Consequently $\gamma(M) = 0$ in particular for $M = S^4, \mathbb{R}^4$ and unfortunately $\gamma(M) = 0$ if $M = R^4$ is any exotic (or fake) $\mathbb{R}^4$, too.

**Lemma 2.5.** *(Excision principle.) Let $M$ be a connected oriented smooth 4-manifold and $\emptyset \subsetneqq Y \subset M$ a submanifold so that $M \setminus Y \subsetneqq M$ 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 2nd 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*

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

*Proof.* Regarding the first 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 $\mathrm{II}_1$ factor. Applying the extension-by-zero operation on compactly supported 2-forms the embedding $M \setminus Y \subsetneqq M$ induces $\Omega^2_c(M \setminus Y; \mathbb{C}) \subsetneqq \Omega^2_c(M; \mathbb{C})$ hence $V(M \setminus Y) \subsetneqq V(M)$ for the corresponding endomorphisms; taking closures we eventually come up with a subfactor $\mathfrak{R}(M \setminus Y) \subsetneqq \mathfrak{R}(M)$ with some Jones index $[\mathfrak{R}(M) : \mathfrak{R}(M \setminus Y)]$. By definition $\mathfrak{R}(M \setminus Y)$ acts via $\rho_{M\setminus Y}$ on $\mathscr{K}(M \setminus Y)$ and $\mathfrak{R}(M)$ acts via $\rho_M$ on $\mathscr{K}(M)$. By assumption $M \setminus Y \subsetneqq M$ induces an isomorphism on the 2nd homology hence $\mathscr{K}(M \setminus Y)$ and $\mathscr{K}(M)$ are simultaneously trivial or not; moreover $Y$ has zero 4 dimensional Lebesgue measure therefore $h^+(M \setminus Y) \oplus h^-(M \setminus Y) = h^+(M) \oplus h^-(M)$ for the completion of the 2-form spaces with respect to any splitting (2). This implies $\mathscr{K}(M \setminus Y) = \mathscr{K}(M)$ because otherwise for instance the element $A\omega \in h^+(M) \oplus h^-(M)$ with some $0 \neq A \in \mathscr{K}(M \setminus Y)^\perp \subset \mathscr{K}(M)$ and $0 \neq \omega \in \Omega^2_c(M; \mathbb{C})$ would be acted upon trivially by $\mathfrak{R}(M \setminus Y)$, a contradiction. Consequently $\rho_{M\setminus Y}$ and $\rho_M$ are both representations on $\mathscr{K}(M)$. We know that $\gamma(M)$ is the $\mathfrak{R}(M)$-dimension of $\mathscr{K}(M)$ i.e., $\gamma(M) = \dim_{\mathfrak{R}(M)} \mathscr{K}(M)$ and likewise $\gamma(M \setminus Y) = \dim_{\mathfrak{R}(M\setminus Y)} \mathscr{K}(M)$. Therefore they are related as $\gamma(M \setminus Y) = [\mathfrak{R}(M) : \mathfrak{R}(M \setminus Y)]\gamma(M)$. However

$$[\mathfrak{R}(M) : \mathfrak{R}(M \setminus Y)] = \left[(\rho_{M\setminus Y}(\mathfrak{R}(M \setminus Y)))' : (\rho_M(\mathfrak{R}(M)))'\right] = 1$$

yielding $\gamma(M \setminus Y) = \gamma(M)$.

Concerning the second assertion note that the $\gamma$-invariant is a well-defined map from (the category) $\mathscr{M}$ of all orientation-preserving diffeomorphism classes of connected, oriented smooth 4-mamifolds into the real interval $[0,1) \subset \mathbb{R}$. But $\mathscr{M}$ forms a commutative semigroup with unit $S^4$ under the connected sum operation $\#$. That is, if $X, Y, Z \in \mathscr{M}$ and $S^4 \in \mathscr{M}$ 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 $M, N \in \mathscr{M}$ with their connected sum $M\#N \in \mathscr{M}$ 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 $\gamma(M\#N) =: \gamma(M) \bullet \gamma(N)$. The $\bullet$-operation therefore satisfies $\gamma(X) \bullet \gamma(Y) = \gamma(Y) \bullet \gamma(X)$ 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 $\gamma : (\mathscr{M}, \#) \to ([0,1), \bullet)$ is a unital semigroup homomorphism. Quite surprisingly there exists a unique structure of this kind on $[0,1)$ yielding the shape for $\gamma(M) \bullet \gamma(N)$ as stated. $\diamondsuit$

**Remark.** Since $0 \leqq \gamma(S^4) < 1$ the gluing principle applied to $S^4 \cong S^4 \# S^4$ shows that $\gamma(S^4) = 0$ hence by the excision principle we obtain $\gamma(\mathbb{R}^4) = \gamma(S^4 \setminus \{\infty\}) = \gamma(S^4) = 0$ too, as we already know. More generally one can demonstrate by techniques from 4-manifold theory and the gluing principle the following two things. First, if $M'$ and $M''$ are connected smooth 4-manifolds which are homeomorphic then (unfortunately) $\gamma(M') = \gamma(M'')$. Second, put $R_0(x) := 0, R_1(x) := x, \dots, R_k(x) := \frac{x + R_{k-1}(x)}{1 + xR_{k-1}(x)}, \dots$ and $y := \gamma(\mathbb{C}P^2) = \gamma(\overline{\mathbb{C}P^2})$; note that surely $y \neq 0$ since $H^2(\mathbb{C}P^2; \mathbb{C}) \ncong \{0\}$. Then if $M$ is any connected, simply connected, closed smooth 4-manifold then $\gamma(M) = R_n(y)$ with some $n \in \{0\} \cup \mathbb{N}$; for instance $\gamma(S^4) = R_0(y), \gamma(\mathbb{C}P^2) = R_1(y), \gamma(\mathbb{C}P^1 \times \mathbb{C}P^1) = R_2(y)$ and $\gamma(K3) = R_{22}(y)$.

*Proof of Theorem 1.1.* This theorem follows from Lemmata 2.1, 2.2, 2.3, 2.4 and 2.5. $\diamondsuit$

## 3. Physical interpretation

In this section we cannot resist the temptation and 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. Accepting Theorem 1.1 this construction is simple, self-contained and is based upon reversing its content or more generally the approach of Section 2. Namely, here in Section 3 not *one particular 4-manifold*—physically regarded as a particular classical space-time—but *the unique hyperfinte* $\mathrm{II}_1$ *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.

1. *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 hyperfinite factor von Neumann algebra of type $\mathrm{II}_1$ 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 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.

Two remarks are in order. The first is: 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 Lemma 2.3 that $\mathfrak{R}$ contains a dense (in the sense of (7)) subalgebra whose members can be interpreted as bounded local (complexified) algebraic curvature tensors along a smooth manifold whose real dimension is precisely *four*; consequently up to arbitrary finite experimental accuracy we can *demand* that the abstract bounded linear operators in $\mathfrak{R}$ be four dimensional bounded curvature tensors. Recall that in classical general relativity *local* gravitational phenomena are caused by the curvature of space-time; hence by demanding $\mathfrak{R}$ to consist of local observables the corresponding quantum theory is declared to be a *four dimensional quantum theory of pure gravity*. In this way we fulfill the *Heisenberg dictum* that a quantum theory should completely and unambigously be formulated in terms of its local physical observables. In modern understanding by a *physical theory* one means a two-level description of a certain class of 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. It is important to point out that in this context our quantum theory is not plainly a mathematical theory anymore but a physical theory. This is because the bare mathematical structure $\mathfrak{R}$ (together with a representation $\pi$ 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, the demand of being conformal anomaly free) and it turns out to be much higher than the phenomenological dimension of space-time.

The second remark concerns the priority between $\mathfrak{R}$ and $\mathscr{H}$ i.e., observables and states. It is an evergreen question in quantum field theory whether the operator algebra of its observables or the Hilbert space of its states should be considered as the primordial structure when constructing it? For example, in the conventional quantum mechanics or Wightman axiomatic quantum field theory the state space is considered to be fundamental while in the more recent algebraic quantum field theory approach [9] the algebra of observables is declared to be the fundamental structure and then the problem arises how to find that representation of this algebra which describes the physical states of the sought quantum theory. In our approach here the following special features occur. The first observation concerns $\mathfrak{R}$ and $\mathscr{H}$ as bare vector spaces: as a by-product of Lemma 2.2, the inclusion $\mathfrak{R} \subset \mathfrak{B}(\mathscr{H})$ in fact stems from a representation $\pi : \mathfrak{R} \to \mathfrak{B}(\mathscr{H})$ which is simply the left multiplication of $\mathfrak{R}$ on itself; hence this so-called *standard representation* is continuous, irreducible, faithful and $\mathfrak{R}$ and $\mathscr{H}$ are isomorphic as complete complex vector spaces. Our second observation concerns $\mathfrak{R}$ and $\mathscr{H}$ as gauge group modules: the unit $1 \in \mathfrak{R}$ is invariant under the gauge group but its image $\hat{1} \in \mathscr{H}$ is a cyclic and separating vector hence is not invariant under the gauge group; consequently $\mathfrak{R}$ and $\mathscr{H}$ are *not* isomorphic as gauge group modules.[4] Nevertheless we can see that this representation $\pi$ meets physical demands and we encounter here a sort of *field-state correspondence* as in conformal field theory. In spite of this "self-duality" however, we will see shortly that for our purposes it will be more convenient to consider $\mathfrak{R}$ (and not $\mathscr{H}$) as the primordial structure.

2. *Observables as the universal space of all space-times.* Taking into account the *Remark* after Lemma 2.3 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. More precisely: given an oriented closed 4-manifold $M$, by the aid of a fixed Riemannian metric $g$ on it and a real parameter $t > 0$ we can identify the point $x \in M$ with a self-adjoint scalar operator of $\mathfrak{R}$ as in (9); moreover the metric can also be recovered from this embedding in the sense of (10) i.e. by pulling back the standard scalar product on $\mathscr{H}$. Orientation-preserving diffeomorphisms of $M$ interchange its points and act on the corresponding operators in $\mathfrak{R}$ by unitary inner automorphisms i.e. $\mathrm{Diff}^+(M)$ embeds into $\mathrm{U}(\mathscr{H}) \cap \mathfrak{R} \subsetneqq \mathrm{Aut}\mathfrak{R}$. Reformulating this in a more geometric way we can say that classical space-times appear as special orbits within $\mathfrak{R}$ of its gauge group.

All the operators in $\mathfrak{R}$ representing geometric points are of the form (9) i.e. are scalar operators hence commute. Consequently the full non-commutative algebra $\mathfrak{R}$ is not exhausted by operators representing points of

[4]In accordance with the notation in Section 2 the norm on $\mathfrak{R}$ is $[[\cdot]]$ and the scalar product on $\mathscr{H}$ is $(\cdot\, ,\, \cdot)$ with induced equivalent norm hence also denoted by $[[\cdot]]$ yielding $\mathfrak{R} \cong \mathscr{H}$ as complex complete vector spaces however not as gauge group modules; we write therefore $A \in \mathfrak{R}$ but $\hat{A} \in \mathscr{H}$, cf. Lemma 2.2 and the *Remark* after it.

space-time; it certainly contains much more operators—e.g. various projectors which are not of geometric origin—therefore this “universal quantum space-time” is more than a bunch of all classical space-times. We remark that this operator algebraic “enumeration” of all (closed, oriented) smooth 4-manifolds might shed a new light onto the predicted non-computability issues of quantum gravity [8].

As a comparison, note that in *algebraic quantum field theory* [9] 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 $\emptyset \subsetneqq U \subsetneqq M$. 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.

3. *Examples of observables and fields.* Let us take a closer look of the elements of $\mathfrak{R}$ and $T_1\mathfrak{R} \supset \mathfrak{R}$. Taking into account the two items above concerning the interpretation of the mathematical results of Section 2 we agree to identify the 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 particular if $(M, g)$ is a solution of the classical (Riemannian or Lorentzian) Einstein's equation with (complexified) curvature tensor $R_g$ as in (8) which is an element of $\mathfrak{R}$ that is, $[[R_g]] < +\infty$ then the full classical solution $(M, g)$ can be identified with a single geometric observable $R_g \in \mathfrak{R}$ in our quantum theory. For instance smooth solutions like the flat $\mathbb{R}^4$ or more generally, the 27 connected, compact orientable flat 4-geometries (in total there exist 74 connected compact flat Riemannian 4-manifolds, cf.[11]) certainly satisfy this boundedness condition and in fact give rise to the same unique zero observable $R_g = 0 \in \mathfrak{R}$. Hence the mapping $(M, g) \mapsto R_g$, when $(M, g)$ runs through all bounded 4-geometries, is not injective in general. It is important to note that, on the 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, etc. give rise to not *observables* in $\mathfrak{R}$ but rather *fields* in $T_1\mathfrak{R} \supset \mathfrak{R}$.

4. *Questions and answers.* First let us clarify what the *answers* in this quantum theory are because this is easier. Staying within the orthodox framework i.e., the *Coppenhagean 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 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 $\mathrm{II}_1$ factor von Neumann algebra $\mathfrak{R}$. Note that in order not to be short sighted, at this level of generality we require neither $A \in \mathfrak{R}$ to be self-adjoint nor $B \in \mathfrak{R}$ to be positive and normalized (however these can be imposed if they turn out to be necessary) 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 fully at the semantical level. The orthodox approach says that a question should 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 is obtained through a *measurement*. Let us make a short digression concerning the measurement. Should we assume that, after performing the physical experiment designed to answer the question above, 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 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 provisory 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 Lemma 2.2 that if $T \in \mathfrak{R}$ is an operator over some space $X$ then for any error term $0 < \varepsilon$ there exists $\omega \in \Omega^2_c(X;\mathbb{C})$ with the property $1 = (\omega,\omega)_{L^2(X)} := \langle\omega^+,\omega^+\rangle_{L^2(X)} - \langle\omega^-,\omega^-\rangle_{L^2(X)}$ given by integration (1) over $X$ such that $\left|\frac{1}{2}\big(T\omega,\omega\big)_{L^2(X)} + \frac{1}{2}\overline{\big(T^*\omega,\omega\big)}_{L^2(X)} - \tau(T)\right| \leqq \varepsilon$ hence $\tau$ is a highly non-local and complex-valued object. This approximation allows to compute $\tau(R_gR_h)$ by integrating $R_gR_h\omega\wedge\overline{\omega}$ along $X = M\cap N$. Note that making use of the aforemtioned embeddings of $M, N$ into $\mathfrak{R}$ taking intersection 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 $|\tau(R_gR_h)| \approx 0$ 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).

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 $c=\hbar=G=1$). This dynamics is therefore trivial leading to the usual "problem of time" in general relativity [2, 3, 6]. Nevertheless, quite interestingly, this dynamics also coincides with the modular dynamics introduced by Connes and Rovelli [6] because it is associated with the *tracial* state $\tau$ on $\mathfrak{R}$ having the identity as modular operator. The *rest energy* or *mass* of a state $B\in\mathfrak{R}$ is defined to be the expectation value of the modular Hamiltonian $1\in\mathfrak{R}$ in this state more precisely $m(B) := \tau(1B)\in\mathbb{C}$ (cf. the smooth invariant $\gamma(M) = \tau(1P_M)$ in Lemma 2.4). If $B$ is non-negative and self-adjoint (as an operator on $\mathscr{H}$) then it has non-negative real number energy because (cf. [16, Sect. VII.104]) in this case there always exists a unique self-adjoint operator $B^{\frac{1}{2}}\in\mathfrak{R}$ satisfying $(B^{\frac{1}{2}})^*B^{\frac{1}{2}} = (B^{\frac{1}{2}})^2 = B$ consequently $m(B) = \tau(B) = \tau\big((B^{\frac{1}{2}})^*B^{\frac{1}{2}}\big) = [\![B^{\frac{1}{2}}]\!]^2 \geqq 0$.

We conclude with a comment on the *cosmological constant problem* namely its *small* but surely *positive* value $\Lambda\approx 0.7$ found in recent observations [15]. Consider a special state which is the curvature operator $R_{g_U}$ of the restricted Friedman–Lemaître–Robertson–Walker metric [18, Chapter 5] modeling the late-time nearly flat and "dust dominated" portion $(U, g_U)$ of the cosmologial space-time we live in. It follows that $R_{g_U}$ has no Weyl component and its non-zero Ricci component is diagonal consequently when considered as a map $R_{g_U} : \Omega^2_c(U;\mathbb{C})\to\Omega^2_c(U;\mathbb{C})$ it gives rise to an element

$R_{g_U} \in \mathfrak{R}$ which is a non-negative self-adjoint operator.[5] Consequently its energy content $m(R_{g_U}) = \tau(1R_{g_U}) \geqq 0$. However $R_{g_U} \neq 0$ and $1 \neq 0$ within $\mathfrak{R}$ imply that in fact $m(R_{g_U}) > 0$ moreover $m(R_{g_U}) \approx 0$ because according to our experience $|R_{g_U}| \approx 0$. Therefore $m(R_{g_U})$ is an energylike *small positive* number hence it is challenging to set $\Lambda := m(R_{g_U})$ in this model i.e. identify the cosmological constant (or dark energy) with the full quantum gravitational energy of the space-time portion $(U, R_{g_U})$. This offers a qualitative understanding of the experimental value of the cosmological constant.

5. *Recovering classical general relativity.* Being this model a genuine quantum theory its matching with known classical field theories is not expected to be a simply limiting process (in the sense of sending some parameter(s) to some special value(s) like $\hbar \to 0$, etc.). Rather, following Haag [9], we expect to recover general relativity by seeking representations $\pi_{\text{classical}}$ of the once and for all given observable algebra $\mathfrak{R}$ which are *different* from (i.e., not unitarily equivalent to) the one $\pi = \pi_{\text{quantum}}$ (the so-called standard representation of $\mathfrak{R}$) used so far. These $\pi_{\text{classical}}$ representations are naturally exhibited by Lemma 2.4. Namely, given a smooth 4-manifold $M$ a Hilbert space $\mathscr{K}(M) \subset \mathscr{H}$ and a reprsentation $\rho_M : \mathfrak{R} \to \mathfrak{B}(\mathscr{K}(M))$ can be assigned to it. These representations already possess non-trivial modular dynamics at finite temperatures in the sense of [6]. We can regard their superabundance as describing the spontaneously broken classical limits of the unique quantum theory, namely the *quantum representation* $\pi : \mathfrak{R} \to \mathfrak{B}(\mathscr{H})$ provided by the trace $\tau : \mathfrak{R} \to \mathbb{C}$ possessing trivial modular dynamics at infinite temperature in the sense of [6]. Therefore the passage from the quantum to the classical regime represents some sort of "cooling down" process in this theory, resembling history after the Big Bang.

## References

[1] Anantharaman, C., Popa, S.: *An introduction to* $\mathrm{II}_1$ *factors*, perprint, 331 pp., available at `http://www.math.ucla.edu/~popa/books.html`;

[2] Ashtekar , A. (ed.): *Loop quantum gravity. The first 30 years*, 100 Years of General Relativity **4**, World Scientific, New Jersey (2017);

[3] Barbour, J.: *The end of time*, Oxford Univ. Press, Oxford (1999);

[5]Indeed, the curvature operator of the FLRW metric is simply scalar multiplication on $\Omega^2_c(U;\mathbb{C})$ by its scalar curvature, cf. (8). However by [18, Equations 5.2.10, 5.2.14 and 5.2.15] this scalar curvature is equal to $8\pi\rho - 24\pi P - \frac{6k}{a^2}$ which is positive in the dust dominated $P = 0$ regime for all $k = -1, 0, +1$ (this is not obvious from this formula for $k = +1$ but is true, cf. [18, Chapter 5]).

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

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

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

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

[8] Geroch, R.P., Hartle, J.: *Computability and physical theories*, Found. Phys. **16**, 533–550 (1986);

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

[10] Hedrich, R.: *String theory—nomological unification and the epicycles of the quantum field theory paradigm*, arXiv: `1101.0690 [physics]` (preprint), 23pp. (2011);

[11] Hillman, J.: *Four-manifolds, geometries and knots*, Geometry and Topology Monographs **5**, Geometry and Topology Publications, Warick (2002);

[12] Landi, G., Lizzi, F., Szabo, R.J.: *From large $N$ matrices to the noncommutative torus*, Commun. Math. Phys. **217**, 181–201 (2001);

[13] Percacci, R.: *An introduction to covariant quantum gravity and asymptotic safety*, 100 Years of General Relativity **3**, World Scientific, New Jersey (2017);

[14] Rieffel, M.A.: *Matricial bridges for "Matrix algebras converge to the sphere"*, Contemporary Mathematics **671**, 209–233 (2016);

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

[16] Riesz, F., Szőkefalvi-Nagy, B.: *Leçons d'analyse fonctionnelle*, Gauthier-Villars, Paris, Akadémiai Kiadó, Budapest (1965);

[17] 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);

[18] Wald, R.M.: *General relativity*, Univ. Chicago Press, Chicago (1984);

[19] von Weizsäcker, C.F.: *Der zweite Hauptsatz und der Unterschied von Vergangenheit und Zukunft*, Ann. Physik **36**, 275–283 (1939), reprinted in: *Die Einheit der Natur*, pp. 172–182, Hanser, München, Wien (1971).

Department of Geometry, Institute of Mathematics
Budapest University of Technology and Economics
Műegyetem rkp. 3., H-1111 Budapest, Hungary
*E-mail address*: etesi@math.bme.hu