Kielnarowa Review (2025), **2**, 1–95

Kielnarowa
REVIEW

JOURNAL

Contents

1. The Architecture of Relativistic Space-Time: A Critical Review

Michał Eckstein and Michał Heller

Kielnarowa Review (2025), **2**, 3–14
doi: https://doi.org/10.57656/kr-2025-0001

Kielnarowa
REVIEW

ARTICLE

The Architecture of Relativistic Space-Time: A Critical Review

Michał Eckstein *†‡ and Michał Heller ‡¶

†Institute of Theoretical Physics, Jagiellonian University, Łojasiewicza 11, 30-348 Kraków, Poland
‡Copernicus Center for Interdisciplinary Studies, ul. Szczepańska 1, 31-016 Kraków, Poland
¶Vatican Observatory, V-00120 Vatican City State
*Corresponding author. Email: michal.eckstein@uj.edu.pl

Abstract

In this critical review, we try to extract from the rich space-time architecture those aspects that are responsible for time and causality. We start with a short review of some attempts to axiomatise the theory of relativity. We focus on the classical approach by Ehlers, Pirani and Schild. While not uncontroversial, it shows — from a quasi-operational perspective — the relations between various space-time substructures. Then we analyse temporal and causal threads in the fabric of the contemporary space-time theory. The so-called causal structure of space-time, nowadays elaborated in great detail, has entered into relativistic physics as a part of its theoretical tool-kit, and can be regarded as complementary with respect to the method of axiomatisation. We argue that a local time flow and an elementary concept of causality are a necessary minimum upon which further richer and richer space-time substructures are built. Our philosophical message is that the relativistic model supports neither the attempts à la Hume to reduce causality to time order, nor the endeavours à la Leibniz to derive time from causal order. Instead, the theory of relativity pictures space-time as a rich edifice which, when looked upon from various angles, displays different logical patterns of its superstructure.

Keywords: space-time, causality, time, general relativity

1. Introduction

In 1950, Cambridge University Press published a small book by Erwin Schrödinger entitled “Space-Time Structure” (Schrödinger 1985). It can be considered a forerunner of the later global approach to the geometry of space-time, which a decade later led to the famous theorems on the existence of singularities (Penrose 1965). According to Schrödinger, Einstein’s theory of space-time is, to use today’s phraseology, a “*theory of everything*”. In the introduction to this book he wrote (Schrödinger 1985): “A four-dimensional continuum endowed with a certain intrinsic geometric structure, a structure that is subject to certain inherent purely geometrical laws, is to be an adequate model or picture of the ‘real world around us in space and time’ with all that it contains and including its total behaviour, the display of all events going on in it.” Schrödinger expressed his belief that it would be possible to conceive other forces, besides gravity, as “purely geometrical restrictions on the structure of space-time”. He himself explored only three levels or aspects of this structure: its general invariance, its connection and metric substructures.

Thanks to the research direction initiated by Schrödinger, we can now reconstruct the rich architecture of space-time, with the wealth of its substructures and their corresponding physical interpretations. Although a complete unification of quantum theory with the geometric structure of space-time is still missing, large areas of physics are indeed encoded in space-time geometry. The purpose of this critical review is to extract from this rich space-time architecture those aspects that are associated with the concepts of time and causality. It seems that the very possibility of doing physics as a science based on measurements is essentially related to causal relationships and measurements of time.

The advent of the theory of relativity has shed a completely new light on the role of time and causality in the structure of modern physic. Since the structural construction of a physical theory is best visible in its axiomatisation, we start, in Section 2, with a short review of some early attempts to axiomatise theory of relativity. Then, in Section 3, we focus on the classical axiomatisation proposed by Elhers, Pirani, and Schild (1972). We give a brief overview of this axiomatic system, which is now well-established, and thoroughly discussed in the philosophical literature. While not uncontroversial, it nicely discloses the structuring of space-time — i.e. shows mutual relations between its various substructures. We quote some critical remarks on the EPS axiomatics from (Linnemann and Read 2021; Adlam, Linnemann, and Read 2022a, 2022b) and summarise its current empirical status.

In Section 4 we give a selected account of what physicists call causal or cone structure of space-time. Our approach consists not so much in analysing separate problems involved in this structure, but rather in investigating temporal and causal threads in the fabric of contemporary space-time theory. The causal structure of space-time has been worked out in great detail — see, for instance, (Carter 1971; Hawking and Ellis 1973) or (Minguzzi and Sánchez 2008; Minguzzi 2019). In (Carter 1971, Sec. 12) one can find a (non-constructivist) axiomatisation of causal spaces, called there "etiological spaces". This approach entered into relativistic physics as a part of its theoretical tool-kit. The fact that this tool-kit was used by Woodhouse (1973) to improve the EPS axiomatisation, shows that these two approaches are, in a sense, complementary.

In Section 5, we argue that, from the geometric point of view, the notion of a local time flow and an elementary concept of causality are the minimum upon which a further hierarchy of stronger and stronger conditions are built. Finally, in Section 6 we revisit the traditional philosophical standpoints à la Leibniz (causality implies time) and à la Hume (time implies causality) in view of the contemporary relativistic model of space-time. We argue that neither of these ideas is reflected within theory of relativity. In contrast, the relativistic space-time is a rich holistic structure, in which the concepts of time and causality are irreducibly interwoven.

2. Some early attempts to axiomatise relativity theory

Perhaps the earliest axiomatisation of relativity theory is attributed to Robb (1914) who proposed an axiomatic system based on the concept of "conic order", and was able to derive both topological and metric properties of space-time from the "invariant succession of events". In the axiomatic systems of both Carnap (1925) and Reichenbach (1924) it is the temporal order that is reduced to the causal order. The same is true for the Mehlberg's system (Mehlberg 1980), which was elaborated within the broader setting of an interdisciplinary study of the causal theory of time[1].

The latter three authors were associated with logical empiricism and, in agreement with its philosophical ideology, aimed at clarifying logically the conceptual situation in the theory of relativity, which was extensively discussed at that time. In doing so, they *a priori* eliminated, with the help of their axioms, some "pathological situations". These included the logical paradoxes induced by causal loops. What they did not take into account was that the later development of physics might need such pathologies (see e.g. (Deutsch 1991; Lloyd et al. 2011)). In 1949 Gödel (1949) published his solution to Einstein's field equations with closed timelike curves (soon after more solutions with similar "time anomalies;' were found). A few years later the first general theorem concerning the global structure of space-time was proved stating that every compact space-time must contain closed timelike curves (Bass and Witten 1957).

1. Although Mehlberg's book appeared in 1980, it is based on his works dating back to 1935 and 1937.

A different strategy for the axiomatisation of space-time was proposed by Elhers, Pirani, and Schild (1972) (in the following, we refer to it as to the EPS axiomatisation). In contrast to the conceptual approaches of Reichenbach and Carnap, the EPS axiomatisation is based on mathematical structures inherent in general relativity and the operational character of physics.

It is instructive at this point to recall the distinction, in the spirit of Carnap (1967) and Reichenbach (1924), between constructive axiomatisation and deductive axiomatisation. The former "builds on a basis of empirically supposedly indubitable posits" (Linnemann and Read 2021), whereas the latter "proceeds semantically in a linear, non-circular fashion" (ibid.), i.e. its primary purpose is to establish the logical order of inference. For obvious reasons, we are interested in the constructivist axiomatisation of space-time; however, in the case of relativistic space-time, as we shall see, it is not fully achievable. So we have to be content with what physicists call a quasi-operational approach. The EPS axiomatisation belongs to this type of axiomatic systems.

3. Space-time architecture

From the mathematical point of view, a space-time in general relativity is a pair (M, g), where M is a four-dimensional differential manifold, and g a Lorentzian metric defined on it. Furthermore, it is typically assumed that (M, g) is connected and time-oriented (see (Bielińska and Read 2023) for a nice overview on the time-orientability issue).

We say that the manifold M carries a Lorentzian structure. This structure not only contains several other mathematical substructures, which interact with each other creating a subtle hierarchical edifice, but also admits, on each of its levels, a physical interpretation, making out of the whole one of the most beautiful models of contemporary physics. We shall briefly describe the 'internal design' of this model. Our analysis will be based on the classical paper by Elhers, Pirani, and Schild (1972), in which the authors presented a quasi-operationistic axiomatic system for the Lorentz structure of space-time showing both its mathematical architecture and physical meaning.

The building blocks (primitive concepts) of this axiomatisation are: (1) a set $M = (p, q, \ldots)$, the elements of which are called events; and two collections of subsets of M: (2) $L = (L_1, L_2, \ldots)$, the elements of which are called histories of light rays or of photons (light rays or photons, for the sake of brevity); (3) $P = (P_1, P_2, \ldots)$, the elements of which are called histories of test particles or of observers (particles or observers, for brevity).

The axioms describe the following scenario: A light ray is sent from an event p, situated on a particle history P_1, and received at an event q, situated on a particle history P_2 (a message from p to q). The message can be reflected at q and sent to the event p', situated on the particle history P_1 (an echo on P_1 from P_2). By a suitable combination of messages and echoes one can ascribe four coordinates to any event, and construct local coordinate systems. A bit of mathematical gymnastics, sanctioned by suitable axioms, allows one to organise the set of events, M, into a differential manifold (with the usual manifold topology).

The next set of axioms equips the manifold M with a conformal metric. This is the usual metric (with Lorentzian signature) defined up to a multiplicative factor (the conformal factor). This metric allows one to distinguish timelike, null (lightlike) and spacelike histories (curves), and completely determines the geometry of null-curves.

To determine the geometry of timelike curves one needs additional axioms defining the so-called projective structure. Its task is to determine a distinguished class of timelike curves that are physically interpreted as representing histories of particles (or observers) moving with no acceleration (freely falling particles or observers).

Conformal and projective structures are, in principle, independent and they need to be synchronised. This is done with the help of suitable axioms, which enforce null geodesics passing through an event p to form the light cone at p, and projective timelike curves to fill in the interior of this light cone. If this is the case, we speak of the Weyl structure.

In a Weyl space (a manifold equipped with the Weyl structure), there is a natural method to define length – the arc length – along any timelike curve. Such a length is interpreted as a time interval measured by a clock carried by the corresponding particle (proper time of the particle). However, proper times of different particles are unrelated; to synchronise them a suitable metric must be introduced on M. This can be done with the help of the following axiom: Let $p_1, p_2, \ldots$ be equidistant events on the history of a freely falling particle P, and let them be correspondingly simultaneous with events $q_1, q_2, \ldots$ on the history of a freely falling particle Q. Simultaneity is understood here in the Einstein sense: two events, p and q, are simultaneous if an observer situated half-way between them sees the light signal emitted by p and q at the same instant as shown by the observer's clock. The metric structure is established if the events $q_1, q_2, \ldots$ on Q are also equidistant. The metric g, constructed in this way, contains in itself the Weyl structure; it is, therefore a Lorentz metric. This completes the construction of the relativistic model of space-time.

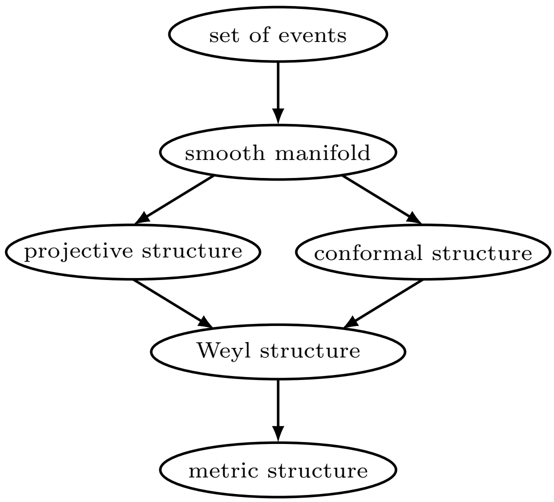


Figure 1. A diagram illustrating the subsequent layers of axiomatisation of general relativity following (Elhers, Pirani, and Schild 1972).

The EPS axiomatisation was recently thoroughly analysed in (Linnemann and Read 2021; Adlam, Linnemann, and Read 2022a, 2022b). The authors showed its limitations and gaps, and proposed some amendments. Here we quote only a few critical remarks, which seem relevant from the point of view of our purposes.

- In the EPS axiomatisation the topological Hausdorff condition is not explicitly postulated. However, when using a smooth manifold as a space-time model, the Hausdorff condition is usually assumed (see e.g. (Hawking and Ellis 1973, p. 13)). The consequences of dropping this condition are studied in detail in (Adlam, Linnemann, and Read 2022b) in the context of possible 'quantum' versions of the EPS axiomatics, in particular in relation to branching space-times (Luc 2020; Luc and Placek 2020; Belnap, Müller, and Placek 2021).
- The axiom assuring the existence of messages and echoes is not valid globally in strong gravitational fields (Linnemann and Read 2021; Pfister and King 2015).
- Lorentz metric is sufficient, but not necessary for the construction of a standard clock measuring the proper time. Indeed, Perlick (1987) has discussed such a clock, with the help of the message-and-echo method, just like in the EPS axiomatisation, entirely within the Weyl structure. This issue is discussed in detail in (Adlam, Linnemann, and Read 2022a, 2022b).

- The transition from Weyl metric to Lorentz metric can also be done by postulating the so-called "no second-clock effect". This effect occurs if a vector in a Weyl space, after being parallelly transported along two different curves to the same point, preserves its length, and if this is valid for all such configurations. This procedure, just like the procedure of simultaneity of two series of equidistant events, described above, requires the assumption that the corresponding space-time domain should be simply connected. This must be guaranteed by the axioms.

For more of constructive criticism on EPS approach, see for example: (Linnemann and Read 2021; Trautman 2012; Dewar, Linnemann, and Read 2022; Pfister and King 2015). A large amount of literature that has grown up around the EPS axiomatisation testifies to the pertinence of the issue.

It is important to stress that the axioms of general relativity, as formalised in (Elhers, Pirani, and Schild 1972), can be and have been tested against empirical data. For example, the axiom on the conformal structure of space-time would be spoiled if the velocity of photons would depend upon its energy (Amelino-Camelia et al. 1998; Amelino-Camelia 2013). No such effect has been observed to a high degree of accuracy (Abdo et al. 2009; Perlman et al. 2015; Pan et al. 2020). In the same vein, one could seek deviations from the universality of the projective structure by inspecting the dispersion relations of high-energy massive particles (Amelino-Camelia et al. 2016). Very recently, such possible effects were refuted basing on the scrutiny of astrophysical neutrinos (The IceCube Collaboration 2022).

The existence of the Lorentzian structure on top of the Weyl structure could be undermined through the observation of the "second clock effect" (Elhers, Pirani, and Schild 1972). The latter arises commonly in various modified-gravity theories (see e.g. (De Felice and Tsujikawa 2010)). Recently, the second clock effect was constrained using the CERN data on muon anomalous magnetic moment (Lobo and Romero 2018). For newer experimental searches for Lorentz structure violation coming possibly from torsion effects see: (Kostelecký, Russell, and Tasson 2008; Schettino et al. 2020; Delhom-Latorre, Olmo, and Ronco 2018).

Eventually, one can question the model of a space-time as a smooth manifold. This is typically done on the basis of some quantum gravity theory (Oriti 2009). Although the possible effects of these models are extremely hard to observe, the experimental efforts in this direction are progressing (Amelino-Camelia 2013; Addazi et al. 2022). In particular, some limits on space-time granularity have recently been established (Chou et al. 2016).

4. Causal structure and global time

Every text-book on relativity tells us that causality (the causal structure) is identical with the cone structure of space-time, that is to say with the Weyl structure of the above scheme (synchronised projective and conformal structures). Chronology relation[2] tells us how causal influences can be propagated along timelike curves, whereas causality relation[3] takes also into account causal influences propagating along null geodesics. Both determine channels through which causal influences can propagate rather than actual interactions between cause and effect. With the help of these relations we define the chronological future $I^{+}(p)$ and the chronological past $I^{-}(p)$ of an event p as the set of all events which chronologically follow (resp. are followed by) p; and analogously, the causal future $J^{+}(p)$ and the causal past $J^{-}(p)$ of p.

If M is a space-time manifold, at each of its points (events) p there is a tangent space T_pM equipped with the structure of the Minkowski space-time. Its causal structure is the familiar light cone structure of special relativity. This structure, essentially unchanged, is inherited by any local

2. An event p is said to chronologically precede an event q, written $p \ll q$, if there is a future-directed timelike curve from p to q; for the full account of chronology and causal relations see (Minguzzi and Sánchez 2008; Minguzzi 2019).

3. An event p is said to causally precede an event q, written $p \preceq q$, if there is a future-directed timelike or null curve from p to q, or $p = q$.

neighbourhood (called normal neighbourhood) of the space-time manifold M.[4] However, globally (outside normal neighbourhoods) causal structure can be very different from that of Minkowski space-time, sometimes extremely exotic and full of pathologies (see (Carter 1971; Minguzzi 2019)).

There exists an interaction between the causal structure of space-time and its topology. Sets $I^+(p)$ and $I^-(p)$ are always open, but sets $J^+(p)$ and $J^-(p)$ are not always closed. This allows one to define topology "innate" for causal spaces (i.e. defined entirely in terms of causal relations). It is the so-called Alexandrov topology[5]. This topology is weaker than the usual manifold topology[6].

This mathematical apparatus proved to be very efficient in disentangling various problems related to the structure of space-time. Engaging it into subtleties of space-time architecture has allowed Woodhause to improve the EPS axiomatisation (Woodhouse 1973). He was able to derive the differential and causal structures from more or less EPS axioms expressed in terms of chronology and causality relations. The main advantage of Woodhouse's approach is that no assumption has to be made concerning paths along which light signals propagate. They are deduced from statements regarding emission and absorption of light signals. In his approach, one has to assume that there is exactly one history of a particle through each point in each direction in space-time. Physically this means that it is possible to define the history of a freely falling particle through any point in space-time.

The above conceptual machinery provides a powerful tool for studying various global aspects of space-time. In what follows, we focus on those of them that elucidate mutual dependencies between time and causality.

The first condition that has to be imposed in order to have something resembling a temporal order is to guarantee the absence of closed timelike and causal curves. The corresponding conditions are called the chronology condition (the absence of closed timelike curves) and the causality condition (the absence of closed timelike or null curves[7]), respectively. The motivation is obvious: with such loops there is no clear distinction between the future and the past. Although the existence of relativistic world models with closed timelike curves (such as the famous Gödel's model 1949) shows that the idea of "closed time" is not a logical contradiction, yet the point is that the space-time structure strongly interacts with the rest of physics, what might be a source of many logical perplexities. Not only in a space-time with closed timelike curves one might kill one's ancestor to prevent one's birth, but also — more prosaically — in a space-time with causality violations no global Maxwell field could exist that would match a given local field configuration (Geroch and Horowitz 1979).

There is a rich hierarchy of stronger and stronger conditions[8] which improve temporal and causal properties (Carter 1971; Hawking and Ellis 1973; Minguzzi and Sánchez 2008; Minguzzi 2019). Among them there is an important condition, called strong causality condition, that excludes almost closed causal curves[9]. In such a space-time, the nonexistence of closed timelike curves is satisfied with a certain safety margin. Owing to this margin, the Alexandrov topology improves to the manifold topology.

However, this is not enough. The fact that any measurement can be done only within certain unavoidable error limits, prevents us from measuring the space-time metric exactly. Many "near-by metrics" are always within the measurement's "error bar". If measurements are to have any meaning at all — and the very existence of physics depends on this — we must postulate a certain stability of measurements. This postulate, as regarding relativistic causality, assumes the form of the stable causality condition; it precludes small perturbations of space-time metric to produce closed causal

4. Through the so-called exponential mapping, exp: $(T_pM) \to M$.

5. Sets are defined to be open in this topology if they are unions of the sets of the form $I^+(p) \cap I^-(p)$.

6. Alexandrov topology coincides with manifold topology if the strong causality condition is satisfied; see below.

7. Timelike and null curves are jointly called causal curves.

8. In fact, the hierarchy is nondenumerable.

9. More precisely, it states that each neighbourhood of any event in space-time contains a neighbourhood which no causal curve intersects more than once.

curves[10].

The condition of stable causality has therefore a certain philosophical significance. When it is not satisfied, measurements of physical quantities become essentially meaningless. As put by Hawking, "Thus the only properties of space-time that are physically significant are those that are stable in some appropriate topology" (Hawking 1971) [11].

It is a nice surprise that this condition of "physical reasonability" meets with another very "reasonable" property. To Hawking we owe the theorem: In a space-time M there exists global time, measured by a global time function, if and only if M is stably causal (Hawking 1969). The history of any clock in the universe (for instance, the history of a vibrating particle) is a timelike curve in space-time M. Indications of such a clock can mathematically be represented by a monotonically increasing function along this curve – the time function for this clock. If there exists a single function $\mathcal{T}$, which is a time function for a family of clocks filling the space-time M, then such a function $\mathcal{T}$ is called a global time function. It measures a global time in the universe. As we can see, there exists a deep connection between global time and the above mentioned measurement stability property. It shows that doing physics automatically requires both: global time and stable causality.

We should notice that the spacelike hypersurfaces $\mathcal{T}$ = const. give surfaces of simultaneity in the universe, but they are not unique. However, one can "synchronise the universe" by imposing a yet stronger causality condition. This is done in the following way. First, we define the Cauchy surface of space-time M as a subset S of M such that no inextendible causal curve crosses S more than once. This definition was first introduced in the theory of partial differential equations (Leray 1953). The initial data for a given equation are given on a Cauchy surface. A space-time M is said to be globally hyperbolic if it can be presented as a Cartesian product $M \simeq T \times S$, where $T = \mathbb{R}$ is a global time and S a Cauchy surface in M (see (Geroch 1970; Bernal and Sánchez 2006)). The global time function $\mathcal{T}$, such that $\mathcal{T}^{-1}(t),\ t \in T$ is a Cauchy surface, is called the Cauchy time function.

A globally hyperbolic space-time is the closest to the Newtonian absolute space we can get within the framework of general relativity. It is "deterministic" in the sense that initial data given on a Cauchy surface determine, in principle, the entire history of the universe. However, the so-called Cauchy problem in general relativity, in all its mathematical details, is far from being simple and easy (see (Hawking and Ellis 1973, Chapter 7) or (Ringström 2009)).

5. Time and Causality as Primitive Notions

It is clear from the inspection of the axioms of general relativity that they presuppose some primitive notions of time and causality. Indeed, in order to make sense out of the "echos and messages" axioms one needs to assume that the signal is emitted *before* it is received. This statement finds its formal justification in the following reasoning. It is not an exaggeration to assume that each history of any test particle carries a C^0-structure, which assures the local homeomorphism with $\mathbb{R}$. This can be interpreted as a local flow of time with no preferred time orientation. Local time orientation (local arrow of time) is a property that has to come "from outside". Time arrow is not a part of the space-time architecture.

In the EPS axiomatic system, echos and messages are operationally defined in terms of coincidences of clock readings and acts of emission or reception of light rays. However, intuition smuggles into this operationistic picture the idea that the received echo is actually *caused* by the radar signal. It seems, therefore, that some elementary notion of causality underlies the entire EPS system.

10. To define this condition precisely, one should consider the space Lor(M) of all Lorentz metrics on a given space-time manifold M, and equip it with a suitable topology. Only then we are able to determine what a small perturbation of a metric means (Hawking 1969).

11. One could argue that in order to "do physics" it is enough for the stable causality condition to hold only locally. However, since the existence of closed causal curves is essentially a non-local condition, we would need some barrier protecting a given "locality" from causal anomalies.

Therefore, the notion of local time flow and some elementary concept of causality are the minimum upon which a further hierarchy of stronger and stronger conditions is built. As we have seen, the stable causality condition plays a special role in this hierarchy. It guarantees both the existence of global time and allows for stable measurement results.

All of the structures described above, suitably synchronised with each other, are contained in the Lorentz metric structure. It is a standard result that the Lorentz metric structure exists globally on a space-time manifold M if and only if a non-vanishing direction field exists on M. Moreover, the Lorentz metric can always by chosen in such a way as to make this direction field timelike (Geroch 1971). Since such a field on a space-time manifold locally always exists, the same is true for the Lorentz metric. The existence of a Lorentz metric is strictly related to the possibility of performing space and time measurements; therefore, it is almost synonymous with the possibility of doing physics. Above, we have identified such a possibility with the existence a local "topological time" (a C^0-structure on each history of a test particle), here we have the same condition raised to the metric level[12].

A word of warning is to be made. Our conclusions are valid only within the conceptual framework of what we have called the relativistic model of space-time, and only within its reconstruction as it is presented above. Other axiomatic approaches to the geometry of space-time are possible (see, for, instance, (Andréka et al. 2013; Covarrubias 1993; Guts 1995)) and they can give rise to different interpretations[13]. However, we should take into account the fact that it is the general theory of relativity that is deeply rooted in this model, and since this theory is very well founded on empirical data it would be unwise to look for a different model (within the limits of its empirical verifications). We should also emphasise that the EPS axiomatic approach should not be easily replaced by other approaches since it renders justice, and does this very well, to both "theoretical practice" of mathematical physicists and operational demands of experimentalists (it has a strong (quasi-)operationistic flavour).

6. Message

It is perhaps commonplace to attribute the relational concept of time to Leibniz. Every philosopher knows that, according to Leibniz, time is but a relation ordering events one after the other. It is less known that at the end of his life Leibniz supplemented his relational conception of time with what later resulted in the causal theory of time[14]. His new idea attempted to identify the nature of relations constituting the time order. When we read Leibniz on ordering relations, we should not ascribe to him our present concept of formal order, since his own aim was to stress the difference between his own understanding and that of his English opponents (Newton and Clarke). In the "English theory", there exist two classes of entities: instances and events, and the "natural order" of instances[15] determines the order of events. Events happen at certain instances, but instances are independent of events. In Leibniz's approach, the events are the only class of entities, and time is a derived concept given by relations ordering events. The causal conception of time adds a new idea to Leibniz's philosophy. In his metaphysical-literary style: "the present is always pregnant with the future" (Leibniz 1969, p. 557), and as explained by Mehlberg: "…if one arranges phenomena in a series such that every term contains the reason for all those which come after it in the series, the

12. It is interesting to ask how this condition looks from the global point of view. The answer is that if M is noncompact, such a nonvanishing direction field, and consequently a Lorentz metric, always exists, but if M is compact it exists if and only if the Euler–Poincaré characteristic of M vanishes (Geroch 1967).

13. Different – within certain limits. There is one important constraint: The mathematical structure of space-time must be preserved by all interpretations. One could say that the mathematical structure is "invariant" with respect to all admissible interpretations.

14. The note concerning this conception is found in Leibniz's essay entitled "The metaphysical foundations of mathematics" (Leibniz 1989) and published posthumously; see the extensive study by Henry Mehlberg devoted to the causal theory of time (Mehlberg 1980).

15. We would today say the order determined by the metric structure of time.

causal order of the phenomena so defined will coincide with their temporal order of succession." (Mehlberg 1980, p. 46). Leibniz's idea contrasts with that of Hume who attempted to reduce causal order to temporal order. Whitrow puts this in the following way: In Hume's view "the only possible test of cause and effect is their 'constant union', the invariable succession of the one after the other" (Whitrow 1980, p. 323).

The advent of the theory of relativity showed the role played by time and causality in the structure of modern physics in a radically new light and had a profound impact on our understanding of the world. "Time" was unified with "space" within the concept of space-time, which belongs to the basic formalism of general relativity, and the so-called causal structure of space-time is an indispensable tool for theoretical studies in gravitational physics. The confrontation of traditional disputes about the relationship between space and time with what the theory of relativity has to say about them is an important thread of foundational research. But when it comes to foundational research, the impact of quantum physics cannot be ignored. In the theory of relativity cause–effect interactions occur between definite events and the events themselves are identified with points in space-time. In quantum physics such a viewpoint on space-time points could change radically — this problem calls for a separate study.

The above analysis shows that, within the considered relativistic model, temporal and causal properties are strongly coupled with each other, and one cannot say which is logically (or ontologically) prior with respect to the other. They are unified in the Weyl structure to provide a basis for the fully-fledged concept of time and causality. In this sense, causal theory of time à la Leibniz (causality implies time) is not supported by the relativistic model considered here. The same should be said about an attempt, à la Hume, to reduce causal interactions to a merely temporal succession. It should be taken into account that axiomatic systems can be composed in various ways: various concepts can be selected as primitive concepts and various statements can be accepted as axioms of the system, depending on criteria one adopts. The EPS axiomatic system has the advantage over others that it is quasi-operational, i.e. its axioms describe some simple empirical procedures, although they do so in a highly idealised way. We could conclude that, from the philosophical point of view, space-time is a rich holistic structure, and the choice of a specific axiomatics corresponds to the choice of the angle at which we contemplate the whole.

Acknowledgement

M.E. acknowledges support from the National Science Centre in Poland under the research grant Sonata BIS *Beyond Quantum Gravity* (2023/50/E/ST2/00472).

Competing Interests None.

References

Abdo, A.A., et al. 2009. Testing Einstein's special relativity with Fermi's short hard γ-ray burst GRB090510. *Nature* 462:331. https://doi.org/10.1038/nature08574.

Addazi, A., et al. 2022. Quantum gravity phenomenology at the dawn of the multi-messenger era — a review. *Progress in Particle and Nuclear Physics* 125:103948. https://doi.org/https://doi.org/10.1016/j.ppnp.2022.103948.

Adlam, Emily, Niels Linnemann, and James Read. 2022a. Constructive axiomatics in spacetime physics part II: constructive axiomatics in context. *preprint arXiv:2211.05672.*

———. 2022b. Constructive axiomatics in spacetime physics part III: a constructive axiomatic approach to quantum spacetime. *preprint arXiv:2208.07249.*

Amelino-Camelia, Giovanni. 2013. Quantum-spacetime phenomenology. *Living Reviews in Relativity* 16 (5). https://doi.org/10.12942/lrr-2013-5.

Amelino-Camelia, Giovanni, Leonardo Barcaroli, Giacomo D'Amico, Niccoló Loret, and Giacomo Rosati. 2016. IceCube and GRB neutrinos propagating in quantum spacetime. *Physics Letters B* 761:318–325. ISSN: 0370-2693. https://doi.org/10.1016/j.physletb.2016.07.075.

Amelino-Camelia, Giovanni, J. Ellis, N.E. Mavromatos, D.V. Nanopoulos, and S. Sarkar. 1998. Tests of quantum gravity from observations of γ-ray bursts. *Nature* 393 (6687): 763–765. https://doi.org/10.1038/31647.

Andréka, Hajnal, Judit X Madarász, István Németi, and Gergely Székely. 2013. An axiom system for general relativity complete with respect to lorentzian manifolds. *arXiv preprint arXiv:1310.1475.*

Bass, R. W., and L. Witten. 1957. Remark on cosmological models. *Reviews of Modern Physics* 29 (3): 452–453. https://doi.org/10.1103/RevModPhys.29.452.

Belnap, Nuel, Thomas Müller, and Tomasz Placek. 2021. *Branching Space-Times: Theory and Applications.* Oxford: Oxford University Press, December. ISBN: 9780190884314. https://doi.org/10.1093/oso/9780190884314.001.0001.

Bernal, A.N., and M. Sánchez. 2006. Further results on the smoothability of Cauchy hypersurfaces and Cauchy time functions [in English]. *Letters in Mathematical Physics* 77 (2): 183–197. ISSN: 0377-9017. https://doi.org/10.1007/s11005-006-0091-5.

Bielińska, Marta, and James Read. 2023. Testing spacetime orientability. *Foundations of Physics* 53 (1): 8. https://doi.org/10.1007/s10701-022-00649-8.

Carnap, Rudolf. 1925. Über die Abhängigkeit der Eigenschaften des Raumes von denen der Zeit. *Kant-Studien* 30 (1-2): 331–345. https://doi.org/10.1515/kant.1925.30.1-2.331.

———. 1967. *The Logical Structure of the World and Pseudoproblems in Philosophy.* London: Routledge K. Paul.

Carter, Brandon. 1971. Causal structure in space-time. *General Relativity and Gravitation* 1, no. 4 (December): 349–391. ISSN: 1572-9532. https://doi.org/10.1007/BF00759217.

Chou, Aaron S., et al. 2016. First measurements of high frequency cross-spectra from a pair of large Michelson interferometers. *Physical Review Letters* 117 (11): 111102. https://doi.org/10.1103/PhysRevLett.117.111102.

Covarrubias, Guillermo M. 1993. An axiomatization of general relativity. *International Journal of Theoretical Physics* 32 (11): 2135–2154. https://doi.org/10.1007/BF00675025.

De Felice, Antonio, and Shinji Tsujikawa. 2010. $f(R)$ theories. *Living Reviews in Relativity* 13, no. 1 (June): 3. ISSN: 1433-8351. https://doi.org/10.12942/lrr-2010-3.

Delhom-Latorre, Adrià, Gonzalo J. Olmo, and Michele Ronco. 2018. Observable traces of non-metricity: new constraints on metric-affine gravity. *Physics Letters B* 780:294–299. ISSN: 0370-2693. https://doi.org/https://doi.org/10.1016/j.physletb.2018.03.002.

Deutsch, David. 1991. Quantum mechanics near closed timelike lines. *Physical Review D* 44 (10): 3197–3217. https://doi.org/10.1103/PhysRevD.44.3197. http://link.aps.org/doi/10.1103/PhysRevD.44.3197.

Dewar, Neil, Niels Linnemann, and James Read. 2022. The epistemology of spacetime. *Philosophy Compass* 17 (4): e12821. https://doi.org/https://doi.org/10.1111/phc3.12821.

Elhers, J, FAE Pirani, and A Schild. 1972. The geometry of free fall and light propagation. In *General Relativity, papers in honour of J. L. Synge,* edited by L O'Raifeartaigh. Republished in: *General Relativity and Gravitation* **44**, 1587 (2012). Clarendon Press, Oxford–London. https://doi.org/10.1007/s10714-012-1353-4.

Geroch, Robert. 1967. Topology in general relativity. *Journal of Mathematical Physics* 8 (4): 782–786. https://doi.org/10.1063/1.1705276.

———. 1970. Domain of dependence. *Journal of Mathematical Physics* 11 (2): 437–449. https://doi.org/10.1063/1.1665157.

———. 1971. Space-time structure from a global viewpoint. In *General Relativity and Cosmology: proceedings of the 47th international school of physics "enrico fermi", varena, italy, 30 june – 12 july 1969,* 71–103. New York: Academic Press.

Geroch, Robert, and Gary T Horowitz. 1979. Global structure of spacetimes. In *General Relativity – An Einstein Centenary Survey,* edited by S.W. Hawking and W. Israel, 212–293. Cambridge University Press.

Gödel, Kurt. 1949. An example of a new type of cosmological solutions of einstein's field equations of gravitation. *Reviews of Modern Physics* 21 (3): 447–450. https://doi.org/10.1103/RevModPhys.21.447.

Guts, Alexander K. 1995. Axiomatic causal theory of space-time. *Gravitation and Cosmology* 1 (3): 211–215.

Hawking, Stephen. 1969. The existence of cosmic time functions. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 308 (1494): 433–435. https://doi.org/10.1098/rspa.1969.0018.

———. 1971. Stable and generic properties in general relativity. *General Relativity and Gravitation* 1 (4): 393–400.

Hawking, Stephen W, and George Francis Rayner Ellis. 1973. *The Large Scale Structure of Space-time.* Cambridge: Cambridge University Press.

Kostelecký, V. Alan, Neil Russell, and Jay D. Tasson. 2008. Constraints on torsion from bounds on Lorentz violation. *Physical Review Letters* 100 (11): 111102. https://doi.org/10.1103/PhysRevLett.100.111102.

Leibniz, Gottfried Wilhelm. 1969. *Hauptschriften zur Grundlegung der Philosophie.* Edited by E Cassirer. Vol. 1. Hamburg: Meiner.

———. 1989. The metaphysical foundations of mathematics. In *Philosophical papers and letters,* edited by Leroy E. Loemker, 666–674. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-010-1426-7_71.

Leray, J. 1953. *Hyperbolic differential equations.* Princeton: Institute for Advanced Study.

Linnemann, Niels, and James Read. 2021. Constructive axiomatics in spacetime physics part I: walkthrough to the Ehlers-Pirani-Schild axiomatisation. *preprint arXiv:2112.14063.*

Lloyd, Seth, Lorenzo Maccone, Raul Garcia-Patron, Vittorio Giovannetti, Yutaka Shikano, Stefano Pirandola, Lee A. Rozema, et al. 2011. Closed timelike curves via postselection: theory and experimental test of consistency. *Physical Review Letters* 106 (4): 040403. https://doi.org/10.1103/PhysRevLett.106.040403.

Lobo, I.P., and C. Romero. 2018. Experimental constraints on the second clock effect. *Physics Letters B* 783:306–310. ISSN: 0370-2693. https://doi.org/10.1016/j.physletb.2018.07.019.

Luc, Joanna. 2020. Generalised manifolds as basic objects of general relativity. *Foundations of Physics* 50 (6): 621–643. https://doi.org/10.1007/s10701-019-00292-w.

Luc, Joanna, and Tomasz Placek. 2020. Interpreting non-hausdorff (generalized) manifolds in general relativity. *Philosophy of Science* 87 (1): 21–42. https://doi.org/10.1086/706116.

Mehlberg, H. 1980. *Time, Causality, and the Quantum Theory.* Vol. One: Essay on the Causal Theory of Time. Dordrecht–Boston–London: Reidel.

Minguzzi, Ettore. 2019. Lorentzian causality theory. *Living Reviews in Relativity* 22 (1): 1–202. https://doi.org/10.1007/s41114-019-0019-x.

Minguzzi, Ettore, and Miguel Sánchez. 2008. The causal hierarchy of spacetimes. In *Recent Developments in Pseudo-Riemannian Geometry, ESI Lectures in Mathematics and Physics,* edited by Dmitri V. Alekseevsky and Helga Baum, 299–358. Berlin: European Mathematical Society Publishing House.

Oriti, Daniele, ed. 2009. *Approaches to Quantum Gravity: Toward a New Understanding of Space, Time and Matter.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511575549.

Pan, Yu, Jingzhao Qi, Shuo Cao, Tonghua Liu, Yuting Liu, Shuaibo Geng, Yujie Lian, and Zong-Hong Zhu. 2020. Model-independent constraints on Lorentz invariance violation: implication from updated gamma-ray burst observations. *The Astrophysical Journal* 890, no. 2 (February): 169. https://doi.org/10.3847/1538-4357/ab6ef5.

Penrose, Roger. 1965. Gravitational collapse and space-time singularities. *Physical Review Letters* 14 (3): 57–59. https://doi.org/10.1103/PhysRevLett.14.57.

Perlick, Volker. 1987. Characterization of standard clocks by means of light rays and freely falling particles. *General Relativity and Gravitation* 19 (11): 1059–1073. https://doi.org/10.1007/BF00759142.

Perlman, E. S., S. A. Rappaport, W. A. Christiansen, Y. J. Ng, J. DeVore, and D. Pooley. 2015. New constraints on quantum gravity from X-ray and gamma-ray observations. *The Astrophysical Journal* 805 (1): 10. https://doi.org/10.1088/0004-637X/805/1/10.

Pfister, Herbert, and Markus King. 2015. *Inertia and Gravitation: The Fundamental Nature and Structure of Space-time.* Cham: Springer. https://doi.org/10.1007/978-3-319-15036-9.

Reichenbach, Hans. 1924. *Axiomatik der relativistischen, Raum-Zeit-Lehre.* Braunschweig: F. Vieweg & Sohn Akt.-Ges.

Ringström, Hans. 2009. *The Cauchy Problem in General Relativity.* Vol. 6. ESI Lectures in Mathematics and Physics. European Mathematical Society. https://doi.org/10.4171/053.

Robb, Alfred Arthur. 1914. *A Theory of Time and Space.* Cambridge: Cambridge University Press.

Schettino, Giulia, Daniele Serra, Giacomo Tommei, and Vincenzo Di Pierri. 2020. A test of gravitational theories including torsion with the BepiColombo radio science experiment. *Universe* 6 (10). ISSN: 2218-1997. https://doi.org/10.3390/universe6100175.

Schrödinger, Erwin. 1985. *Space-time Structure.* First published in 1950. Cambridge University Press.

The IceCube Collaboration. 2022. Search for quantum gravity using astrophysical neutrino flavour with IceCube. *Nature Physics* 18:1287–1292. https://doi.org/10.1038/s41567-022-01762-1.

Trautman, Andrzej. 2012. Editorial note to: J. Ehlers, F.A.E. Pirani and A. Schild, The geometry of free fall and light propagation. *General Relativity and Gravitation* 44 (6): 1581–1586. https://doi.org/10.1007/s10714-012-1352-5.

Whitrow, Gerald James. 1980. *The Natural Philosophy of Time.* Second. Oxford: Clarendon Press.

Woodhouse, N. M. J. 1973. The differentiable and causal structures of space-time. *Journal of Mathematical Physics* 14 (4): 495–501. https://doi.org/10.1063/1.1666344.

2. Entanglement and Smooth Geometry in 4-spacetime

Jerzy Król

Kielnarowa Review (2025), **2**, 16–27
doi: https://doi.org/10.57656/kr-2025-0002

Kielnarowa
REVIEW

ARTICLE

Entanglement and smooth geometry in 4-spacetime

Jerzy Król*

University of Information Technology and Management, Chair of Cognitive Science and Mathematical Modelling, ul. Sucharskiego 2, 35-22 Rzeszów, Poland
*Corresponding author. Email: jkrol@wsiz.edu.pl

Abstract

We try to understand quantum entanglement by geometric relations of spatially separated regions of 4-spacetime. The relations become detectable by working in the Euclidean smoothness structure underlying the Lorentzian structure. There are 5-dimensional bridges, i.e. 5-dimensional nontrivial smooth h-cobordisms connecting spatially separated 4-regions of spacetime. These connections are nonlocal in spacetime. At the quantum regime spacetime is reduced to its smooth atlases of charts which are related by automorphisms of the maksimal Boolean algebra in the quantum lattice of projections. Quantum entanglement in 4-spacetime can be represented by exotic smoothness structures on $\mathbb{R}^4$, which are determined by the h-cobordisms due to the results in particular by Casson, Akbulut, Freedman, Donaldson or Gompf. The involutions of corks correspond to the phases between the Boolean ZFC-models and to the change of the exotic R^4 in W^5. This work is more a description of the ongoing project than a detailed presentation of the results. The discussion focuses on certain general contexts and even philosophical features.

Keywords: exotic R^4s and 5-cobordisms in spacetime, quantum entanglement, boolean ZFC models

1. Introduction

The breakthrough discoveries in differential topology and geometry from 1980s led to the existence of exotic smooth versions of many compact and noncompact 4-manifolds. Since then, researchers have been trying to understand the importance of this phenomenon in physics. A particular interest is the case of exotic R^4 and $S^3 \times \mathbb{R}$. Both are open smooth 4-manifolds. Exotic R^4s are topologically equivalent (homeomorphic) with $\mathbb{R}^4$ but nondiffeomorphic with standard smooth $\mathbb{R}^4$. Similarly, exotic $S^3 \times \mathbb{R}$ are homeomorphic with $S^3 \times \mathbb{R}$ but nondiffeomorphic with standard smooth $S^3 \times \mathbb{R}$. All known exotic R^4s fall into two uncountably infinite classes, small and large exotic R^4s. Small are embeddable in standard smooth $\mathbb{R}^4$, while large cannot be embedded in $\mathbb{R}^4$.

Consider a topological 5-cobordism $W^5(N_1, N_2)$, where the boundary of the manifold W^5, ∂W^5, is the disjoint sum of two topological 4-manifolds N_1, N_2, i.e., $\partial W^5 = N_1 \cup N_2$. When N_1 is topologically equivalent (homeomorphic) to N_2, $N_1 \simeq N_2$, then $W^5 \simeq N_1 \times I$ where $I = [0, 1] \subset \mathbb{R}$ and we call such cobordism $W^5(N_1, N_2)$ *topologically trivial*. To be more precise, we require that the topological manifolds W^5, N_1, N_2 be simply-connected and compact and $W^5, N_1 \times I, N_2 \times I$ be homotopically equivalent to each other. Then the statement that they are homeomorhic is the topological h-cobordism theorem for W^n, $n = 1, 2, ..., k, ...$ and N_1^{n-1}, N_2^{n-1} (h stays for the above mentioned homotopical equivalence). The proof for the W^5 case was given in Freedman 1982.

We want to model nonlocal connections between regions of spacetime manifold by 5-cobordisms, such that this can shed light on the phenomenon of quantum entanglement. However, certain clarifications are in order. Spacetime is a smooth Lorentzian 4-manifold, but its underlying manifold (mathematically, which can be non-physical) is an Euclidean open smooth 4-manifold on which the Lorentzian structure is introduced. Thus, open regions U_1, U_2 in spacetime should be subsets of N_1, N_2 correspondingly, and all manifolds W^5, N_1, N_2, U_1, U_2 should be smooth. Then we deal with smooth h-cobordism 'theorem' rather than the topological one. However, in dimension $n = 5$ the statement:

Given a smooth h-cobordism W^5 between smooth, compact, Euclidean, simply-connected 4-manifolds N_1, N_2 such that $W^5, N_1 \times I, N_2 \times I$ all are pairwise homeomorphic, the smooth cobordism $W^5(N_1, N_2)$ is diffeomorphic to $N_1 \times I$

is, in general, false. This was shown in Donaldson 1983. The understanding of this phenomenon is deeply rooted in modern differential topology and geometry and is crucial for this work.

The quantum uncertainty principle between some pairs of observables and quantum entanglement between two systems are nontrivially related. Entanglement can be detected as non-classical (e.g. breaking the Bell's bound) if there are noncommuting observables $[X, Y] \neq 0$ at Alice and Bob sides. Certainly, entanglement cannot erase the uncertainty of the observables applied to each system separately, but it can reduce the quantum bound for the entangled state when correlations are between the results of Alice $a(X), a(Y)$ on the entangled state and the results of Bob $b(X), b(Y)$ on it. An entangled state is just a state and it does not require any noncommuting observables (or uncertainty between them). But to witness entanglement as a quantum phenomenon we need noncommuting X, Y. Similarly, entanglement can be detected in spacetime by measuring the momentum and position P, Q. For a single system, the entanglement of states cannot reduce the Heisenberg uncertainty, however, for two entangled subsystems $s_1 \otimes s_2$ the P, Q measurement in the Alice and Bob laboratories can reduce the uncertainty, which, though, does not contradict the uncertainty relations in each subsystem individually (Horodecki et al. 2009).

Here we try to understand the nonlocality in spacetime by incorporating 5-dimensional bridges corresponding to nontrivial W^5 cobordisms. These bridges introduce additional correlations between spatially separated noncommuting observables.

2. 5-dimensional bridges in 4-spacetime and exotic smoothness

This and the following sections collect arguments for seeing nontrivial smooth cobordisms W^5 as modeling the spacetime bridges that connect 4-regions U_1, U_2 . Here we focus on the classical differential geometric point of view, while in the next section we connect this with the formalism of quantum mechanics (QM). As we mentioned above, we work in Euclidean metrics, since they underlie the Lorentzian case, and the smoothness structures assigned to Euclidean $\mathbb{R}^4$ allow for uncovering a fundamental mathematical layer that has connections to QM. We use the symbol R^4 for exotic $\mathbb{R}^4$s where $\mathbb{R}^4$ refers to the standard smoothness structure on $\mathbb{R}^4$. Let us follow the idea in Król and Asselmeyer-Maluga, 2025 that under certain, rather extreme conditions in spacetime, one can separate $p = (p_1, p_2, p_3, E)$ and $q = (q_1, q_2, q_3, t)$ as coordinates of a single physical object, and they contribute as local coordinate frames in the reconstructed defragmented new spacetime. Here we describe the separation process as the evolution of these 4-dimensional patches $p, q \simeq \mathbb{R}^4$ that are represented by trivial open 5-cobordisms $C_0 = \mathbb{R}^4 \times I$ as the subcobordism of certain trivial compact $W^5(N_1, N_2) \simeq N_1 \times I$, thus, $C_0 \subset W^5$. So, the evolution of p or q can be represented by trivial cobordisms as in Fig. 1a) when the momenta and positions to be measured are actually assigned to different particles. However, in the case where momentum and positions are observables of a single particle, the uncertainty relation constrains their simultaneous measurement

$$\Delta\overline{p} \cdot \Delta\overline{q} \geq \frac{\hbar}{2}.$$

This 3-dimensional version can be formally extended over dimension four as follows

$$p^{\mu} \cdot q_{\nu} \geq \frac{\hbar}{2}\delta^{\mu}_{\nu}, \mu, \nu \in \{0, 1, 2, 3\},\ p_{\mu} = (E, \overline{p}), q_{\nu} = (t, \overline{x}).$$

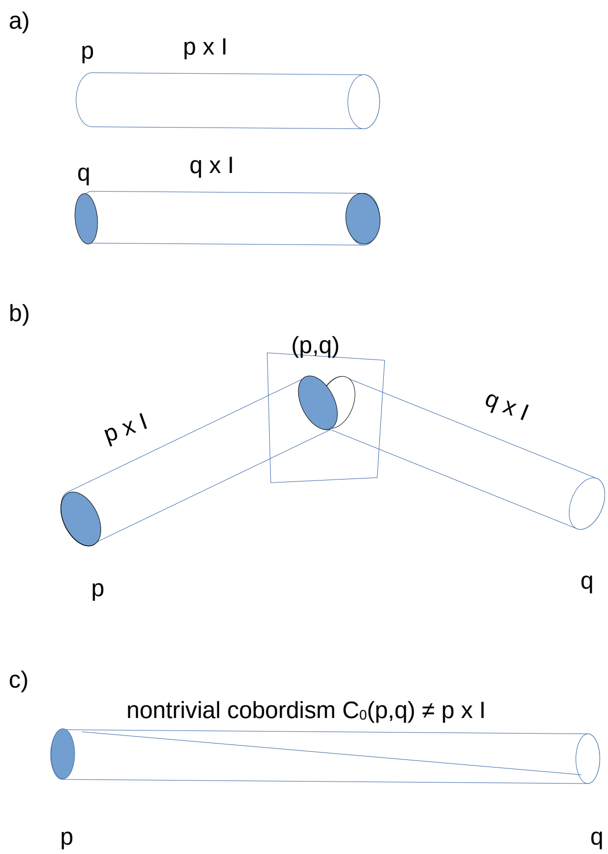


Figure 1. 5-dimensional evolutions of 4-regions. a) p, q evolve via trivial cobordisms $p \times I, q \times I$. This corresponds to the independent shifts of 4-dimensional local regions in spacetime as can be the case for two separated particles such that their positions and momenta are not confined by the uncertainty reltion. b) Uncertainty relations in a small 4-region (p, q) affect the evolutions of the regions p and q and enforces the correlations. c) There results a nontrivial 5-cobordism $C_0(p, q)$ as a subcobordism of certain $W^5(N_1, N_2)$.

These 4-dimensional coordinate patches $\mathbb{R}^4 \simeq p, q$ can evolve in 5-dimensions as 5-cobordisms, leading to their spatial separation with a degree of uncertainty being transferred to entanglement. The quantum side of the process will be discussed in the next section. Here we want to determine the smooth geometric component of it. The evolution of the p, q patches that undergo the uncertainty principle in the micro initial state is represented in Fig. 1b). In Król and Asselmeyer-Maluga, 2025 conditions were given for the destruction of the integrity and causality of the spacetime manifold and this resulted in the fragments of spacetime, each being a $\mathbb{R}^4$ local patch. Then, the reverse process was described to retain the integrity, causality, and the underlying smoothness structure. The smooth regions of the reconstructed Euclidean spacetime become necessarily exotic smooth $\mathbb{R}^4$, i.e. R^4. Here we focus on this classical geometric picture (without touching quantum level) and describe the appearance of R^4s in spacetime due to nontrivial cobordisms W^5 connecting 4-

regions p, q in spacetime. This is represented in Fig. 1c) where nontrivial open 5-subcobordism $C_0(p, q) \subset W^5(N_1, N_2)$ is visualized. The classical counterpart to losing the causality and integrity of 4-spacetime is to consider the 5-dimensional C_0 that cannot sit in M^4 and thus cannot be local in spacetime. We will see that indeed C_0 contains information about quantum entanglement.

Why C_0 has to deal with exotic R^4s? This is really a deep result in mathematics first observed and proved in Donaldson 1983 by applying gauge field theory methods to the differential geometry of 4-manifolds and making use of another profound result in the topology of 4-manifolds obtained in Freedman 1982. The analysis of the nontrivial cobordism $W^5(N_1, N_2)$ for compact 4-manifolds N_1, N_2 that were simply connected such that $N_1 = K3\#\overline{CP^2}$ and $N_2 = \#CP^2\#20\overline{CP^2}$ led Akbulut to the discovery of compact cork (Akbulut cork) $K \subset N_1$ with a nonempty boundary ∂K. Here $K3$ is the celebrated $K3$ - surface (complex 2-dimensional compact smooth manifold), CP^2 is the 2-dimensional complex projective space, $\overline{CP^2}$ is CP^2 with the reversed orientation and # is the connected sum of two manifolds (obtained by cutting off 4-balls from each manifold and gluing the remnants manifolds by homeomorphism, or difffeomorphism, of their common boundary). It appears that in the above case the Akbulut cork K has an explicit description via handle decomposition, and this is known as Mazur manifold (Akbulut, 1991(a), 1991(b)) and $\partial K = \Sigma(2, 5, 7)$ is the homology 3-sphere, one of Brieskorn's spheres (Gompf and Stipsicz 1999). The first contractible 4-manifolds whose nonempty boundary would not be S^3 were constructed by Mazur in 1961. Then many other examples were built and it was shown that the boundary of such generalized Mazur manifolds can be Brieskorn homology 3-spheres $\Sigma(2, 5, 7), \Sigma(3, 4, 5), \Sigma(2, 3, 13)$ (Akbulut and Kirby, 1979). Since then it has been shown (in particular by Casson) that many other Brieskorn spheres appear as the boundary of such contractible 4-manifolds. Thus, the point is that for any nontrivial smooth 5-cobordism $W^5(N_1, N_2)$ between two smooth nondiffeomorphic, compact, simply-connected 4-manifolds N_1, N_2, there is always an embedded cork $K \subset N_1$ and the special role is played by ∂K. The celebrated example is $N_1 = K3\#\overline{CP^2}$ and $N_2 = \#CP^2\#20\overline{CP^2}$, where these manifolds are homeomorphic but nondiffeomorphic, and the Akbulut cork K, which is a compact, contractible 4-manifold, has a boundary $\partial K = \Sigma(2, 5, 7)$.

The following decomposition result is basic for understanding 5-h-cobordisms and their relation to exotic R^4s (Curtis et al., 1996, p.343). Let N_1, N_2 be two smooth compact simply-connected 4-manifolds and $W^5(N_1, N_2)$ be a smooth nontrivial simply-connected 5-cobordism between them.

Theorem 1 *There exist decompositions:* $N_1 = K \cup_{\partial K} N_1'$ *and* $N_2 = K \cup_{\partial K} N_2'$ *such that* $(N_1', \partial K)$ *and* $(N_2', \partial K)$ *are diffeomorphic and the subcobordism* $W^{5\prime}(N_1', N_2') = N_1' \times I$ *is trivial.*

K is the compact contractible 4-manifold (a cork) with a nonempty boundary ∂K. In conclusion, we have

Theorem 2 *a) There exist two nondiffeomorphic exotic* $\mathbb{R}^4$*s,* R_1^4 *and* R_2^4 *such that* $K \subset R_1^4 \subset N_1$ *and* $K' \subset R_2^4 \subset N_2$*. b)* K' *is diffeomorphic to* K *relative to the nontrivial involution of its boundary* $\tau(\partial K)$*.*

The decomposition of an h-cobordism $W^5(N_1, N_2)$ is shown schematically in Fig. 2

A. Casson also showed that the nontriviality of 5-cobordism can be obtained by cutting off R_1^4 in N_1 and gluing it back by involution of its end in infinity. In fact, the reduction of this Casson 'non-compact' procedure to compact K and the involution of ∂K was quite a surprise. Finally, it was shown in Gompf, 2018 that these two procedures can be used interchangeably and the subsequent generalization of this is possible to the G-slice corks and G-slice exotic R^4 for quite general groups G.

For the canonical example recognized by Akbulut we have $\partial K = \Sigma(2, 5, 7)$ and its involutions give rise to the change of the smoothness structures from R_1^4 to R_2^4. Note that such a change of the smoothness structure on $\mathbb{R}^4$ cannot be described within GR formalsim, as well as the 5-cobordisms connecting regions of 4-dimensional spacetime are not entities describable on M^4. This is the motivation for considering some extension of GR which would lead to a close relation with the QM formalsim.

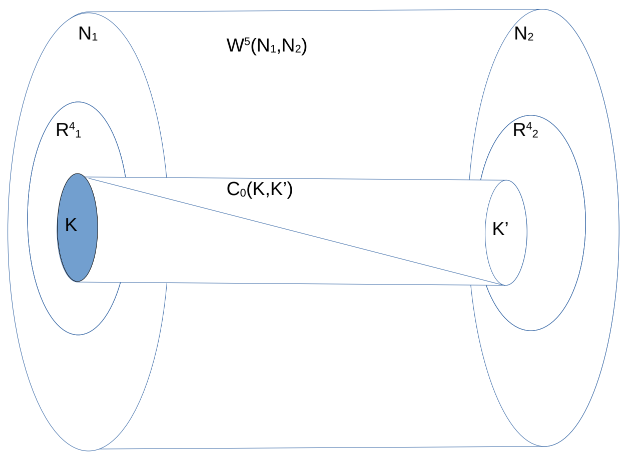


Figure 2. The decomposition of a nontrivial h-cobordism $W^5(N_1, N_2)$.

3. QM and the evolution of exotic spacetime 4-regions

In this section, we present arguments that the QM formalism can support the existence of exotic 4-regions in spacetime and their evolution as 5-cobordisms. The exotic R^4s here are small, i.e., embeddable in standard $\mathbb{R}^4$. The 5-cobordisms certainly cannot be embedded in 4-spacetime and can be a source of certain nonlocality detected in spacetime. However, some additional conditions must be met. We start with the local modeling of spacetime by R^4s patches that belong to local Boolean models V^B rather than to any universal a priori established universe of sets like the Von Neumann cumulative proper class model V (reference to the constant V is a usual practice in physics.). The varying local ZFC models do not follow any sheaf category localizations but rather are hybrid methods of Boolean models and the relations between them. Reference to the local patches of the spacetime 4-manifold (before the Lorentzian twist) as embedded also in local Boolean-valued models V^B has a nice advantage when one thinks about unification of GR and QM. This is the extension of the diffeomorphism group of M^4 by the group of morphisms between V^Bs where the latter are generated by $Aut\,B$.

$$Diff(\mathbb{R}^4) \oplus Aut\, V^B.$$

Thus, the equivalence principle of GR gains additional degrees of freedom connected with the automorphisms of V^B such that the local choice of flat $\mathbb{R}^4$ where gravitational energy vanishes, is augmented by the local phase connected with the automorphism of models where $\mathbb{R}^4$ patches are living, $R^4_{V^{B1}} \in V^{B1}$. Two flat local coordinate patches $\mathbb{R}^4_1, \mathbb{R}^4_2$ in addition to their flatness can be distinguished by their local phase:

$$R^4_{V^{B1}} \simeq_F R^4_{V^{B2}} \text{ where } F : V^{B1} \to V^{B2} \text{ is the relative phase.}$$

Instead of $R^4_{V^{B1}}$ we can write R^4_1 and this is the product $R \times R \times R \times R$ in the ZFC model V^{B1} and this R^4_1 has nothing to do with exotic R^4_1 as before. The context will clearly distinguish both objects. Now the point is that this phase F has QM origins and on infinite-dimensional Hilbert spaces $\mathcal{H}^\infty$ can be further reduced to the automorphisms of a Boolean algebra B. So, the extended equivalence principle taking into account the relative phase F would read as follows.

> [ExtEP] There is always possible in the microscale to choose a local coordinate frame $U = \mathbb{R}^4$ and a Boolean ZFC model V^B that it would erase the gravitational effects in U and U would be entangled with another possibly spatially separated 4-region $U' \simeq U$ such that their entanglement induces gravitational effects nonlocalized to U.

We want to show that this ExtEP is probable by giving step by step explanation, but a more rigorous justification will be published elsewhere. Let us reformulate ExtEP as

> [ExtEP'] The choice of the local $\mathbb{R}^4$ patch in microscale in spacetime that erases all gravitational effects always factorizes through exotic R^4 and leads to the quantum entanglement effects.

A full justification of this very strong statement will be the subject of a separate publication. Here we focus on some elements of ExtEP' making it probable. Let $U = \mathbb{R}^4$ and $\dim \mathcal{H} = \infty$. The ZFC twist of QM (Król and Asselmeyer-Maluga, 2025) determines $F : V^{B1} \to V^{B2}, F \neq Id$ and:

a) a Boolean ZFC model V^B where $U = R^4_B$;
b) a Boolean ZFC model $V^{F(B)}$ where there is an accompanied flat $U_F = R^4_{F(B)}$;
c) exotic R^4_1 where U_B is a flat local coordinate patch;
d) exotic R^4_2 where $U_{F(B)}$ is a local flat coordinate patch.

a) means that once local flat $U' = \mathbb{R}^4$ is obtained in spacetime erasing gravitational effects, which is legitimate due to EP, at the micro level there is room for a deeper representation of $U' = \mathbb{R}^4$ as $U = R^4_B$ in certain ZFC Boolean model V^B. Even though the models V^Bs vary depending on the specific region of spacetime, for $\dim \mathcal{H} = \infty$ there exists just a single model V^B where B is the atomless Boolean measure algebra. The different regions in spacetime refer to different automorphic copies of $F'(V^B)$. However, $F'(V^B)$ are determined by the automorphisms F of B

$$\forall_{F'} \exists_F F'(V^B) = V^{F(B)}, F \in Aut\, B.$$

Thus b) holds true.

The justification of c) and d) is based on the relation of the quantum lattice of projections $\mathbb{L}$ for $\mathcal{H}^\infty$ and the spacetime manifold M^4 with a smooth atlas $\{U_\alpha\}_{\alpha \in I}$. Again, crucial is the ZFC twist of QM that assigns models V^{B_α} to the local patches U_α such that

$$U_\alpha \simeq R^4_\alpha \text{ in } V^{B_\alpha} \text{ and } V^{B_\beta} = V^{F_{\alpha\beta}(B_\alpha)} \text{ where } F_{\alpha\beta} \in Aut\, B.$$

We can characterize the border line between quantum and classical on the basis of a smooth atlas $\mathcal{U} = \{U_\alpha\}_{\alpha\in I}$ subordinated to the lattice $\mathbb{L}(\mathcal{H}^\infty)$. Let B be a maximal atomless Boolean subalgebra of projections chosen from $\mathbb{L}$. B is the measure algebra. We call $\mathcal{U}$ a quantum atlas of $\mathbb{R}^4$ if the corresponding projections from $\mathbb{L}$ belong to at least two different maximal algebras.

Lemma 1 *If* $|\mathcal{U}| \geq 2$ *then* $\mathcal{U}$ *is quantum.*

This follows from the $1:1$ correspondence between U_α from $\mathcal{U}$ and V^{B_α} which gives at least two maximal Boolen algebras $B \subset \mathbb{L}$ and consequently, it has to be determined at least a pair of noncommuting observables subordinated to the different algebras correspondingly. In addition, we have

Lemma 2 *(Król et al., 2017, Corr. 1, Th. 3)* If any smooth $\mathcal{U}$ on some $\mathbb{R}^4$ is quantum, then such a smooth $(\mathbb{R}^4, \mathcal{U})$ has to be exotic R^4.

On the contrary, $|\mathcal{U}| = 1$ corresponds to the classical case and the standard smooth $\mathbb{R}^4$. Thus we arrive at the conclusion that a quantum system in spacetime can modify its local standard smoothness toward exotic $\mathbb{R}^4$ and this exotic R^4 encodes information about the quantum system. The fundamentals are the following

Lemma 3 *A nontrivial 5-cobordism* W^5 *can encode the change of base in Hilbert space* $\mathcal{H}^\infty$.

In particular, the projectors in one base are sent to some projectors in another base.

Theorem 3 *The quantum entanglement of two spatially separated quantum systems at microscale is represented in 4-spacetime by the relation of two exotic smooth local 4-patches linked by a nontrivial 5-cobordism* W^5 *outside of spacetime.*

Instead of the proof, we collect the main points with short commentaries leading to Theorem 3.

P1. Entanglement as a quantum effect in spacetime requires multiple local patches at the initial interaction stage and in the final spatially separated stage.
P2. Any quantum system in spacetime determines a collection of local coordinate frames $\{qU_p\}_{p\in M^4}$ in a way that: Any observable $\mathcal{O}_i, i \in O$ when measured in a state $\psi_0 \in \mathcal{H}^\infty$ determines the value $o_i(\psi_0)$ and the local flat $\mathbb{R}^4_{o_i}$.
P3. On the microscale $\mathbb{R}^4_{o_i}$ becomes R^4_i in the ZFC Boolean model V^{B_i}.
P4. When fragmented, the spacetime manifold M^4 becomes a collection of flat local $R^4_i, i \in I$ and a set of relations $F'_{ij} : V^{B_i} \to V^{B_j}$ between ZFC models hosting R^4_is.
P5. If the observables $\mathcal{O}_i, \mathcal{O}_j, i \neq j$ are compatible, i.e., measured simultaneously, then $V^{B_i} = V^{B_j}, F'_{ij} = id$. If $[\mathcal{O}_i, \mathcal{O}_j] \neq 0$ then $F'_{ij} \neq id$ and $R^4_{o_i}$ is not identically diffeomorphic to $R^4_{o_j}$.

Fragmentation in P4. was explained in detail in Król and Asselmeyer-Maluga, 2025. The microscale appearance of ZFC Boolean models in spacetime regions (P3.) has been systematically explored in Król and Asselmeyer-Maluga, 2020; Król, Bielas, and Asselmeyer-Maluga, 2023. P2. and P5. are direct consequences of the construction, and P1. follows from them and Lemma 1.

4. Discussion and perspectives

Theorem 3 opens several opportunities related to entanglement and its relation to spacetime regions; however, they need a more thorough explanation. Nonlocality in this context means a relation of two 4-regions that is nonlocal in spacetime. Note that the bridge $W^5(N_1, N_2)$ connecting the regions is 5-dimensional and cannot be local 4-dimensional. However, these geometric data of W^5

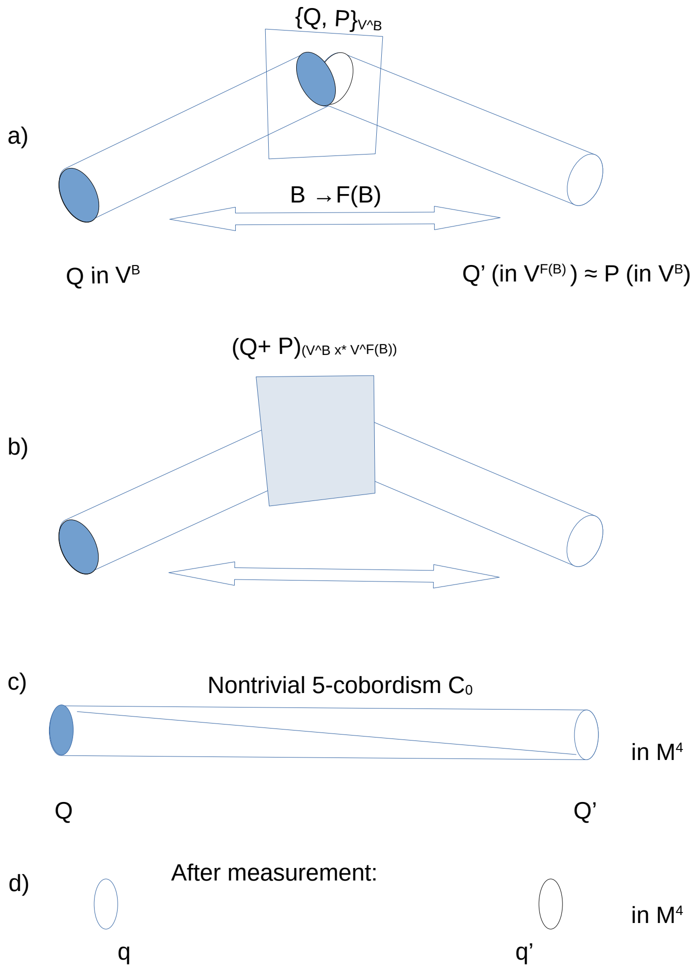


Figure 3. Entanglement and uncertainty in spacetime. a) In some conditions q-coordinates and p-coordinates assigned to a single particle can become separated in 4-dimensions. Here p, q refer to $(\overline{p}, E)$, $(\overline{x}, t)$ correspondingly. They become 4-dimensional local $\mathbb{R}^4$ patches, however with the relative phase $F \in Aut(B)$. b) Q, P become entangled before measurement in 4-spacetime M^4. This leads to $(P + Q)$ in the composed ZFC model $V^B x * V^{F(B)}$. c) The spacetime separation of P, Q and the entanglement of them leads to the entangled regions in M^4. The measurement of a particle's P in V^B is entangled with the measurement of Q' in $V^{F(B)}$ since Q' in $V^{F(B)}$ corresponds to P' in V^B and the momentum is preserved in V^B. d) After measurement the 4-regions of P and Q coordinates contribute to spacetime as independent local regions.

should be augmented with a suitable quantum content to represent the entanglement. We will see that it is possible and that W^5 can represent a quantum entanglement.

First, following Lemma 1 if any open atlas $\mathcal{U}$ of $\mathbb{R}^4$ fulfils $|\mathcal{U}| \geq 2$ then it can encode quantumness. This relies on the direct observation that if there is $1 : 1$ correspondence of the maximal Boolean algebras in $\mathbb{L}$ with open maps in $\mathcal{U}$ then given at least two distinct such maps we have to have at least a pair of noncommuting observables $\mathcal{O}_1, \mathcal{O}_2, [\mathcal{O}_1, \mathcal{O}_2] \neq 0$ or an observable $\mathcal{O}$ with at least two nonperpendicular but different positive operator-valued measure (POVM) eigenspaces. However, to measure entanglement in spacetime we need noncommuting observables let them be $X, Y, [X, Y] \neq 0$. Otherwise, the Bell-like inequality (e.g. CHSH) will never be broken, or the classical bound would never be exceeded. That is why to see entanglement as witnessed by X, Y in spacetime one needs $[X, Y] \neq 0$ and X, Y to be applied to both entangled systems. This is the reason for $|\mathcal{U}| \geq 2$ as above.

It follows another strong consequence, as stated in Lemma 2: *Any smooth quantum $\mathbb{R}^4$ has to be exotic R^4*. Thus, we find another heuristic justification for Theorem 3.

Let us derive a direct description of the nonlocal entanglement of observables $[X, Y] \neq 0$ in terms of W^5. Let $a(X), a(Y), b(X), b(Y) \in \{a_1, a_2, b_1, b_2\}$ be the possible outcomes of Alice and Bob. Let the phase $F' : V^{B1} \to V^{B2}$ sends X to Y. Thus, measuring X on Alice side gives rise to measuring Y on Bob's side. There are sets of local patches U_1, U_2, U_1', U_2' for Alice and Bob (primes), respectively. To clarify presentation, let us refine the correspondence (which does not affect the construction) *opens in M^4 to operators*: 'subfamily $\{U_i, i \in J\}$ of open local patches of M^4' $\leftrightarrow$ 'eigenvalues $\{o_i \in \mathbb{R}, i \in J\}$ of any observable $\mathcal{O}$'. This subfamily has relative phase $F' : V^B \to V^B$ identity on V^B, since the projections on the eigenspaces of the same $\mathcal{O}$ commute. Still, the relative diffeomorphisms are not identities $f : \mathbb{R}^4 \to \mathbb{R}^4, f \neq id$ and this is the refinement that extends. The minimal case states that there are at least two eigenvalues and two corresponding open charts $(o_1, o_2; U_1, U_2)$ all in the same model V^{B1}. So, thus, U_1, U_2 are opens of $R_1^4 \in V^{B1}$ and any open cover of R_1^4 has a cover that does not allow for a single local chart R_1^4. This also follows from taking two noncommuting X, Y (which is essential to detect entanglement) and that X, Y are in the same model V^{B1} (Alice) and V^{B2} (Bob). This means R_1^4 is already exotic in V^{B1} (Lemma 2). A smooth open nontrivial subcobordism of W^5 sends R_1^4 to R_2^4 in V^{B2} with the nontrivial relative phase $F' : V^{B1} \to V^{B2}, F' \neq Id$. Then R_2^4 is exotic in V^{B2}. The minimal extension as above on the commuting eigenprojections of $\mathcal{O}$ assigns the local charts U_1, U_2 of M^4 (for any smooth atlas) with nonidentity diffeomorphisms as transition maps. After measurement, the result is described in one frame, U_1 corresponding to a_1 or U_2 for a_2 in V^{B1}. Similarly for the image $U_1'(b_1)$ or $U_2'(b_2)$ in V^{B2}. Let us assume that U_1, U_2 in exotic R_1^4 correspond to the ends E_1, E_2 of R_1^4 (Gompf and Stipsicz 1999). Then, in R_2^4 they are assigned to the ends with the following twist

$$U_1 \to E_1 \to U_1' \to E_2' \text{ in exotic } R_2^4 \text{ in } V^{B2}$$
$$U_2 \to E_2 \to U_2' \to E_1' \text{ in exotic } R_2^4 \text{ in } V^{B2}.$$

The entanglement of 4-regions in spacetime causes the instantaneous action at a distance: *a measurement of a_1 in U_1 in R_1^4 enforces the measurement of b_2 in U_2' in R_2^4*. The subtlety is a hidden, though general property of spacetime: in the quantum regime open covers replace the spacetime manifolds. In other words, *spacetime is defined categorically in the classical regime from open covers*. Thus, quantum measurement in spacetime determines local micro patches which are related by the 5-dimensional bridges-cobordisms.

The entire process behind the twist of local patches at spatially separated regions is based on Theorem 2 and its equivalence to the twist of the ends at infinity of exotic R_1^4 to obtain R_2^4. This twist is generated by the nontrivial phase $F_{12}' : V^B \to V^B$ (Figs. 2, 3).

The appearance of such 5-dimensional bridges bears certain similarity to the recently formulated proposal that possibly entanglement is nonlocal in spacetime due to connections via wormholes. Currently we do not have arguments allowing for association of our geometric 5-bridges with Suskind's proposal. Let us note only that the source of exotic smoothness in spacetime has been assigned to fragmentation due to extreme conditions in the singular regions like in black holes (Król and Asselmeyer-Maluga, 2025). Whether similar conditions can generate an exotic matter needed for opening wormholes has not been specified so far.

The appearance of exotic R^4s in the quantum regime of spacetime is quite important and requires discussion. Let us turn our attention to topological quantum field theory (TQFT) where the relation of n-cobordisms and Hilbert spaces is in the heart of the constructions. The original formulation in Atiyah, 1988 defines a TQFT symmetric monoidal functor $Z : \mathbf{Cob}_n \to \mathbf{Hilb}$ from the category

Cob$_n$ of n-cobordisms (morphisms) between $n-1$ compact manifolds (objects) to the category of finite-dimensional Hilbert spaces (objects) with linear operators (morphisms). This axiomatic version of TQFT does not allow infinite-dimensional $\mathcal{H}^{\infty}$. The reason is dualizability in **Hilb**, i.e. the existence of evaluation ϵ and coevaluation η

$$\epsilon : \mathcal{H}^* \otimes \mathcal{H} \to \mathbb{C}; \quad \eta : \mathbb{C} \to \mathcal{H}^* \otimes \mathcal{H}.$$

To be well-defined, theese operations require finite-dimensional $\mathcal{H}$ and they allow for essential for Atiyah - Segal TQFT gluing behavior

$$Z(M \cup_N M') = Z(M) \circ Z(N^{-1}) \circ Z(M'). \tag{1}$$

So, dualizability hence finite-dimensional $\mathcal{H}$ is essential. In the context of physics, ϵ corresponds to anihilation, while η corresponds to the birth of states or particles. In our case of nontrivial 5-cobordisms, local representability of atlases in V^B requires $\dim \mathcal{H} = \infty$ since then $B = Bor\mathbb{R}/Null$ is the measure algebra that is universal among all maximal projection algebras in $\mathbb{L}(\mathcal{H}^{\infty})$. We claim that this is not an accident, but rather a very crucial property of the formalism.

First, the extension of TQFT over $\mathcal{H}^{\infty}$ would require complete extensions of the formalism, i.e. the use of higher categories up to ∞-categories (Lurie 2009). Then one has to replace the target category of finite-dimensional Hilbert spaces $(Hilb, \otimes)$ by a symmetric monoidal category in which the objects assigned to 4-manifolds are dualizable in a way to support the gluing property (1). This would be a way toward grasping the nontrivial smooth 5-cobordisms in the TQFT formalism and eventually understanding the nonlocality of entanglement between the regions of spacetime. In fact, the realization of the above would require quitting the TQFT structure and making smooth variants of rigourous categorical field theories where various modifications of the tangential categorical structure would be required.

Second, the inclusion of exotic smoothness on $\mathbb{R}^4$ in the TQFT formalism is again (if possible at all) based on the extension of TQFT over dualizability realized in higher categories, as mentioned above (Grady and Pavlov 2021). In particular, the extended TQFT should refer to infinite dimensional $\mathcal{H}$. This kind of cobordisms between open 4-manifolds would become the main player of the approach.

Our approach strongly indicates the role of $\mathcal{H}^{\infty}$, exotic R^4s and nontrivial 5-cobordisms. However, this comes from completely different points of view based on the automorphisms of B. As we presented before, the ∞ dimension of $\mathcal{H}$ gives rise to the universal $B = Bor(\mathbb{R})/Null$ for entire $\mathbb{L}(\mathcal{H}^{\infty})$ but there is also the level of $Aut\, B$ that distinguishes cases of finite and infinite dimensions.

Lemma 4 *If a maximal complete Boolean algebra of projections in $\mathbb{L}$ is atomic, then all automorphisms of B are extendable to the global automorphisms of $\mathbb{L}$.*

Lemma 5 *If a maximal complete Boolean algebra of projections is atomless then there exist automorphisms of B that are nonextandable over $\mathbb{L}$.*

But $B = Bor(\mathbb{R})/Null$ is atomless, and thus there are local nonextendable automorphisms of B for $\mathcal{H}^{\infty}$. Moreover, these local automorphisms apply as local phases (gauges) in the description of exotic R^4s. Schematically, the local phase represented in Fig. 3a as F where $Aut\, B$ generates $Aut(V^B)$ and gives rise to the change of the exotic structure on R^4_1. The nontrivial 5 subcobordism of two open exotic R^4s emerges (see Fig. 2) which is generated by the inversion of the boundary of the Akbulut cork $\Sigma(2,5,7)$. This inversion $\tau : \Sigma(2,5,7) \to \Sigma(2,5,7)$ is due to the nontrivial local phase between $R^4_1 \to R^4_2$. Thus finally, the extension over infinite dimensional Hilbert spaces and allowing for 5-cobordisms between open 4-manifolds result from the bottom up approach reviewed in this work and lies in the heart of top down approaches in higher categories and TQFT. We believe that this phenomenon is the key toward better understanding the discripancy of QM and GR.

Let us indicate yet another feature connected to the above. The negation of the Tsirelson conjecture says roughly that the infinite dimension of $\mathcal{H}$ allows us to distinguish by finitely many quantum correlation the situation where on $\mathcal{H}^{\infty}$ two parties' measurements of commuting observables and the case where we have factorized Hilbert space $\mathcal{H}_A^{\infty,\text{finite}} \otimes \mathcal{H}_B^{\text{finite},\infty}$ and the measurements are on the factors independently. It appears that the nonfactorizable case is more strongly correlated than the factorizable case. The original prove goes through Touring uncomputability of certain sets of formulas describing both cases and showing that they differ by the degree in Touring classes. Recently, an alternative proof was given by methods of maximal Boolean algebras $Bor(\mathbb{R})/Null$ in the lattice $\mathbb{L}$ and forcing relation in models of ZFC (Król and Asselmeyer-Maluga 2024). This indicates that the coding of infinite dimension of $\mathcal{H}^{\infty}$ along with the ZFC twist of QM are well-suited for analyzing subtle phenomena of randomness in QM.

Acknowledgement

JK approciates the conversations with the participants of the Mathematical Workshop in Kielnarowa in Sept. 2025, since they influenced the ideas presented here.

Competing Interests None.

References

Akbulut, S. 1991(a). A fake compact contractible 4-manifold. *J. Diff. Geom.* 33:335–356.

———. 1991(b). An exotic 4-manifold. *J. Diff. Geom.* 33:357–361.

Akbulut, S., and R. Kirby. 1979. Mazur manifolds. *Michigan Math. Journal.* 26:259.

Atiyah, M. 1988. Topological quantum field theories. *Publ. Math. IHÉS* 68:175–186.

Curtis, C. L., M. H. Freedman, W. C. Hsiang, and R. Stong. 1996. A decomposition theorem for *h*-cobordant smooth simply connected compact 4-manifolds. *Inventiones Mathematicae.* 123:343–348.

Donaldson, S. 1983. An application of gauge theory to four dimensional topology. *J. Diff. Geom.* 18:279–315. https://doi.org/10.4310/jdg/1214437665.

Freedman, H. 1982. The topology of four-dimensional manifolds. *J. Diff. Geom.* 17:357–453. https://doi.org/10.4310/jdg/1214437136.

Gompf, R. 2018. Group actions, corks and exotic smoothings of R^4. *Invent. Math.* 214 (3): 1131–1168.

Gompf, R. E., and A. I. Stipsicz. 1999. *4-Manifolds and Kirby Calculus.* Graduate Studies in Mathematics. Providence, Rhode Island, USA: American Mathematical Society.

Grady, D., and D. Pavlov. 2021. *The geometric cobordism hypothesis.* arXiv: 2111.01095.

Horodecki, R., P. Horodecki, M. Horodecki, and K. Horodecki. 2009. Quantum entanglement. *Rev. Mod. Phys.* 81:865–942. https://doi.org/https://doi.org/10.1103/RevModPhys.81.865.

Król, J., and T. Asselmeyer-Maluga. 2024. Negation of the Tsirelson's conjecture and genericity in quantum mechanics. *Kielnarowa Review* 1 (1): 46–70.

———. 2025. Negation of the Smooth Poincare Conjecture in dimension 4 and negation of the Tsirelson's Conjecture shed light on quantum gravity. *Universe* 11:126. https://doi.org/{10.3390/universe11040126}.

———. 2020. Quantum Mechanics, Formalization and the Cosmological Constant Problem. *Foundations of Science* 25:879–904. https://doi.org/https://doi.org/10.1007/s10699-019-09646-z.

Król, J., T. Asselmeyer-Maluga, K. Bielas, and P. Klimasara. 2017. From Quantum to Cosmological Regime. The Role of Forcing and Exotic 4-Smoothness. *Universe* 3 (31): https://doi.org/https://doi.org/10.3390/universe3020031.

Król, J., K. Bielas, and T. Asselmeyer-Maluga. 2023. Random World and Quantum Mechanics. *Foundations of Science* 28:575–625. https://doi.org/https://doi.org/10.1007/s10699-022-09852-2.

Lurie, J. 2009. On the classification of topological field theories. In *Current developments in mathematics 2008,* 129–280. International Press, Somerville, MA. https://doi.org/10.4310/CDM.2008.v2008.n1.a3. eprint: arXiv:0905.0465.

3. Through the Singularity

Michał Heller, Tomasz Miller and Wiesław Sasin

Kielnarowa Review (2025), **2**, 29–35
doi: https://doi.org/10.57656/kr-2025-0003

Kielnarowa
REVIEW

ARTICLE

Through the Singularity

Michael Heller,[*†] Tomasz Miller,[†] and Wiesław Sasin[‡]

†Copernicus Center for Interdisciplinary Studies, Jagiellonian University, Szczepańska 1/5, 31-011 Cracow, Poland
‡Institute of Mathematics and Cryptology, Military University of Technology, Kaliskiego 2, 00-908 Warsaw, Poland
*Corresponding author. Email: mheller@wsd.tarnow.pl

Abstract

In this work, we propose a dangerous journey – a journey through the strong singularity from one universe to another or from inside of a black hole to its 'inverse' as a white hole. Such singularities are hidden in the Friedman and Schwarzschild solutions; we call them malicious singularities. The journey is made possible owing to two generalizations. The first generalization consists in considering spaces with differential structures on them (so-called ringed spaces) rather than the usual manifolds. This entails a generalization of the concept of smoothness, which allows us to think about a smooth passage through the singularity. The second generalization is related to the concept of curve. We show that if a kind of singularity is implanted in the set of curve's parameters, along with an appropriate topology, in such a way that the structure of the set of parameters corresponds to the structure of the singular space-time, the curve can smoothly – in a generalized sense – pass through the singularity.

Keywords: concept of smoothness, *b*-boundary of space-time,malicious singularity, curves through singularity

1. Introduction

The rapidly developing physics of black holes, both theoretically and observationally, brings the singularity problem into sharp focus once again. In the classical period of studying singularities and formulating theorems about their existence (the 1970s and 1980s of the previous century), it was still possible to neutralize the problem by eliminating the beginning of the universe with the help of various tricks. Today, when observations clearly suggest that there are black holes at the centers of most galaxies, the problem must be faced in all its severity. Until we have a final quantum theory of gravity, classical methods should be fully exploited. They may not only be valuable in themselves, but also point the way to the future quantum gravity theory.

In this work, we want to do the impossible: not only reach the singularity, but also go through it to a new universe, if one exists 'on the other side'. This is not an easy undertaking. All the time, we have in mind strong singularities, such as the ones hidden in the Friedman or Schwarzschild solutions; in the following we will call them malicious singularities. So far, such singularities have been understood as points of a space-time boundary (not belonging to space-time, so its points are not points in the usual sense of the word) at which the histories of observers, photons or other particles terminate. If so, then these histories – by definition – cannot go through the singularity. However, in mathematics, sometimes a small generalization is enough to overcome another conceptual barrier.

In this work, two such generalizations turn out to be crucial. The first concerns the concept of smoothness. We use this concept, as it is understood in the theory of Sikorski's differential spaces Sikorski 1972 and the theory of Grothendieck's sheaf spaces Grothendieck 1958. The essence of this approach is to consider, instead of the usual differential manifolds, a more general class of spaces with differential structures defined on them. These so-called *ringed spaces* will be recalled in Section 2. This method is employed to analyze malicious singularities, regarded as points of the *b*-boundary of space-time. A short review of this issue can be found in Section 3. The second generalization is related to the concept of curve. A smooth curve (in the traditional sense) has no chance of passing through a malicious singularity, but if a kind of singularity is implanted in its set of parameters (along with an appropriate topology), so as the structure of the set of parameters corresponds to the structure of the singular space-time, the curve can smoothly – in a generalized sense – pass through the singularity. This is shown in Section 4.

2. The Concept of Smoothness

Let us recall the standard concept of smoothness as we encounter it in every analysis course. Let M be an n-dimensional differential manifold, and $\mathcal{A} = \{(U_\alpha, \phi_\alpha)\}_{\alpha \in I}$ an atlas on M. A map $f : M \to \mathbb{R}$ is said to be a smooth function on M if, for every $p \in M$, there exixts a chart $(U, \phi) \in \mathcal{A}$ such that $p \in U$ and $f \circ \phi^{-1} : \phi(U) \to \mathbb{R}$ is smooth in the usual sense. The smoothness condition for the transition maps between charts in the atlas guarantees that if f is smooth in one chart (in a neighbourhood of p), it is smooth in any other chart.

Let N be another n-dimensional manifold, and let us consider a map $F : M \to N$. The map F is said to be smooth if, for every $p \in M$, there exists a chart (U, ϕ) on M with $p \in U$, and a chart (V, ψ) on N with $F(p) \in V$, such that $F(U) \subset V$ and $\psi \circ F \circ \phi^{-1} : \phi(U) \to \psi(V)$ is a smooth function on $\mathbb{R}^n$.

Any smooth mapping $F : M \to N$ between manifolds induces linear maps between their tangent spaces $F_{*p} : T_pM \to T_{F(p)}N$. With the help of this map and its dual, vector fields, differential forms and other local geometric data can be transported between manifolds.

To deal with singularities we need a generalization of this smoothness concept. We find it in Sikorski's theory of differential spaces. His idea was to collect all functions that we would like to consider 'smooth' in a family defined on the space M called its differential structure. Of course, for this concept to be reasonable, the family must satisfy suitable conditions. This idea has been implemented in the following way (for a short introduction to the theory of differential spaces see Gruszczak and Heller 1993; Heller and Sasin 1995).

Let $(M, \operatorname{top} M)$ be a topological space, and C a family of functions defined on M (with values in any set). A function f, defined on an open subset $V \subset M$, is said to be a local C-function on V if for every $p \in V$ there exists its open neighbourhood $U \subset V$ and a function $g \in C$ such that $f|U = g|U$. The set of all local C-functions on $V \subset M$ is denoted by C_V.

From the definition it follows that $C|V \subset C_V$, and if $V = M$ then $C \subset C_M$. If $C = C_M$ (every local C-function on M belongs to C), we say that the family C is closed with respect to localization. This is the property we would require for smooth functions. Another such property is the following.

The non-empty family C of real valued functions defined on a set M is said to be closed with respect to composition with smooth functions provided the following condition is satisfied: if $\omega \in C^\infty(\mathbb{R}^n)$ and $f_1, \ldots, f_n \in C$, then $\omega(f_1, \ldots, f_n) \in C$. The family C satisfying this condition is denoted by $\operatorname{sc} C$.

A family C of real functions on M, which is closed with respect to localization and closed with respect to composition with smooth functions is called the differential structure on M, and the pair (M, C) the differential space. All functions belonging to C are smooth *ex definitione*.

A mapping between differential spaces is smooth if it preserves the smoothness understood in the above way, i.e. if (M, C) and (N, D) are differential spaces, then the map $f : M \to N$ is smooth if, for every function $g \in D$, $g \circ f \in C$.

The above understanding of smoothness works well in many cases, but is still ineffective when applied to stronger type singularities. What we need is a 'localization' of this concept. This is achieved in the following way.

Let $(M, \operatorname{top} M)$ be a topological space. The sheaf $\mathcal{C}$ of real continuous functions on $(M, \operatorname{top} M)$ is said to be a differential structure if, for any $U \in \operatorname{top} M$, one has $\operatorname{sc} \mathcal{C}(U) = \mathcal{C}(U)$, i.e. every $\mathcal{C}(U)$ is closed with respect to composition with smooth functions. The pair $(M, \mathcal{C})$ is called the structured or sheaf space (for a theory of structured spaces see Heller and Sasin 1995).

Since the sheaf definition implies that each $\mathcal{C}(U)$ is closed with respect to localization, every $\mathcal{C}(U)$ is a differential structure on U (in the sense of Sikorski).

Let $(M, \mathcal{C})$ and $(N, \mathcal{D})$ be two structured spaces. A mapping

$$F : (M, \mathcal{C}) \to (N, \mathcal{D})$$

is said to be smooth if $F^*(\mathcal{D}) \subset \mathcal{C}$, i.e. for every $U \in \operatorname{top} N$ and $g \in \mathcal{D}(U)$ one has $F^*g := g \circ F \in \mathcal{C}(F^{-1}(U))$.

Structured spaces as objects and smooth mappings between them as morphisms form a category of structured spaces. We will use the concept of smoothness involved in this category in our further considerations regarding going through a malicious singularity.

3. Malicious Singularity

The problem is that the singularity cannot belong to space-time, because the metric and even manifold structures of space-time break down in it but, on the other hand, it must be defined using theoretical tools available within space-time, otherwise, it would not even have a quasi-operational significance. Moreover, the definition of a singularity should also be 'theoretically operational', i.e. it should be possible to prove statements about its occurrence and properties based on it.

As a result of long discussions (described in detail in Tipler, Clarke, and Ellis 1980), it turned out that these criteria are met by the definition of a singularity as a geodesic incompleteness of space-time. An inextendible space-time[1] is geodesically complete if all (timelike and null) geodesics in it can be continued to arbitrarily large values of an affine parameter (in both directions). If this is not the case, a given space-time is said to be geodesically incomplete. An incomplete geodesic represents the history of an object that 'runs into a premature end' Tipler, Clarke, and Ellis 1980, p. 139, and may be regarded as ending at a singularity.

Singularities, understood in this way, can be organized into a kind of boundary of space-time, the so-called g-boundary. The point of the g-boundary of space-time is defined by an equivalence class of incomplete geodesics that meet the condition allowing us to regard them as ending at the same point.

As we can see, geodesical incompleteness is a sufficient condition for the existence of a singularity rather than its full fledged definition, but it well served its purpose since it was possible, with the help of it, to prove several theorems on the existence of singularities Hawking and Ellis 1973. And mostly for this reason, the g-boundary construction has outdistanced other similar constructions such as causal boundary Kronheimer and Penrose 1967, p-boundary Dodson 1967 or essential boundary Clarke 1979, but it was not without problems itself. One of them was that geodesic (in)completeness applies, as the name suggests, only to geodesics, while also timelike curves can, in principle, 'break off' at a singularity. It was Geroch who constructed a geodesically complete space-time containing an inextendible timelike curve of finite length and bounded acceleration Geroch 1968. The latter property is essential since 'an observer with a suitable rocketship and a finite amount of fuel could traverse this curve' Hawking and Ellis 1973, p. 258.

This shortcoming was supposed to be remedied by the construction proposed by Schmidt 1971. Since it plays an important role in our considerations, let us present it briefly (for a detailed account Clarke 1993; Dodson 1978; Schmidt 1971). A connection on space-time M induces a Riemann metric on a connected component of the fibre bundle $O(M)$ of orthonormal frames over M. It does it in such a way as to make horizontal and vertical subspaces orthogonal. This leads to the

1. Space-time (M', g') is an extension of space-time (M, g) if there is an isometric embedding $\rho : M \to M'$. Space-time is inextendible if there is no such embedding with $\rho(M) \neq M'$.

existence of the Riemann metric G on $O(M)$. This metric is not unique but the subsequent steps of the construction do not depend of the choice of the one of admissible metrics. With the help of G we define the distance function and construct the Cauchy completion $\overline{O(M)}$ of $O(M)$. We extend, by continuity, the action of the structural group $O(3,1)$ of the bundle from $O(M)$ to $\overline{O(M)}$ and we form the quotient space $\overline{O(M)}/O(3,1)$ to define

$$\partial_b M := M_* - M = \pi(\overline{O(M)}) - \pi(O(M))$$

where $\pi : \overline{O(M)} \to \overline{O(M)}/O(3,1)$ is the canonical projection, and $M_* = M \cup \partial_b M$ the space-time with its b-boundary.

The elements of the b-boundary of space-time are now the equivalence classes (of a suitable equivalence relation) of not necessarily causal geodesics, but also the equivalence classes of other curves. If, however, b-boundary is restricted to causal geodesics, it changes into g-boundary of space-time. So it seemed that the problem of a 'working definition' of singularity had been solved. It was an unpleasant surprise when Bosshard 1976 and Johnson 1976 independently discovered the following pathologies in which b-boundary are involved. If the singularity in the closed Friedman solution or Schwarzschild solution of Einstein field equations is represented as a point p in the b-boundary, $p \in \partial_b M$, then the fibre $\pi^{-1}(p)$ of the fibre bundle of linear frames over p degenerates to a single point. This is particularly frustrating with respect to Friedman's closed solution, because in that case the beginning and end of the universe would be the same, and only, point of the boundary. Moreover, this b-boundary point turns out not to be Hausdorff separated from the rest of space-time. In fact, topology of this configuration is highly degenerate: the only open neighbourhood of the singularity is the whole of space-time. These conclusions seemed disastrous for the b-boundary construction: what was supposed to be a successful definition of singularity turned out to be a source of new difficulties.

The situation was clarified in Heller and Sasin 1994, 1995. It turned out to be essential to employ the concept of smoothness as it is integrated into the construction of space-time as a structured space. The following theorem is crucial.

Theorem 1. *Let $(M, \mathcal{C})$ be a structured space with the topology τ, the weakest topology in which functions of $\mathcal{C}$ are continuous, and (M_*, τ) a topological space such that $M_* = M \cup \{*\}$, $* \notin M$. And let the following conditions be satisfied*

- $\tau | M = \tau_{\mathcal{C}(M)}$,
- $* \in U \in \tau \Rightarrow U = M$.

Then on the topological space (M, τ) there exists exactly one differential structure $\mathcal{C}_*$ such that $\mathcal{C}_*(M) = \mathcal{C}(M)$ and $\mathcal{C}_*(M_*) = \mathbb{R}$.

Proof. See Heller and Sasin 1995. □

This theorem says that if we are inside space-time M, i.e. in any open set $U \in \tau$, we will not find any pathologies, everything will happen as in a regular (non-singular) space-time. However, if we want to prolong the differential structure $\mathcal{C}$ from M to M_*, this can only be done in a trivial way, that is only constant functions can be prolonged, and we obtain $\mathcal{C}(M_*) = \mathbb{R}$. Therefore, the whole of M_* collapses to a single point.

We also have the following theorem.

Theorem 2. *Let $p \in \partial_b M$. If the fibre $\pi^{-1}(p)$ in the fibre bundle of orthonormal frames over M_* degenerates to a single point, then only constant functions can be prolonged to M_*, i.e. $\mathcal{C}(M_*) = \mathbb{R}$.*

Proof. See Heller and Sasin 1995[2]. □

Singularites to which Theorems 1 and 2 apply we have in anticipation called malicious singularities.

Theorems 1 and 2 at least partially explain 'pathological' situations analyzed by Bosshard 1976 and Johnson 1976. These theorems confirm these 'pathologies': their differential structures indeed consist of only constant functions, and their topology of only one open set (besides empty set). Therefore, the differential structure of the closed Friedman model encodes the beginning and end of the Friedman universe (and all other its points) as just one point. However, as soon as we do not extend the differential structure to the b-boundary (to the singularity), everything happens as in ordinary space-time.

In our previous works, we stopped at these explanations, now we want to pick up this thread and go further. So we ask the question: if we have the apparatus of structural spaces and its possible sharpening at our disposal, is it possible to go (smoothly or not) through the malicious singularity? This will be the subject of our analysis in the following section. To move forward, we first have to design curves that could do the work.

4. Curves in the Singularity

Let $(M_*, \mathcal{C}_{M_*})$, $M_* = M \cup \{*\}$, be a space-time manifold M with the malicious singularity $* \in \partial_b M$, where $\mathcal{C}_{M_*}$ is the sheaf of smooth functions on M_*. The topology on M_* is

$$\operatorname{top} M_* = \operatorname{top} M \cup \{M_*\}$$

where $\{M_*\}$ is the only open subset containing the singularity $*$. The cross-sections of the sheaf $\mathcal{C}_{M_*}$ are of the form

$$\mathcal{C}_{M_*}(M_*) = \mathbb{R}, \qquad \mathcal{C}_{M_*}(U) = C^\infty(U)$$

for $U \in \operatorname{top} M$.

We will analyze curves passing through the singularity. Let us consider two types of smooth curves (smooth – in the sense of structured spaces):

- $\gamma : I \to M_*$ where $I = \langle 0, 1\rangle$ is endowed with the standard topology, $\gamma|\langle 0, 1) \subset M$ and $\gamma(1) = *$,
- $\gamma_* : I_* \to M_*$ where $I_* = \langle 0, 1\rangle$ is endowed with the topology $\operatorname{top}\langle 0, 1) \cup \{I_*\}$ (mimicking the topology of M_*), along with $\gamma_*|\langle 0, 1) \subset M$ and $\gamma_*(1) = *$ as above.

Let us notice that the injection $\imath : I \to I_*$, defined simply as the identity mapping $\imath(t) = t,\ t \in \langle 0, 1\rangle$, is continuous.

The differential structure on I_* is defined as a sheaf on I_*, $\mathcal{C}_{I_*} = \mathbb{R}$, $\mathcal{C}_{I_*}(U) = C^\infty(U)$ for $U \in \operatorname{top}\langle 0, 1)$. Our curves can be thus regarded as morphisms

$$\gamma : (I, C^\infty) \to (M_*, \mathcal{C}_{M_*}),$$
$$\gamma_* : (I_*, \mathcal{C}_{I_*}) \to (M_*, \mathcal{C}_{M_*}).$$

It can be easily seen that the injection $\imath$ is a smooth mapping of parameter spaces of structured spaces (I, C^∞) and $(I_*, \mathcal{C}_{I_*})$. This allows us to define the transformation of the curve γ_* into the curve γ.

$$\gamma = \gamma_* \circ \imath.$$

(see the commutative diagram below).

2. The set of all linear frames over M_*, if there exist in it degenerate fibres, is not a fibre bundle, in the usual sense, but it is a fibre bundle in the category of structured spaces.

$$\begin{array}{ccc} I & \xrightarrow{\gamma} & M_* \\ \downarrow \iota & \nearrow \gamma_* & \\ I_* & & \end{array}$$

The following lemmas present the smoothness properties of curves γ and γ_*.

Lemma 1. *Curve $\gamma : I \to M_*$ is smooth if and only if curve $\gamma|\langle 0,1) : \langle 0,1) \to M$ is smooth.*

Proof. Recall that the smoothness of γ means that $\gamma^*(\mathcal{C}_{M_*}(U)) \subset C^\infty(\gamma^{-1}(U))$ for any $U \in \mathrm{top}\, M$ and $\gamma^*(\mathcal{C}_{M_*}(M_*)) \subset C^\infty(\gamma^{-1}(M_*))$. But since $\mathcal{C}_{M_*}(U) = C^\infty(U)$, the former condition is equivalent to the smoothness of $\gamma|\langle 0,1)$, whereas the latter condition boils down to $\gamma^*(\mathbb{R}) \subset C^\infty(\langle 0,1))$ (with the elements of $\mathbb{R}$ interpreted as the constant functions), what is trivially true. □

Lemma 2. *Curve $\gamma_* : I_* \to M_*$ is smooth if and only if curve $\gamma : I \to M_*$ is smooth.*

Proof. The smoothness of γ_* means that $(\gamma_*)^*(\mathcal{C}_{M_*}(U)) \subset \mathcal{C}_{I_*}(\gamma_*^{-1}(U))$ for any $U \in \mathrm{top}\, M$ and $(\gamma_*)^*(\mathcal{C}_{M_*}(M_*)) \subset \mathcal{C}_{I_*}(\gamma_*^{-1}(M_*))$. By defition of the respective sheaves, the latter condition boils down to $(\gamma_*)^*(\mathbb{R}) \subset \mathbb{R}$, what is trivially true (even as an equality). On the other hand, the former condition reduces to $\gamma^*(C^\infty(U)) \subset C^\infty(\gamma^{-1}(U))$, where we have replaced γ_* with γ because they are equal as maps ($\gamma_*(t) = \gamma(t)$ for all $t \in \langle 0,1\rangle$). But as we have seen in the proof of Lemma 1, this is already equivalent to the smoothness of γ.

□

As we can see, both curves, γ and γ_* are smooth in the category of structured spaces, even though the topologies of their parameter spaces are different, $\mathrm{top}\, I_* \subset \mathrm{top}\, I$.

Let us now consider a curve γ hitting the malicious singularity, such as the one likely hiding in the Big Bang or beyond the horizon of a black hole; alternatively, a curve passing through the singularity from an expanding phase of the universe to its expanding phase or, if you prefer, entering a black hole singularity, and then emerging from the 'other side of it' as of a white hole. For brevity, we will talk about the transition from the universe M_1, through the singularity $*$, to the universe M_2, where $M_1 \cap M_2 = \emptyset$.

We thus have the following situation: $M_* = M_1 \cup \{*\} \cup M_2$ (M_* will also be denoted by $M_1 \star M_2$) with the topology

$$\mathrm{top}\, M_* = \mathrm{top}(M_1 \cup M_2) \cup \{M_*\}$$

where M_* is the only open set containing the singularity.

For any smooth curves $\gamma_1 : (0,1) \to M_1$ and $\gamma_2 : (1,2) \to M_2$ let us now define the curve $\gamma_1 \star \gamma_2 : (0,2) \to M_*$ "passing through" the singularity

$$(\gamma_1 \star \gamma_2)(t) = \begin{cases} \gamma_1(t) \in M_1 & \text{if } t \in (0,1) \\ * & \text{if } t = 1 \\ \gamma_2(t) \in M_2 & \text{if } t \in (1,2) \end{cases}$$

where the set of parameters, denoted $I_* = (0,2)$ in what follows, is endowed with the topology

$$\mathrm{top} I_* = \mathrm{top}((0,1) \cup (1,2)) \cup \{(0,2)\}.$$

As a kind of formal summary of our considerations, we can formulate the following proposition

Proposition 1. *The curve $\gamma_1 \star \gamma_2 : I_* \to M_*$ is continuous and smooth if and only if γ_1 and γ_2 are continuous and smooth.*

Proof. The continuity part is obvious from the very definition of the topologies involved. As for the smoothness part, the reasoning goes along the same lines as in the proofs of Lemmas 1 & 2 with some simple modifications, such as replacing I and M with $(0,1) \cup (1,2)$ and $M_1 \cup M_2$, respectively. □

This method of passing through the singularity opens up yet another problem. The Noether theorem shows that breaking the translational time symmetry leads to violating the principle of energy conservation (see, e.g. Hanca, Tulejab, and Hancova 2004; Kosmann and Schwarzbach 2010). Introducing a singularity to the reparamerization of the curve $\gamma_1 \star \gamma_2$ and synchronizing it with the cosmological singularity creates just such a situation. Although the curve $\gamma_1 \star \gamma_2$ passes smoothly through the singularity (in the sense of the theory of differential spaces), at the "transition" moment anything can happen.

Even more importantly, let us notice that the above construction is very general. In fact, in this way one can "smoothly join" even the curves which do not approach the b-boundary (in particular, nothing in the above construction prohibits γ_1 and γ_2 to be *constant* curves!). It is thus worth investigating whether the above construction can be made sharper in the case of curves that *do* approach the singularity by explicitly involving the b-boundary definition.

Competing Interests None.

References

Bosshard, B. 1976. On the b-boundary of the closed Friedmann model. *Commun. Math. Phys.* 46:263–966.

Clarke, C. J. S. 1979. Boundary definitions. *General Rel. Gravit.* 10:977–980.

———. 1993. *The analysis of space-time singularities.* Natural Science Foundation Research Project on Geometry of Function Space. Cambbridge, UK: Cambridge University Press.

Dodson, C. T. J. 1967. A new bundle completion for parallelizable space-times. *General Rel. Gravit.* 10:969–979.

———. 1978. Space-time edge geometry. *Int. J. Theor. Phys.* 17:389–504.

Geroch, R. P. 1968. What is a singularity in general relativity? *Ann. Phys. (New York)* 48 (4): 526–540.

Grothendieck, A. 1958. *A general theory of fibre spaces with structure sheaf.* Natural Science Foundation Research Project on Geometry of Function Space, Report No 4. Lawrence, USA: University of Kansas.

Gruszczak, J., and M. Heller. 1993. Differential structure of space-sime and its prolongations to singular boundaries. *Int. J. Theor. Phys.* 32 (4): 625–648.

Hanca, J., S. Tulejab, and M. Hancova. 2004. Symmetries and conservation laws: Consequences of Noether's theorem. *American J. Phys.* 72 (4): 428–435.

Hawking, S. W., and G. F. R. Ellis. 1973. *The large-scale structure of space-time.* Cambridge, UK: Cambridge University Press.

Heller, M., and W. Sasin. 1994. The structure of the b-completion of space-time. *General Rel. Gravit.* 26:797–811.

———. 1995. Structured spaces and their application to relativistic physics. *J. Math. Phys.* 36:3644–3662.

Johnson, R. A. 1976. The bundle boundary in some special cases. *J. Math. Phys.* 18:898–902.

Kosmann and Schwarzbach. 2010. *The Noether theorems: Invariance and conservation laws in the twentieth century.* Sources and Studies in the History of Mathematics and Physical Sciences. Springer-Verlag.

Kronheimer, E. H., and R. Penrose. 1967. On the structure of causal spaces. *Proc. Camb. Philos. Soc.* 63:481–501.

Schmidt, B. G. 1971. A new definition of singular points in general relativity. *General Rel. Gravit.* 1:269–280.

Sikorski, R. 1972. *Introduction to differential geometry.* Sources and Studies in the History of Mathematics and Physical Sciences. In Polish. Warszawa, Poland: PWN.

Tipler, F. J., C. J. S. Clarke, and G. F. Ellis. 1980. Singularities and horizons. In *General relativity and gravitation. one hundred years after the birth of albert einstein,* edited by A. Held, 2:97–206. New York – London: Plenum Press.

4. Quantum Weak Values and Harmonic Analysis on Lie Groups

Jean-Pierre Fréché and Dominique Lambert

Kielnarowa Review (2025), **2**, 37–55
doi:https://doi.org/10.57656/kr-2025-0004

Kielnarowa
REVIEW

ARTICLE

Quantum Weak Values and Harmonic Analysis on Lie Groups

Jean-Pierre Fréché*† and Dominique Lambert*‡

†Esphin, University of Namur, Rue de Bruxelles 61, B-5000 Namur, Belgium
‡Esphin and naXys, University of Namur, Rue de Bruxelles 61, B-5000 Namur, Belgium
*Corresponding author. Email: jean-pierre.freche@unamur.be; dominique.lambert@unamur.be

Abstract

The aim of this contribution is to generalize a formula proved by Maurice de Gosson (de Gosson 2017) about weak values in the context of the phase-space formulation of Quantum Mechanics (Rundle and Everitt 2021), in order to express those weak values using tools coming from the harmonic analysis on Lie Groups (Faraut 2006). A general formula which enables us to compute weak values is proved, in which the integration on a Lie Group is substituted to the integration on phase-space, using Haar measures. Then this formula is applied to SU(2) and SO(3) and also to the quotient group G/H, where H is a normal subgroup of G.

Keywords: Grossman-Royer, Weyl-Heisenberg, weak values, Lie groups, pre- and post-selection, Haar measure, group representation, special unitary group SU(2), special orthogonal group SO(3).

1. Introduction

The aim of this contribution is to generalize a formula proved by Maurice de Gosson (de Gosson 2017) about weak values in the context of the phase-space formulation of Quantum Mechanics (Rundle and Everitt 2021), in order to express those weak values using tools coming from the harmonic analysis on Lie Groups (Faraut 2006)).

Maurice de Gosson (de Gosson 2017, pp.151-153) has shown that we can express a weak value in general using the Cross-Wigner transform:

$$\langle \hat{A} \rangle_{\psi_f \psi_i} := \frac{\langle \psi_f | \hat{A} \psi_i \rangle}{\langle \psi_f | \psi_i \rangle} = \int_{\mathbb{R}^{2n}} dz \, \frac{W(\psi_f, \psi_i)(z)}{\langle \psi_f | \psi_i \rangle} \, a(z) \tag{1}$$

where $a(z)$ is the *Weyl symbol* of the operator $\hat{A}$ and $W(\psi_f, \psi_i)$ the *Cross-Wigner Transform* given by:

$$W(\psi_f, \psi_i) = \left(\frac{1}{2\pi\hbar}\right)^n \int_{\mathbb{R}^n} dy \, e^{-i\frac{py}{\hbar}} \, \overline{\psi_f(x - \frac{y}{2})} \, \psi_i(x + \frac{y}{2}) \tag{2}$$

with $z = (x, p)$ a point of the phase-space $\mathbb{R}^{2n}$; x, y and p are vectors.

In the particular case where $\hat{A}$ is, for example, the projector $\Pi_{\psi_r} = \frac{|\psi_r\rangle\langle\psi_r|}{\langle\psi_r|\psi_r\rangle}$

$$\langle \Pi_{\psi_r} \rangle_{\psi_f \psi_i} := \frac{\langle \psi_f | \Pi_{\psi_r} \psi_i \rangle}{\langle \psi_f | \psi_i \rangle} = (2\pi\hbar)^n \int_{\mathbb{R}^{2n}} dz \, \frac{W(\psi_i \psi_f)(z)}{\langle \psi_f | \psi_i \rangle} \, W(\psi_r)(z) \tag{3}$$

The Cross-Wigner transform is related to the Weyl-Heisenberg groups acting on phase-space. This is clear if we express $\langle \hat{A} \rangle_{\psi_f \psi_i}$ as follows, according to de Gosson (de Gosson 2017), pp. 17 (2.1), 151 (12.18)), successively:

$$W(\psi_f, \psi_i) = (\frac{1}{\pi \hbar})^n \ (\hat{R}(z)\psi|\phi)_{L^2}$$

$$\langle \hat{A} \rangle_{\psi_f \psi_i} = \frac{1}{(\psi|\phi)} \int_{\mathbb{R}^{2n}} dz \ a(z) \ W(\psi, \phi)$$

$$\langle \hat{A} \rangle_{\psi_f \psi_i} = \frac{1}{(\pi \hbar)^n} \int_{\mathbb{R}^{2n}} dz \ a(z) \ \frac{\langle \hat{R}(z)\psi_f|\psi_i \rangle}{\langle \psi_f|\psi_i \rangle} \tag{4}$$

The operator $\hat{R}$ is the so-called Grossman-Royer operator which is nothing but the Stratonovich-Weyl kernel which is well-know in the generalization of Wigner and cross-Wigner transforms (Gadella et al. 1991; Varilly 1989).

Our aim will be to generalize this formula in the case the Weyl-Heisenberg group is replaced by a Lie group (satisfying some constraints in order for the formula to be well-defined). The generalized formula will be:

$$\langle \hat{A} \rangle_{\psi_f \psi_i} = \frac{1}{\lambda^2} \int_G d\mu(g) \ Tr[\hat{A}U(g)] \ \frac{Tr[|\psi_i\rangle\langle\psi_f|U^\dagger(g)]}{\langle \psi_f|\psi_i \rangle} \tag{5}$$

where $d\mu(g)$ is the Haar measure on G (we suppose it unimodular, i.e. its left-invariant Haar measure is equal to its right-invariant Haar measure) and λ is a constant related to the dimension of the irreducible unitary and $\mathbb{C}$-linear representation U of G. The term $\frac{Tr[U^\dagger(g)|\psi_i\rangle\langle\psi_f|}{\langle\psi_f|\psi_i\rangle}$ can be identified to a quasi-distribution of probabilities (Brif and Mann 1998); (Abgaryan, Khvedelidze, and Torosyan 2019) whose values can be negative. If $|\psi_i\rangle$ is an admissible vector in the sense of the generalized coherent states (defined by groups acting on a very specific vector called *an admissible one*, see below), the term $Tr[U^\dagger(g)|\psi_i\rangle\langle\psi_f|] = \langle\psi_f|U^\dagger(g)|\psi_i\rangle$ takes the sense of a coherent state transform (which becomes wavelets transform in some particular choice of group: describing translation and change of scale).We refer here to the book of S. t. Ali, J.-P Antoine, J.-P. Gazeau (Ali, Antoine, and Gazeau 2000). The term $Tr[\hat{A}U(g)]$ plays the role of a Weyl symbol of $\hat{A}$.

One can wonder why it would be amenable to rewrite the weak value $\langle \hat{A} \rangle_{\psi_f \psi_i}$ in these terms. The answer of such a question is that it allows us to perform its harmonic analysis connected with a G-symmetry. Let us note that the Grossman-Royer is not so easy to write when we pass from the Weyl-Heisenberg group to an arbitrary unimodular Lie group with a square-integrable representation. If you give a group, then knowing the irreducible and square integrable representation $G \longrightarrow EndG \ \ g \longrightarrow \ U(g)$, you can immediately write the weak value without having to build an analog of the Grossman-Royer operator if any.

We have to say that some attempts were made to get the formula we give here. The work of F. Antonsen (Antonsen 1998) is very interesting and inspiring, from this point of view, but the formula he proposed does not seem to be the right one (it differs by a hermitian conjugate, but is crucial to us).

Our formula could also be adapted in the case of a symmetric space described by a coset G/H, where H is a subgroup of G. In some interesting particular cases, this coset can be endowed with a Kählerian structure and thus could mimic a generalized phase-space as it is done in the geometric quantization. It is worth noting that if our vector $|\psi_i\rangle$ is an admissible vector invariant under the subgroup H, the coset G/H is nothing but a set of coherent states.

This formula leads finally to corollaries, one of them being a formula introducing a kind of Moyal product (Varilly and Gracia-Bonda 1989).

2. The main formula and its proof

In this section we have to use a lemma:

Lemma. Let $\mathfrak{H}$ be a Hilbert space and $\{|i\rangle, |j\rangle, |k\rangle, \ldots\}$ an infinitely countable orthogonal basis of $\mathfrak{H}$. Let also $\hat{A}$ be a linear (bounded) operator $\mathfrak{H} \longrightarrow \mathfrak{H}$. It is always possible to write

$$\hat{A} = \sum_{ij} \alpha^{ij} |i\rangle\langle j| \tag{6}$$

where $\alpha^{ij} \in \mathbb{C}$.

Indeed, let $|\psi\rangle$ be a ket of $\mathfrak{H}$. We have:

$$|\psi\rangle = \sum_{j} \psi^j |j\rangle \tag{7}$$

where $\psi^j = \langle j|\psi\rangle$. We can express $\hat{A}|j\rangle$ in terms of his components:

$$\hat{A}|j\rangle = \sum_{i} \alpha^{ij} |i\rangle \tag{8}$$

Thus we can write successively:

$$\begin{aligned} \hat{A}|\psi\rangle &= \sum_{j} \psi^j \sum_{i} \alpha^{ij} |i\rangle \\ &= \sum_{ij} \alpha^{ij} |i\rangle\langle j|\psi\rangle \\ &= \left(\sum_{ij} \alpha^{ij} |i\rangle\langle j| \right) |\psi\rangle \end{aligned}$$

And so we obtain, as expected:

$$\hat{A} = \sum_{ij} \alpha^{ij} |i\rangle\langle j|$$

Theorem. Let $|\eta\rangle$ and $|\varphi\rangle$ be two states of the system: $|\eta\rangle, |\varphi\rangle \in \mathfrak{H}$. Then one has:

$$\lambda^2 \langle \eta|\hat{A}|\varphi\rangle = \int_G d\mu(g)\ Tr[\hat{A}\ U(g)]\ Tr[\ |\varphi\rangle\langle\eta|\ U^{\dagger}(g)] \tag{9}$$

where $\mu(g)$ is the so-called *left-invariant Haar-measure*.

In order to prove this theorem, let us begin – it is important – by writing the *orthogonality relations* for these two states explicitly and clarify the context[1] :

$$\langle \hat{C}\eta|\hat{C}\eta'\rangle \ \langle \varphi'|\varphi\rangle = \int_G d\mu(g) \ \overline{\langle \eta'_g|\varphi'\rangle} \ \langle \eta_g|\varphi\rangle \tag{10}$$

where

$$\begin{aligned} |\eta_g\rangle &= U(g)|\eta\rangle \\ |\eta'_g\rangle &= U(g)|\eta'\rangle \\ \hat{C} &= \lambda\mathbb{1} \end{aligned}$$

and where $U(g)$ is a square-integrable representation of a locally compact Lie group G on $\mathfrak{H}$, with $\lambda > 0$. Moreover, $|\eta\rangle$ and $|\eta'\rangle$ must be *admissible kets*, i.e.

$$I(\eta) := \int_G d\mu(g)|\langle U(g)\eta|\eta\rangle|^2 = \int_G d\mu(g)|\langle \eta|U(g)|\eta\rangle|^2 < \infty \tag{11}$$

Starting from (10) we can write

$$\begin{aligned} \langle \lambda\mathbb{1}\eta|\lambda\mathbb{1}\eta'\rangle \ \langle \varphi'|\varphi\rangle &= \int_G d\mu(g) \ \langle \varphi'|U(g)|\eta'\rangle \ \langle \eta|U^\dagger(g)|\varphi\rangle \\ \lambda^2 \ \langle \eta|\eta'\rangle \ \langle \varphi'|\varphi\rangle &= \int_G d\mu(g) \ \langle \varphi'|U(g)\mathbb{1}|\eta'\rangle \ \langle \eta|U^\dagger(g)\mathbb{1}|\varphi\rangle \end{aligned}$$

Let us put, as a particular case, $|\eta'\rangle = |i\rangle$ and $|\varphi'\rangle = |j\rangle$ (two basis kets of $\mathfrak{H}$), and insert the resolution of the identity:

$$\begin{aligned} \lambda^2 \ \langle \eta|i\rangle \ \langle j|\varphi\rangle &= \int_G d\mu(g) \ \langle j|U(g) \left(\sum_k |k\rangle \ \langle k|\right) |i\rangle \ \langle \eta|U^\dagger \left(\sum_l |l\rangle \ \langle l|\right) |\varphi\rangle \\ &= \int_G d\mu(g) \ \sum_k \langle k|i\rangle \ \langle j|U(g)|k\rangle \ \sum_l \langle l|\varphi\rangle\langle \eta|U^\dagger(g)|l\rangle \end{aligned}$$

Two traces appear clearly on the right side ; so we may conclude :

$$\lambda^2 \ \langle \eta|i\rangle \ \langle j|\varphi\rangle = \int_G d\mu(g) \ Tr[|i\rangle\langle j|U(g)] \ Tr[|\varphi\rangle\langle \eta|U^\dagger(g)] \tag{12}$$

This formula expresses a property of the operator $|i\rangle\langle j|$. Using our lemma, we can generalize. Let α^{ij} be complex numbers for every pair i, j of indexes labelling the basis-kets of $\mathfrak{H}$; We can write successively :

$$\lambda^2 \ \langle \eta|\alpha^{ij}|i\rangle \ \langle j|\varphi\rangle = \int_G d\mu(g) \ Tr[\alpha^{ij}|i\rangle\langle j|U(g)] \ Tr[|\varphi\rangle\langle \eta|U^\dagger(g)]$$

$$\lambda^2 \ \langle \eta| \left(\sum_{ij} \alpha^{ij}|i\rangle \ \langle j|\right) |\varphi\rangle = \int_G d\mu(g) \ Tr[\left(\sum_{ij} \alpha^{ij}|i\rangle\langle j|\right) U(g)] \ Tr[|\varphi\rangle\langle \eta|U^\dagger(g)]$$

1. This theorem can be found in (Ali, Antoine, and Gazeau 2000), p.156.

The two pairs of brackets contain the general expression of the operator $\hat{A}$; so we have obtained the following :

$$\lambda^2 \langle\eta|\hat{A}|\varphi\rangle = \int_G d\mu(g)\ Tr[U(g)\hat{A}]\ Tr[|\varphi\rangle\langle\eta|U^\dagger(g)] \tag{13}$$

This formula enables us to obtain the *weak values* of the operator $\hat{A}$ by means of traces and of the representation $U(g)$. Indeed, taking $|\eta\rangle = |\psi_f\rangle$ and $|\phi\rangle = |\psi_i\rangle$ (respectively, post- and pre-selected states) and dividing by $\langle\psi_f|\psi_i\rangle$, we also obtain:

$$\lambda^2 \frac{\langle\psi_f|\hat{A}|\psi_i\rangle}{\langle\psi_f|\psi_i\rangle} = \langle\hat{A}\rangle_{\psi_f\psi_i} = \int_G d\mu(g)\ Tr[U(g)\hat{A}]\ \frac{Tr[|\psi_i\rangle\langle\psi_f|U^\dagger(g)]}{\langle\psi_f|\psi_i\rangle} \tag{14}$$

which is nothing but (5).

3. Some corollaries

We enunciate the corollaries : the proofs are obvious, and we do not give them. Let us first introduce a new function: the so-called *generalised Weyl function* (gWf) of the operator $\hat{A}$:

$$W_{\hat{A}}(g) := Tr[\hat{A}\ U(g)] \tag{15}$$

Corollary 1 We get the following relation between gWf and traces of operator. As above, such traces will play an important role:

$$\lambda^2\ Tr[\hat{A}\ \hat{B}^\dagger] = \int_G d\mu(g)\ W_{\hat{A}}(g)\ W_{\hat{B}}(g) \tag{16}$$

Corollary 2 Another relationship the gWf and traces of operators can be proved:

$$\lambda^2\ W_{\hat{A}}(g') = \int_G d\mu(g)\ W_{\hat{A}}(g)\ Tr[U(g)\ U^\dagger(g')] \tag{17}$$

It should be noted that the trace on the left-hand side plays the role of *reproducing kernel* for $W_{\hat{A}}(g)$ if we define:

$$\lambda^2\ K(g,g') = Tr[U(g)\ U^\dagger(g')] \tag{18}$$

then in this way we can write:

$$\int_G d\mu(g)\ K(g,g')\ W_{\hat{A}}(g) = W_{\hat{A}}(g') \tag{19}$$

Now a third formula, wich can be introduced by defining first a new product (similar to a Moyal-Product).

Corollary 3 Let F and G be two functions belonging to $L^2(G, d\mu)$. We put

$$(F\star L)(g) := \int_G\int_G d\mu(g')\ d\mu(g'')\ \frac{1}{\lambda^4}\ \overline{Tr[Ug')\ U(g'')\ U^\dagger(g)]}\ F(g')\ L(g'') \tag{20}$$

We can establish the following:

$$W_{\hat{A}\hat{B}}(g) = (W_{\hat{A}} \star W_{\hat{B}})(g) \tag{21}$$

wich means that the gWf of the product of two functions is equal to $\star$ − *product* of the gWf of the functions.

Corollary 4. It is a particular case of (17) for $\hat{B} = \mathbb{1}$:

$$\lambda^2 \, Tr\hat{A} = \int_G d\mu(g) \, W_{\hat{A}}(g) \tag{22}$$

Corollary 5. It is another particular case of (13) for $\hat{A} = U^\dagger(g')$:

$$\lambda^2 \, W_{\hat{A}}(g') = \int_G d\mu(g) \, W_{\hat{A}}(g) \, Tr[U(g) \, U^\dagger(g')] \tag{23}$$

Corollary 6. Let H be the maximal compact subgroup of G, let $\Omega \in G/H$ $(G = \Omega\, H)$. Let also the two vectors $|\eta\rangle$ and $|\eta'\rangle$ be such that $U(h)|\eta\rangle = |\eta\rangle$ and $U(h)|\eta'\rangle = |\eta'\rangle$. Then we get :

$$\frac{\lambda^2}{Vol(H)} \langle\eta|\eta'\rangle \, \langle\varphi'|\varphi\rangle = \int_{G/H} d\mu(\Omega) \, (\overline{W(\eta', \varphi')(\Omega)} \, W(\eta, \varphi)(\Omega) \tag{24}$$

where

$$W(\eta, \varphi)(\Omega) := Tr[U^\dagger(\Omega)|\varphi\rangle\langle\eta|]U^\dagger(g)|\varphi\rangle = \langle\eta|U^\dagger(\Omega)|\varphi\rangle \tag{25}$$

and hence

$$\frac{\lambda^2}{Vol(H)} Tr(\hat{B}\hat{A}) = \int_{G/H} d\mu(\Omega) \, \overline{W_{\hat{B}^\dagger}(\Omega)} \, W_{\hat{A}}(\Omega) \tag{26}$$

4. Two examples, G = SU(2) and G = SO(3).

In this section we try to show that our formula (13) can be applied to two important groups in Physics; we prove in every case that, with a suitable choice of the measure μ, both left-hand and right-hand sides of the equation are truly equal.

4.1. The main formula and SU(2). Let G be an abstract group represented by $SU(2)$, i.e. there exists a morphism $U: \; G \longrightarrow SU(2) \quad g \longrightarrow U(g)$ and two conditions: $U(g_1 g_2) = U(g_1)U(g_2)$ and $U(g)U^\dagger(g) = U^\dagger(g)U(g) = \hat{I}$ (identity) for all $g, \; g_1, \; g_2 \; \in$ G. We can write

$$U(g) = \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix} \tag{27}$$

$$U^\dagger(g) = \begin{pmatrix} \overline{\alpha} & -\beta \\ \overline{\beta} & \alpha \end{pmatrix} \tag{28}$$

And thus we have:

$$U(g)U^\dagger(g) = U^\dagger(g)U(g) = \begin{pmatrix} \alpha\overline{\alpha} + \beta\overline{\beta} & 0 \\ 0 & \alpha\overline{\alpha} + \beta\overline{\beta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

if and only if $\alpha\overline{\alpha} + \beta\overline{\beta} = 1$. This condition is fulfilled with:

$$\begin{aligned} \alpha &= x_1 + ix_2 \\ \beta &= x_3 + ix_4 \end{aligned} \tag{29}$$

where

$$\begin{aligned} x_1 &= cos\theta \\ x_2 &= sin\theta\ cos\phi \\ x_3 &= sin\theta\ sin\phi\ cos\psi \\ x_4 &= sin\theta\ sin\phi\ sin\psi \end{aligned} \tag{30}$$

Indeed, $\alpha\overline{\alpha} + \beta\overline{\beta} = (x_1^2 + x_2^2) + (x_3^2 + x_4^2) = 1$.
When G is *the group of rotations*, this definition seems to be very natural and we have $\theta \in [0, \pi]$, $\phi \in [0, \pi]$ and $\psi \in [0, 2\pi]$.
Let

$$U(g_1) = \begin{pmatrix} \alpha_1 & \beta_1 \\ -\overline{\beta}_1 & \overline{\alpha}_2 \end{pmatrix}$$

and

$$U(g_2) = \begin{pmatrix} \alpha_2 & \beta_2 \\ -\overline{\beta}_2 & \overline{\alpha}_2 \end{pmatrix}$$

The product is

$$U(g_1)U(g_2) = \begin{pmatrix} \alpha_1\alpha_2 - \beta_1\overline{\beta}_2 & \alpha_1\beta_2 + \beta_1\overline{\alpha}_2 \\ -\overline{\beta}_1\alpha_2 - \overline{\alpha}_1\overline{\beta}_2 & -\overline{\beta}_1\beta_2 + \overline{\alpha}_1\overline{\alpha}_2 \end{pmatrix}$$

If we define $\alpha_3 = \alpha_1\alpha_2 - \beta_1\overline{\beta}_2$ and $\alpha_1\beta_2 + \beta_1\overline{\alpha}_2$, we see that

$$U(g_1)U(g_2) = \begin{pmatrix} \alpha_3 & \beta_3 \\ -\overline{\beta}_3 & \overline{\alpha}_3 \end{pmatrix} \tag{31}$$

That is exactly the same form as (27), as needed.
Our choice of the measure will be the so-called *Haar measure* :

$$d\mu(g) = \frac{1}{2\pi^2}\ sin^2\theta\ sin\phi\ d\theta\ d\phi\ d\psi \tag{32}$$

The vectors will be of the form

$$|\eta\rangle = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} \tag{33}$$

$$|\varphi\rangle = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}$$

where $\eta_i, \varphi_i \in \mathbb{C}$. Their scalar product $\langle\ .\ |\ .\ \rangle$ will be *represented* by the matrix product. Provided with the usual addition law of two vectors, the set of all these vectors forms a two-dimensional complex

Hilbert space $\mathfrak{H}$, whose basis vectors are $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$. However, are these vectors admissible in the sense of (11) ? Yes. Indeed, it is easy to write

$$|\langle\eta|U(g)\eta\rangle|^2 = |\overline{\eta}_1\alpha\eta_1 \; - \overline{\eta}_2\overline{\beta}\eta_1 \; + \; \overline{\eta}_1\beta\eta_2 \; + \; \overline{\eta}_2\overline{\alpha}\eta_2|^2$$

Clearly, the right-hand side is finite (being composed only of a sinus and a cosinus). The integration on $SU(2)$ is made of integrations between 0 and π or 2π, whose results are necessarily finite.
As $\hat{A}$, we choose:

$$\hat{A} = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \tag{34}$$

where all $a_i \in \mathbb{C}$. We must ensure that, representing an observable, the matrix $\hat{A}$ is *hermitian*, i.e. $\hat{A}^\dagger = \hat{A}$, which can be written as follows:

$$\hat{A} = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} = \hat{A}^\dagger = \begin{pmatrix} \overline{a}_1 & \overline{a}_3 \\ \overline{a}_2 & \overline{a}_4 \end{pmatrix}$$

Thus $a_1, a_4 \in \mathbb{R}$. Moreover, $a_2 = \overline{a}_3$ and (which is the same), $a_3 = \overline{a}_2$. So, $\hat{A}$ must be rewritten:

$$\hat{A} = \begin{pmatrix} a_1 & a_2 \\ \overline{a}_2 & a_4 \end{pmatrix} \tag{35}$$

Let us now compute all the detailed elements of the relation (13). Successively:

$$|\varphi\rangle\langle\eta| = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} \begin{pmatrix} \overline{\eta}_1 & \overline{\eta}_2 \end{pmatrix} = \begin{pmatrix} \varphi_1\overline{\eta}_1 & \varphi_1\overline{\eta}_2 \\ \varphi_2\overline{\eta}_1 & \varphi_2\overline{\eta}_2 \end{pmatrix} \tag{36}$$

It should be noted that the product $|\varphi\rangle\langle\eta|$ has been *represented* here by the usual tensor product of two matrices.
Our task is now to establish the relevance of the formula (13). Let us compute separately her left-hand side and her right-hand side.

Left-hand side:

$$\begin{aligned} \lambda^2\langle\eta|\hat{A}|\varphi\rangle &= \lambda^2 \begin{pmatrix} \eta_1 \; \eta_2 \end{pmatrix} \begin{pmatrix} a_1 & a_2 \\ \overline{a}_2 & a_4 \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} \\ &= \lambda^2(\overline{\eta}_1 a_1 \varphi_1 + \overline{\eta}_2\overline{a}_2\varphi_1 + \overline{\eta}_1 a_2\varphi_2 + \overline{\eta}_2 a_4 \varphi_2) \end{aligned} \tag{37}$$

Right-hand side
Successively :

$$U(g)\hat{A} = \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix} \begin{pmatrix} a_1 & a_2 \\ \overline{a}_2 & a_4 \end{pmatrix} = \begin{pmatrix} \alpha a_1 + \beta\overline{a}_2 & \alpha a_2 + \beta a_4 \\ -\overline{\beta}a_1 + \overline{\alpha}\overline{a}_2 & -\overline{\beta}a_2 + \overline{\alpha}a_4 \end{pmatrix}$$

Note that we have *represented* the composed of the two operators $U(g)$ and $\hat{A}$ by their usual matrix product. We would insist on the fact that it is *a choice*. An other choice, for instance their tensor product, would have been possible. The ultimate justification of our choice and all the other choices we have made lies in the relevance of the formula (13) that we try to establish. Now:

$$Tr[U(g)\hat{A}] = \alpha a_1 + \beta \bar{a}_2 - \bar{\beta} a_2 + \bar{\alpha} a_4 \tag{38}$$

$$\begin{aligned}|\varphi\rangle\langle\eta|U^\dagger(g) &= \begin{pmatrix}\varphi_1\bar{\eta}_1 & \varphi_1\bar{\eta}_2\\ \varphi_2\bar{\eta}_1 & \varphi_2\bar{\eta}_2\end{pmatrix}\begin{pmatrix}\bar{\alpha} & -\beta\\ \bar{\beta} & \alpha\end{pmatrix}\\ &= \begin{pmatrix}\varphi_1\bar{\eta}_1\bar{\alpha}+\varphi_1\bar{\eta}_2\bar{\beta} & -\varphi_1\bar{\eta}_1\beta+\varphi_1\bar{\eta}_2\alpha\\ \varphi_2\bar{\eta}_1\bar{\alpha}+\varphi_2\bar{\eta}_2\bar{\beta} & -\varphi_2\bar{\eta}_1\beta+\varphi_2\bar{\eta}_2\alpha\end{pmatrix}\end{aligned}$$

$$Tr[|\varphi\rangle\langle\eta|U^\dagger] = \varphi_1\bar{\eta}_1\bar{\alpha} + \varphi_1\bar{\eta}_2\bar{\beta} - \varphi_2\bar{\eta}_1\beta + \varphi_2\bar{\eta}_2\alpha \tag{39}$$

On the right-hand side of (13), the integrand is a product of three factors:

$$d\mu(g) \quad\times\quad Tr[U(g)\hat{A}] \quad\times\quad Tr[|\phi\rangle\langle\eta|U^\dagger(g)]$$

If we are explicit, we must write by introducing (29) and (30):

$$\begin{aligned}&d\theta\ d\phi\ d\psi\ \frac{1}{2\pi^2}\ sin^2\theta\ sin\phi\\ &\times\Big(a_1(cos\theta + sin\theta\ cos\phi) + \bar{a}_2(sin\theta\ sin\phi\ cos\psi + i\ sin\theta\ sin\phi\ sin\psi\\ &-a_2(sin\theta\ sin\phi\ cos\psi - i\ sin\theta\ sin\phi\ sin\psi) + a_4(cos\theta - i\ sin\theta\ cos\phi)\Big)\\ &\times\Big(\varphi_1\bar{\eta}_1(cos\theta - i\ sin\theta\ cos\phi) + \varphi_1\bar{\eta}_2(sin\theta\ sin\phi\ cos\psi - i\ sin\theta\ sin\phi\ sin\psi\\ &-\varphi_2\bar{\eta}_1(sin\theta\ sin\phi\ cos\psi + i\ sin\theta\ sin\phi\ sin\psi) + \varphi_2\bar{\eta}_2(cos\theta + i\ sin\theta\ cos\phi)\Big)\end{aligned}$$

Each pair of big brakets contains 8 terms; thus the product contains 64 terms, but unfortunately it is not useful to write each of them. Why? In the course of the triple integration, many of them will give a null contribution. The reason lies in the following definite integrals (which all clearly occur in the integration process).

$$\int_0^\pi d\theta\ sin^3\theta\ cos\theta = 0$$

$$\int_0^\pi d\phi\ cos\phi = \int_0^\pi d\phi\ cos\phi\ sin\phi = 0$$

$$\int_0^{2\pi} d\psi\ cos\psi = \int_0^{2\pi} d\psi\ sin\psi = \int_0^{2\pi} d\psi\ sin\psi\ cos\psi = 0$$

The only terms of the integrand that contribute to the final value of the right-hand side of (13) can now be written:

$$\begin{aligned}d\theta\ d\phi\ d\psi\ \frac{1}{2\pi^2}\ sin^2\theta\ sin\phi\ \Big(&a_1\varphi_1\bar{\eta}_1(cos^2\theta + sin^2\theta\ cos^2\phi) + a_1\varphi_2\bar{\eta}_2(cos^2\theta - sin^2\theta\ cos^2\phi)\\ &+\bar{a}_2\varphi_1\bar{\eta}_2(sin^2\theta\ sin^2\phi\ cos^2\psi + sin^2\theta\ sin^2\phi\ sin^2\psi)\\ &-\bar{a}_2\varphi_2\bar{\eta}_1(sin^2\theta\ sin^2\phi\ cos^2\psi - sin^2\theta\ sin^2\phi\ sin^2\psi)\\ &-a_2\varphi_1\bar{\eta}_2(sin^2\theta\ sin^2\phi\ cos^2\psi - sin^2\theta\ sin^2\phi\ sin^2\psi)\\ &-a_2\varphi_2\bar{\eta}_1(sin^2\theta\ sin^2\phi\ cos^2\psi + sin^2\theta\ sin^2\phi\ sin^2\psi)\\ &+a_4\varphi_2\bar{\eta}_2(cos^2\theta + \sin^2\theta\ cos^2\phi)\Big)\end{aligned}$$

The integration on ψ introduces a multiplicative factor 2π ; if we take into account that $cos^2\psi + sin^2\psi = 1$ and $cos^2\psi - sin^2\psi = cos2\psi$, the new integrand must now be rewritten as

$$
\begin{aligned}
d\theta\, d\phi\, \frac{1}{2\pi^2}\, sin^2\theta\, sin\phi \times 2\pi\Big(a_1\varphi_1\bar{\eta}_1(cos^2\theta + sin^2\theta\, cos^2\phi) + a_1\varphi_2\bar{\eta}_2(cos^2\theta - sin^2\theta\, cos^2\phi) \\
+\bar{a}_2\varphi_1\bar{\eta}_2(sin^2\theta\, sin^2\phi) \\
-\bar{a}_2\varphi_2\bar{\eta}_1(sin^2\theta\, sin^2\phi\, cos2\psi) \\
-a_2\varphi_1\bar{\eta}_2(sin^2\theta\, sin^2\phi\, cos2\psi) \\
-a_2\varphi_2\bar{\eta}_1(sin^2\theta\, sin^2\phi) \\
+a_4\varphi_2\bar{\eta}_2(cos^2\theta + sin^2\theta\, cos^2\phi)\Big)
\end{aligned}
$$

The fourth and fifth terms of this sum contain $cos2\psi$. The integration on ψ from 0 to 2π gives $\int_0^{2\pi} d\psi\, cos2\psi = 0$. Therefore, these two terms do not contribute to the final result and we can rewrite the terms of the integrand who really contributes to the right-hand side of (13) as

$$
\begin{aligned}
d\theta\, d\phi \times \\
\frac{1}{\pi}\Big(a_1\varphi_1\bar{\eta}_1(sin^2\theta\, \cos^2\theta\, sin\phi + sin^4\theta\, cos^2\phi\, sin\phi) + a_1\varphi_2\eta_2(sin^2\theta\, cos^2\theta\, sin\phi - sin^4\theta\, cos^2\phi\, sin\phi) \\
+ \bar{a}_2\varphi_1\bar{\eta}_2\, sin^4\theta\, sin^3\phi + a_2\varphi_2\bar{\eta}_1 sin^4\theta\, \sin^3\phi \\
a_4\varphi_1\bar{\eta}_1(sin^2\theta\, cos^2\theta\, sin\phi - sin^4\theta\, cos^2\phi\, sin\phi) + a_4\varphi_2\bar{\eta}_2(sin^2\theta\, cos^2\theta\, sin\phi + sin^4\theta\, cos^2\phi\, sin\phi)
\end{aligned}
$$

In order to provide the final result of the calculation, we need the following definite integrals:

$$
\begin{aligned}
\int_0^{\pi} d\theta\, sin^2\theta = \int_0^{\pi} d\phi\, sin^2\phi &= \frac{\pi}{2} \\
\int_0^{\pi} d\theta\, sin^4\theta &= \frac{3\pi}{8} \\
\int_0^{\pi} d\theta\, sin\theta &= 2 \\
\int_0^{\pi} d\phi\, sin\phi\, cos^2\phi &= \frac{2}{3} \\
\int_0^{\pi} d\theta\, sin^2\theta\, cos^2\theta &= \frac{\pi}{8} \\
\int_0^{\pi} d\phi\, sin^3\phi &= \frac{4}{3}
\end{aligned}
$$

Final result : the right-hand side of (13):

$$
\begin{aligned}
\frac{1}{\pi}\Big(a_1\varphi_1\bar{\eta}_1(\frac{\pi}{8} \times 2 + \frac{3\pi}{8} \times \frac{2}{3}) + a_1\varphi_2\bar{\eta}_2(\frac{\pi}{8} \times 2 - \frac{3\pi}{8} \times \frac{2}{3}) + \bar{a}_2\varphi_1\bar{\eta}_2 \times \frac{3\pi}{8} \times \frac{4}{3}) + a_2\varphi_2\bar{\eta}_1\frac{3\pi}{8} \times \frac{4}{3} \\
+ a_4\varphi_1\bar{\eta}_1(\frac{\pi}{8} \times 2 - \frac{3\pi}{8} \times \frac{2}{3}) + a_4\varphi_2\, \bar{\eta}_2(\frac{\pi}{8} \times 2 + \frac{3\pi}{8} \times \frac{2}{3})\Big) \\
= \frac{1}{\pi}\Big(a_1\varphi_1\bar{\eta}_1 \times \frac{\pi}{2} + a_2\varphi_1\bar{\eta}_2 \times \frac{\pi}{2} + a_2\varphi_2\bar{\eta}_1\frac{\pi}{2} + a_4\varphi_2\bar{\eta}_2 \times \frac{\pi}{2}\Big) \\
= \frac{1}{2}\Big(a_1\varphi_1\bar{\eta}_1 + \bar{a}_2\varphi_1\bar{\eta}_2 + a_2\varphi_2\bar{\eta}_1 + a_4\varphi_2\bar{\eta}_2\Big)
\end{aligned}
$$

And thus we get

$$
\int_{SU(2)} d\mu(g)\, Tr[U(g)\hat{A}]\, Tr[|\varphi\rangle\langle\eta|U^{\dagger}(g)] = \frac{1}{2}\Big(a_1\varphi_1\bar{\eta}_1 + \bar{a}_2\varphi_1\bar{\eta}_2 + a_2\varphi_2\bar{\eta}_1 + a_4\varphi_2\bar{\eta}_2\Big) \quad (40)
$$

This is exactly what we had obtained in (37), except for the multiplicative factor λ^2. The comparison between (37) and (40) tells us, as expected, that:

$$\lambda^2 \langle \eta | \hat{A} | \varphi \rangle = \int_{SU(2)} d\mu(g)\; Tr[U(g)\hat{A}]\; Tr[|\varphi\rangle\langle\eta| U^\dagger(g)] \tag{41}$$

provided $\lambda^2 = \frac{1}{2}$ or

$$\lambda = \frac{1}{\sqrt{2}} \tag{42}$$

Furthermore we know that $dim\, \mathfrak{H} = 2$ (the dimension of the Hilbert space $\mathfrak{H}$ we have initially chosen). We may thus conclude that:

$$\lambda = \frac{1}{\sqrt{dim\, \mathfrak{H}}} \tag{43}$$

This last relation is in perfect agreement with the relation (8.49) of (Ali, Antoine, and Gazeau 2000): as far as $SU(2)$ is concerned, our goal is achieved.
Among all observables that are worth considering, are the square and the z-component of the spin, namely $\hat{S}^2 = (\hbar^2/4)\sigma^2$ and $S_z = (\hbar/2)\sigma_z$, for wich:

$$\begin{aligned} \sigma^2 &= 4\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 4\sigma_0 \\ \sigma_z &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned}$$

For the first, we have $a_1 = a_4 = (\hbar^2/4)$, $a_2 = 0$; for the second, $a_1 = a_4 = (\hbar/2)$, $a_2 = 0$. The weak values are easy to write:

$$\begin{aligned} \langle \eta | \hat{S}^2 | \varphi \rangle &= \frac{\hbar^2}{4}(\bar{\eta}_1\varphi_1 + \bar{\eta}_2\varphi_2) \\ \langle \eta | \hat{S}_z | \varphi \rangle &= \frac{\hbar}{2}(\bar{\eta}_1\varphi_1 - \bar{\eta}_2\varphi_2) \end{aligned}$$

And the general formula (13) becomes:

$$\lambda^2 \langle \eta | \hat{S}^2 | \varphi \rangle = \int_{SU(2)} d\mu(g)\; Tr[U(g)\hat{S}^2]\; Tr[|\varphi\rangle\langle\eta| U^\dagger(g)]$$

$$\lambda^2 \langle \eta | \hat{S}_z | \varphi \rangle = \int_{SU(2)} d\mu(g)\; Tr[U(g)\hat{S}_z]\; Tr[|\varphi\rangle\langle\eta| U^\dagger(g)]$$

A similar formula holds for $\hat{S}_y$ and $\hat{S}_z$, but not for $\hat{S}_\pm = \hat{S}_x \pm i\hat{S}_y$, who are not hermitian.

4.2. The main formula and SO(3). Now, let G be an abstract group represented by $SO(3)$ (For instance, such a group could be the group of rotations in three-dimensionnal space). In this case, $U(g)$ will be a 3x3 matrix belonging to $SO(3)$. Generally, we can note this matrix as follows:

$$U(g) = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix} \tag{44}$$

We have written u_{ik} instead of $u_{ik}(g)$ for brevity. Recall that if $U(g) \in SO(3)$ the following relations hold:

$$det\ U(g) = 1 \tag{45}$$

$$U^T(g)U(g) = U(g)U^T(g) = I_3 \tag{46}$$

On the other hand, if g is a rotation, we can use the Euler's angles $\phi \in [0, 2\pi]$, $\theta \in [0, \pi]$, $\psi \in [0, 2\pi]$ and write:

$$U(g) = \begin{pmatrix} cos\psi & -sin\psi & 0 \\ sin\psi & cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos\theta & -sin\theta \\ 0 & sin\theta & cos\theta \end{pmatrix} \begin{pmatrix} cos\phi & -sin\phi & 0 \\ sin\phi & cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

which can be rewritten as

$$U(g) = \begin{pmatrix} cos\psi\ cos\phi - sin\psi\ cos\theta\ sin\phi & -cos\psi\ sin\phi - sin\psi\ cos\theta\ cos\phi & sin\psi\ sin\theta \\ sin\psi\ cos\phi + cos\psi\ cos\theta\ sin\phi & -sin\psi\ sin\phi + cos\psi\ cos\theta cos\phi & -cos\psi\ sin\theta \\ sin\theta\ sin\phi & sin\theta\ cos\phi & cos\theta \end{pmatrix} \tag{47}$$

This form is well known in group theory and we verify (45) and (46). Moreover, if $U(g_1)$ and $U(g_2)$ are orthogonal, have a determinant equal to 1 and represent the rotations g_1 and g_2, the product $U(g_1)U(g_2)$ is also orthogonal, has a determinant equal to 1, therefore belongs also to $SO(3)$ and is also of a form such as (47); it *represents* the rotation g_1g_2 : [2]

$$U(g_1)\ U(g_2) = U(g_1g_2)$$

We can write the state vectors as follows:

$$|\varphi\rangle := \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix} \tag{48}$$

$$|\eta\rangle := \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix} \tag{49}$$

where φ_i and $\eta_k \in \mathbb{C}$ for $i, k = 1, 2, 3$. Consequently it is natural to write:

$$|\varphi\rangle\ \langle\eta| := \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix} \otimes \begin{pmatrix} \overline{\eta}_1 & \overline{\eta}_2 & \overline{\eta}_3 \end{pmatrix} = \begin{pmatrix} \varphi_1\overline{\eta}_1 & \varphi_1\overline{\eta}_2 & \varphi_1\overline{\eta}_3 \\ \varphi_2\overline{\eta}_1 & \varphi_2\overline{\eta}_2 & \varphi_2\overline{\eta}_3 \\ \varphi_3\overline{\eta}_1 & \varphi_3\overline{\eta}_2 & \varphi_3\overline{\eta}_3 \end{pmatrix} \tag{50}$$

We have *represented* the product of the ket $|\varphi\rangle$ and the bra $\langle\eta|$ by the tensor product of two matrices. We must ensure that the kets $|\varphi\rangle$ and $|\eta\rangle$ are admissible, i.e. $|\langle\varphi|U(g)|\eta\rangle|^2$ is finite:

$$|\langle\varphi|U(g)|\varphi\rangle|^2 = |\begin{pmatrix} \overline{\varphi}_1 & \overline{\varphi}_2 & \overline{\varphi}_3 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix}|^2$$

2. Such an affirmation can be found an proved in every standard book on Group Theory, v.g. W.M. Miller, *Symmetry Groups and their Applications*, Academic Press, New York and London, 1972.

That is effectively true because all terms of the matrix product are of the form $\overline{\varphi}_i u_{ik} \varphi_l$, which are clearly finite.
As observable, we choose an hermitian 3x3 complex matrix:

$$\hat{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \tag{51}$$

with $a_{ik} = \overline{a}_{ki}$, $a_{ii} \in \mathbb{R}$. Let us write the detailed expressions that we must compute in order to verify the formula (13) in this particular case. First, on the left-hand side of (13): we use (51):

$$U(g)\hat{A} := \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \tag{52}$$

To compute the trace, we need only to add the diagonal elements of the product:

$$Tr[U(g)\hat{A}] = (u_{11}a_{11} + u_{12}a_{21} + u_{13}a_{31}) + (u_{21}a_{12} + u_{22}a_{22} + u_{23}a_{32}) + (u_{31}a_{13} + u_{32}a_{23} + u_{33}a_{33}) \tag{53}$$

We proceed in the same way with :

$$|\varphi\rangle\langle\eta|U^\dagger(g) = \begin{pmatrix} \varphi_1\overline{\eta}_1 & \varphi_1\overline{\eta}_2 & \varphi_1\overline{\eta}_3 \\ \varphi_2\overline{\eta}_1 & \varphi_2\overline{\eta}_2 & \varphi_2\overline{\eta}_3 \\ \varphi_3\overline{\eta}_1 & \varphi_3\overline{\eta}_2 & \varphi_3\overline{\eta}_3 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix}^T \tag{54}$$

Sum of diagonal elements:

$$\begin{aligned} Tr[|\varphi\rangle\langle\eta|U^\dagger(g)] = & (\varphi_1\overline{\eta}_1 u_{11} + \varphi_1\overline{\eta}_2 u_{12} + \varphi_1\overline{\eta}_3 u_{13}) + (\varphi_2\overline{\eta}_1 u_{21} + \varphi_2\overline{\eta}_2 u_{22} + \varphi_2\overline{\eta}_3 u_{23}) \\ & +(\varphi_3\overline{\eta}_1 u_{31} + \varphi_3\overline{\eta}_2 u_{32} + \varphi_3\overline{\eta}_3 u_{33}) \end{aligned} \tag{55}$$

The left-hand term of (13):

$$\lambda^2\langle\eta|\hat{A}|\varphi\rangle = \lambda^2 \begin{pmatrix} \overline{\eta}_1 & \overline{\eta}_2 & \overline{\eta}_3 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{pmatrix}$$

$$\begin{aligned} \lambda^2\langle\eta|\hat{A}|\varphi\rangle = & \overline{\eta}_1 a_{11}\varphi_1 + \overline{\eta}_2 a_{21}\varphi_1 + \overline{\eta}_3 a_{31}\varphi_1 + \overline{\eta}_1 a_{12}\varphi_2 + \overline{\eta}_2 a_{22}\varphi_2 + \overline{\eta}_3 a_{32}\varphi_2 \\ & +\overline{\eta}_1 a_{13}\varphi_3 + \overline{\eta}_2 a_{23}\varphi_3 + \overline{\eta}_3 a_{33}\varphi_3 \end{aligned} \tag{56}$$

Let us try to express the integrand on the right-hand side of (13):

$$\begin{aligned} & d\mu(g) \times Tr[U(g)\hat{A}] \times Tr[|\varphi\rangle\langle\eta|U^\dagger(g)] = \\ & d\theta\ d\phi\ d\psi\ \frac{1}{8\pi^2}\ sin\theta \\ & \times\Big((u_{11}a_{11} + u_{12}a_{21} + u_{13}a_{31}) + (u_{21}a_{12} + u_{22}a_{22} + u_{23}a_{32}) + (u_{31}a_{13} + u_{32}a_{23} + u_{33}a_{33})\Big) \\ & \times\Big((\varphi_1\overline{\eta}_1 u_{11} + \varphi_1\overline{\eta}_2 u_{12} + \varphi_1\overline{\eta}_3 u_{13}) + (\varphi_2\overline{\eta}_1 u_{21} + \varphi_2\overline{\eta}_2 u_{22} + \varphi_2\overline{\eta}_3 u_{23}) + \\ & + (\varphi_3\overline{\eta}_1 u_{31} + \varphi_3\overline{\eta}_2 u_{32} + \varphi_3\overline{\eta}_3 u_{33})\Big) \end{aligned} \tag{57}$$

Where, following (44) and (47), we must use

$$\begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix} = \begin{pmatrix} cos\psi\ cos\phi - sin\psi\ cos\theta\ sin\phi & -cos\psi\ sin\phi - sin\psi\ cos\theta\ cos\phi & sin\psi\ sin\theta \\ sin\psi\ cos\phi + cos\psi\ cos\theta\ sin\phi & -sin\psi\ sin\phi + cos\psi\ cos\theta cos\phi & -cos\psi\ sin\theta \\ sin\theta\ sin\phi & sin\theta\ cos\phi & cos\theta \end{pmatrix} \quad (58)$$

Fully developed, the expression (57) contains 81 terms to be integrated; they are of the form:

$$d\theta\ d\phi\ d\psi\ \frac{1}{8\pi^2}\ sin\theta\ u_{ij}u_{kl}\ \varphi_k\overline{\eta}_l$$

It seems to be a long and complicated task. Fortunately, it is possible, using *visually* (58) and using the tables of definite integrals presented above, to *see* easily that 72 terms vanish. The remaining 9 terms are easy to compute: all are equal to 1/3, and correspond exactly to the 9 terms of (56). As an example, we could compute:

$$\frac{1}{8\pi^2}\ sin\theta\ u_{13}^2\ a_{31}\varphi_1\overline{\eta}_3 = \frac{1}{8\pi^2}\ sin^3\theta\ sin^2\psi\ a_{31}\varphi_1\overline{\eta}_3$$

After integration:

$$\left(\frac{1}{8\pi^2} \times \frac{4}{3} \times \pi \times 2\pi\right) a_{31}\varphi_1\overline{\eta}_3 = \frac{1}{3}\ a_{31}\varphi_1\overline{\eta}_3$$

We could also compute:

$$\frac{1}{8\pi^2}\ sin\theta\ u_{31}u_{13}\ a_{13}\varphi_1\overline{\eta}_3 = \frac{1}{8\pi^2}\ sin^3\theta\ sin\phi\ sin\psi$$

After integration:

$$\left(\frac{1}{8\pi^2} \times \frac{4}{3} \times 0 \times 0\right) a_{13}\varphi_1\overline{\eta}_3 = 0$$

We may conclude:

$$\lambda^2\ \langle\eta|\hat{A}|\varphi\rangle = \int_{SO(3)} d\mu(g)\ Tr[U(g)\hat{A}]\ Tr[|\varphi\rangle\langle\eta|U^\dagger(g)] \quad (59)$$

provided $\lambda^2 = 1/3$, or

$$\lambda = \frac{1}{\sqrt{3}} = \frac{1}{\sqrt{dim\ SO(3)}} = \frac{1}{\sqrt{dim\ \mathfrak{H}}} \quad (60)$$

in agreement with (8.49) of (Ali, Antoine, and Gazeau 2000).
We have established our main formula (59) in the general case, i.e. for any hermitian operator such as $\hat{A}$ in (51). Among these operators, we can take the particular case of *projectors* $\hat{\Pi}_x$, $\hat{\Pi}_y$, $\hat{\Pi}_z$:

$$\hat{\Pi}_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (61)$$

$$\hat{\Pi}_y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (62)$$

$$\hat{\Pi}_z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{63}$$

And we can take the two normalized vectors:

$$|\eta\rangle = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad |\varphi\,\rangle = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \tag{64}$$

But we can also choose the kets $|\eta\rangle = |\psi_f\rangle$ and $|\varphi\rangle = |\psi_i\rangle$, respectively post-selected and pre-selected states, in a process of measure during which the system is slightly perturbated. In this case, we obtain the weak values we can write in this way:

$$\frac{\langle\psi_f|\hat{A}|\psi_i\rangle}{\langle\psi_f|\psi_i\rangle} = 3\int_{SO(3)} d\mu(g)\; Tr[U(g)\hat{A}]\; \frac{Tr[|\psi_i\rangle\langle\psi_f|U^\dagger(g)]}{\langle\psi_f|\psi_i\rangle} \tag{65}$$

4.3. The main formula and G/H. Let H be a normal subgroup of G, i.e. $gH = Hg$ for all $g \in G$. Let also $|\psi_i\rangle$ be an initial state such as

$$U(h)|\psi_i\rangle = |\psi_i\rangle \tag{66}$$

for all $h \in H$. Then we also have (U is unitary) $U^\dagger(h)|\psi_i\rangle = U^{-1}(h)|\psi_i\rangle = U(h^{-1})|\psi_i\rangle = |\psi_i\rangle$, because $h^{-1} \in$ H. So,

$$U^\dagger(h)|\psi_i\rangle = |\psi_i\rangle \tag{67}$$

Successively,

$$\begin{aligned} \lambda^2\langle\psi_f|A|\psi_i\rangle &= \int_G d\mu(g)\; Tr[U(g)\hat{A}]\; Tr[|\psi_i\rangle\langle\psi_f|U^\dagger(g)] \\ &= \int_G d\mu(g)\; Tr[U(g)\hat{A}]\; \sum_\alpha \langle e_\alpha|\psi_i\rangle\langle\psi_f|U^\dagger(g)|e_\alpha\rangle \\ &= \int_G d\mu(g)\; Tr[U(g)\hat{A}]\; \sum_\alpha \langle\psi_f|U^\dagger(g)|e_\alpha\rangle\langle e_\alpha|\psi_i\rangle \\ &= \int_G d\mu(g)\; Tr[U(g)\hat{A}]\; \langle\psi_f|U^\dagger(g)\sum_\alpha \Big(|e_\alpha\rangle\langle e_\alpha\Big)|\psi_i\rangle \\ &= \int_G d\mu(g)\; Tr[U(g)\hat{A}]\; \langle\psi_f|U^\dagger(g)|\psi_i\rangle \end{aligned} \tag{68}$$

We want to show now how to adapt the main formula (13) in the case where the group G is replaced by the coset G/H where H is an abelian maximal subgroup of G. Recall that every Lie algebra $\mathfrak{g}$ can be broken down into a Cartan subalgebra $\mathfrak{h}$ and another one $\mathfrak{p}$.

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p} \tag{69}$$

wich gives by exponentiation

$$G = H \otimes P \tag{70}$$

Here, $P = G/H$. Let us define $g \in G$ such as $g := hx$ with $h \in H$ and $x \in P = G/H$. One has

$$d\mu(g) = d\mu(h)\, d\mu(x) \tag{71}$$

We may rewrite :

$$\lambda^2 \langle \psi_f | \hat{A} | \psi_i \rangle = \int_H d\mu(h) \int_{G/H} d\mu(x)\ Tr[U(h)U(x)\hat{A}]\ \langle \psi_f | U^\dagger(hx) | \psi_i \rangle \tag{72}$$

$$= \int_H d\mu(h) \int_{G/H} d\mu(x)\ Tr[U(hx)\hat{A}]\ \langle \psi_f | U^\dagger(x) U^\dagger(h) | \psi_i \rangle \tag{73}$$

Now, we must note that the proof of (13) has *not* used the hermiticity of the operator $\hat{A}$. This means that (13) is true whether $\hat{A}$ is hermitian or not: $\hat{A}$ does not need to be observable. Thus the relation (68) is true for every linear operator $\hat{A}$ (however, if we want to give meaning to the notion of weak value, then $\hat{A}$ must be hermitian !). Let us define $\hat{B} = \hat{A}U(h^{-1})$. We can write:

$$\lambda^2 \langle \psi_f | \hat{B} | \psi_i \rangle = \int_H d\mu(h) \int_{G/H} d\mu(x)\ Tr[U(hx)\hat{B}]\ \langle \psi_f | U^\dagger(x) U^\dagger(h) | \psi_i \rangle$$

$$\lambda^2 \langle \psi_f | \hat{A} U(h^{-1}) | \psi_i \rangle = \int_H d\mu(h) \int_{G/H} d\mu(x)\ Tr[\hat{B}U(hx)]\ \langle \psi_f | U^\dagger(x) U^\dagger(h) | \psi_i \rangle \tag{74}$$

But (left-hand side), $U(h^{-1})|\psi_i\rangle = |\psi_i\rangle$ and (right-hand side), $U^\dagger(h)|\psi_i\rangle = |\psi_i\rangle$. Then:

$$\lambda^2 \langle \psi_f | \hat{A} | \psi_i \rangle = \int_H d\mu(h) \int_{G/H} d\mu(x)\ Tr[\hat{A}U(h^{-1})U(h)U(x)]\ \langle \psi_f | U^\dagger(x) | \psi_i \rangle \tag{75}$$

And

$$\lambda^2 \langle \psi_f | \hat{A} | \psi_i \rangle = \int_H d\mu(h) \int_{G/H} d\mu(x)\ Tr[\hat{A}U(x)]\ \langle \psi_f | U^\dagger(x) | \psi_i \rangle \tag{76}$$

We could define "*volume*" of H as the measure of H :

$$Vol(H) := \mu(H) = \int_H d\mu(h) \tag{77}$$

So, the weak value of $\hat{A}$ is given by

$$\frac{\langle \psi_f | \hat{A} | \psi_i \rangle}{\langle \psi_f | \psi_i \rangle} = \frac{Vol(H)}{\lambda^2} \int_{G/H} d\mu(x)\ Tr[\hat{A}U(x)]\ \frac{\langle \psi_f | U^\dagger(x) | \psi_i \rangle}{\langle \psi_f | \psi_i \rangle} \tag{78}$$

It is possible to give another form to this formula. Indeed we have, if the $|k\rangle$ forms a basis of $\mathfrak{H}$:

$$\begin{aligned}
\langle \psi_f | U^\dagger(x) | \psi_i \rangle &= \langle \psi_f | U^\dagger(x) \Big(\sum_k |k\rangle\langle k| \Big) |\psi_i\rangle \\
&= \sum_k \langle \psi_f | U^\dagger(x) | k \rangle \langle k | \psi_i \rangle \\
&= \sum_k \langle k | \psi_i \rangle \langle \psi_f | U^\dagger(x) | k \rangle \\
&= Tr[|\psi_i\rangle\langle \psi_f | U^\dagger(x)]
\end{aligned} \tag{79}$$

Formula (78) becomes :

$$\frac{\langle\psi_f|\hat{A}|\psi_i\rangle}{\langle\psi_f|\psi_i\rangle} = \frac{Vol(H)}{\lambda^2}\int_{G/H} d\mu(x)\; Tr[\hat{A}U(x)]\;\frac{Tr[|\psi_i\rangle\langle\psi_f|U^\dagger(x)]}{\langle\psi_f|\psi_i\rangle} \tag{80}$$

And the right-hand side of (80) contains only traces.

An interesting example is given by:

$$G = SO(3) \qquad H = SO(2) \qquad P = \frac{SO(3)}{SO(2)}$$

where $SO(3)/SO(2)$ is nothing but the sphere S_2.

5. Conclusion.

The main formula (13) and his extension to coset (72) are probably not entirely new (if we consider the theory of generalized coherent stated à la Ali-Antoine-Gazeau), but the formulae we have proved are useful in the quantum weak value context. Furthermore, relation (72) could be interesting if we want to consider quantum theory starting from the phase space, which is taken as homogeneous Kählerian manifolds. In this theory, formula (68) enables us to express the transition amplitudes. The generalization of formulae (13) and (68 - 76) in the cases of noncompact groups is not obvious because we need some square-integrable irreducible representation. But in certain cases it is possible. In particular in situations where the coset spaces are the so-called classical domains (deeply studied by Jean-Pierre Gazeau, (Gazeau 1989): $SO(4,2)/SO(4)\otimes SO(2)$, $SO(3,2)/SO(3)\otimes SO(2)$, $SO(2,2)/SO(2)\otimes SO(2)$, $SO(1,2)/SO(2)$. All these manifolds are kählerian manifolds. The wave functions are to be chosen as elements of Hilbert spaces of analytical functions of such domains and square integrable with an appropriate (Bergman) measure.

6. A general perspective on this work

The work on weak measures is done in the context of a joint Belgian research project (ARC, "Action de Recherche Concertée") gathering physicists (Y. Caudano, L.Ballestros, J.-P. Fréché) mathematicians (T. Carletti, W. Delongha), logicians and philosophers (B. Hespel, V. Degauquier) of the University of Namur (Naxys and Esphin Research Institutes), dedicated to the interpretations of weak values and measurements. The weak value of an observable of a system is obtained during a very weak interaction (implying as little perturbation as possible) with the system constrained by imposing a pre-selected state and a post-selected state. The weak values have strange behaviors (they can be complex and go sometimes outside the usual spectrum) and in some cases they are related to values predicted by Bohm theory (thus they are interesting when you are interested in the study of interpretations of Quantum Theory). Weak values and measurements addressed many philosophical questions because. For example (as it was considered in a thought experiment imagined by Wheeler of a modified two-slit experiment) if you are fixing photons in a preselected state and after a long time (when photons are already flying away the source!), you chose various post-selected states, all these choices change radically the way you are describing the past. How can we interpret this fact? It is difficult to admit that present is influencing the past. Is it maybe rather convenient to think that modification of present knowledge of the future can modify the way we are understanding the past. But there is here the place for a debate on the structure of time (see Thomas Hertogh, On the origin of time. Stephen Hawking's final theory, Penguin Books, 2023). Our personal work is also connected to applications of weak measurements in cosmology (see the seminal works of Brout, Englert and Spindel for example). Post-selected states correspond here to prescribed final state of the universe. Here the knowledge of a final state could influence the way one is telling the origin of the universe… This addressing many interesting philosophical questions. Technically, we have to consider here quantum theory in curved space-time. Without entering completely into this tough subject, we have modestly begun to tackle the problem of defining weak values in the context of a curved space-time. We have chosen to begin with the Wigner phase-space formalism of Quantum Theory (based on functions: Wigner and cross-Wigner transforms, Weyl symbols of operators, …) and adapt it to the case of curved phase-space (being non trivial Kähler manifolds). The paper presented here is to be considered in such a context and try to search a way to express a weak value in the case of phase-space endowing with Lie group symmetry. This could serve as toy-model to explore weak values in the curved phase-space context (see the important work of Maurice de Gosson, The Wigner Transform, World Scientific, 2017). Mathematically our result can also be considered in the field of harmonic analysis on Lie groups and as a way to perform the Lie group harmonic analysis of a weak value.

Acknowledgement

The authors thank Prof. Yves Caudano for his kind help and for having introduced us in the domain of weak measurements in the context of an ARC project. One of the authors (D. L.) thanks Prof. Jean-Pierre Antoine for his generosity in introducing him in the generalized coherent states study. This research was supported by the Action de Recherche Concertée WeaM at the University of Namur (19/23-001)

References

Abgaryan, V., A. Khvedelidze, and A. Torosyan. 2019. On the moduli space of Wigner quasiprobability distributions for N-dimensional quantum systems. *J. Math. Sci.* 240 (5): 617–633. ISSN: 1573-8795. https://doi.org/10.1007/s10958-019-04379-7.

Ali, Syed Twareque, Jean-Pierre Antoine, and Jean-Pierre Gazeau. 2000. *Coherent states, wavelets and their generalizations.* Edited by R. Stephen Berry, Joseph L. Birman, Jeffrey W. Lynn, Mark P. Silverman, H. Eugene Stanley, and Mikhail Voloshin. Graduate Texts in Contemporary Physics. New York, NY: Springer. ISBN: 978-1-4612-7065-2. https://doi.org/10.1007/978-1-4612-1258-4.

Antonsen, Frank. 1998. Wigner–Weyl–Moyal formalism on algebraic structures. *Int. J. Theor. Phys.* 37 (2): 697–757. ISSN: 1572-9575. https://doi.org/10.1023/A:1026612428446.

Brif, C., and A. Mann. 1998. A general theory of phase-space quasiprobability distributions. *J. Phys. A: Math. Gen.* 31 (1): L9. ISSN: 0305-4470. https://doi.org/10.1088/0305-4470/31/1/002.

de Gosson, Maurice. 2017. *The Wigner transform* |. Advanced Textbooks in Mathematics. New York, NY: Wiley. https://doi.org/10.1142/q0089.

Faraut, Jacques. 2006. *Analysis on Lie groups: an introduction.* Paris: Calvage et Mounet.

Gadella, Manuel, Miguel A. Martin, Luis M. Nieto, and Mariano A. Del Olmo. 1991. The Stratonovich–Weyl correspondence for one-dimensional kinematical groups. *J. Math. Phys.* 32 (5): 1182–1192. ISSN: 0022-2488, 1089-7658. https://doi.org/10.1063/1.529315.

Gazeau, Jean-Pierre. 1989. *Lecture notes in complex geometry and elementary quantum systems.* Louvain-la-Neuve, Belgium.

Rundle, Russell P., and Mark J. Everitt. 2021. Overview of the Phase Space Formulation of Quantum Mechanics with Application to Quantum Technologies. *Adv. Quantum Technol.* 4 (6): 2100016. ISSN: 2511-9044. https://doi.org/10.1002/qute.202100016.

Varilly, Joseph C. 1989. *The Stratonovich-Weyl correspondence: A general approach to Wigner functions.* BIBOS preprint # 345, U. Bielefeld, Germany.

Varilly, Joseph C, and M Gracia-Bonda. 1989. The Moyal representation for spin. *Ann. Phys.* 190:107–148. https://doi.org/10.1016/0003-4916(89)90262-5.

5. Logemes and Their Homotopy-Theoretic Foundations

Andrew Schumann

Kielnarowa Review (2025), **2**, 57–66
doi:https://doi.org/10.57656/kr-2025-0005

Kielnarowa
REVIEW

ARTICLE

Logemes and Their Homotopy-Theoretic Foundations

Andrew Schumann*

University of Information Technology and Management, Chair of Cognitive Science and Mathematical Modelling, ul. Sucharskiego 2, 35-22 Rzeszów, Poland
*Corresponding author. Email: andrew.schumann@gmail.com

Abstract

We introduce *logemes*, consistent fragments of reasoning closed under at least one inference rule, as foundational units for different logics. Unlike full logical theories, logemes need not be axiomatized or algebraically structured; instead, they are evaluated via their associated Lindenbaum–Tarski quotients, interpreted as spaces presenting partially ordered sets. We propose that two logemes are homotopy identical precisely when their posetal semantics are homotopy equivalent. This criterion, grounded in the Univalence Axiom of Homotopy Type Theory, allows us to formalize diagrammatic reasoning and compare ancient logical traditions, such as Stoic and Yogācāra, on purely mathematical grounds. We show that both traditions instantiate the same homotopy type of poset, confirming their logical identity despite historical separation.

Keywords: simplicial complex, homotopy type, univalent foundations

1. Introduction

From the standpoint of modern mathematical logic, pre-20th-century logical systems, such as Stoic, Epicurean, Aristotelian, and Buddhist, are not theories in the formal sense, since they lack axiomatizations, completeness results, and algebraic semantics. Nevertheless, they contain coherent doctrines of inference, introducing explicit rules for deriving conclusions from premises.

We propose that such doctrines are best understood as collections of *logemes* – minimal, usable fragments of reasoning. A logeme is not a theory but a *reasoning token*, i.e., a finite, consistent set of formulas on which at least one inference rule can be applied (e.g., *modus ponens* on $\{p, p \Rightarrow q, q\}$). This reflects everyday practice: legal, medical, or scientific argumentation rarely invokes full logical systems, but relies instead on reusable, local inferential patterns. Hence, the central thesis of this paper is as follows: *the history of logic prior to Frege and Russell is the history of logemes, not of logical theories.*

To formalize this approach, we equip logemes with semantic content via Lindenbaum–Tarski quotients, then classify them up to *homotopy equivalence* of their induced posets. This approach, inspired by univalent foundations Univalent Foundations Program 2013, treats logically equivalent reasoning patterns as *homotopy identical* when their diagrammatic realizations share the same homotopy type.

The paper proceeds as follows. In section 2, we define logemes and their semantics. Section 3 surveys tools for computing homotopy types of posets. Section 4 establishes the homotopy-theoretic identification principle. In section 5, we apply the framework to Stoic and Yogācāra logemes, proving their identity. We conclude with historical and methodological implications.

2. Logemes and Their Semantics

Let T be a (classical or intuitionistic) propositional logic, with algebra of formulas A_T. For any formula p, define its *equivalence class*:

$$\|p\|_T := \{q\colon \vdash_T p \equiv q\}.$$

Let $\equiv_T$ be the congruence induced by $\equiv$. The *Lindenbaum-Tarski algebra* of T is the quotient $A_T/{\equiv_T}$. When T is classical, this is a Boolean algebra; for intuitionistic T, a Heyting algebra.

Definition 1 (Logeme). *A finite set of formulas $F \subseteq A_T$ is a* logeme *iff F is* consistent*: $\nvdash_T \bigwedge F \Rightarrow \bot$. Furthermore, this F is not trivial iff it is inferentially closed under at least one rule: there exists an inference rule R of T and a substitution σ such that the premises of $R\sigma$ lie in F, and its conclusion is entailed by F (not necessarily in F itself).*

The set $F = \{p, p \Rightarrow q\}$ is a logeme: it is consistent (in any nontrivial T), and *modus ponens* applies to yield q. The singleton $\{p\}$ is trivial, as no rule applies; and $\{p, \neg p\}$ is inconsistent, hence excluded.

Given a logeme F, we consider the subalgebra $A_F \subseteq A_T$ generated by F, and its quotient $A_F/{\equiv_T}$. However, for historical logics, we cannot assume a background theory T. We thus define semantics intrinsically:

Definition 2 (Meaningful Logeme). *A logeme F is* meaningful *if the Lindenbaum quotient $A_F/{\equiv}$ (for the fragment it generates) carries a partial order $\leq$ such that:*

$$\|p \Rightarrow q\| = 1 \quad \textit{iff} \quad \|p\| \leq \|q\|.$$

In this case, $A_F/{\equiv}$ is a poset*; it is a* lattice *if all binary suprema and infima exist.*

Proposition 1. *The logeme $F = \{(p \,\&\, (p \Rightarrow q)) \Rightarrow q\}$ is meaningful for the 2-element chain $\mathbf{2} = \{0 < 1\}$ under Boolean interpretation.*

Proof. Indeed:

$$\begin{aligned}\|(p \,\&\, (p \Rightarrow q)) \Rightarrow q\| &= \|\neg(p \,\&\, (\neg p \vee q)) \vee q\| \\ &= \|(\neg p \vee \neg(\neg p \vee q)) \vee q\| \\ &= \|((\neg p \vee q) \vee \neg(\neg p \vee q))\| = 1.\end{aligned}$$

Hence, the formula evaluates to 1 (top element), and *modus ponens* holds in $\mathbf{2}$. □

Thus, logical reasoning reduces to operating on *individual meaningful logemes*, each tied to a class of posets. The space of all such logemes vastly exceeds the space of formal logical systems.

3. Computational Tools for Homotopy Types of Posets

To classify logemes via homotopy, we require effective methods to compute or simplify the homotopy type of the order complex $|\Delta(P)|$ of a finite poset P. Below we recall and elaborate the main combinatorial-topological tools.

Definition 3 (Order Complex). *Let P be a (finite) partially ordered set. The* order complex *of P, denoted $\Delta(P)$, is the abstract simplicial complex defined by:*

$$\Delta(P) := \Big\{\{x_0, x_1, \ldots, x_k\} \subseteq P : \ k \geq 0,\ x_0 < x_1 < \cdots < x_k \text{ in } P\Big\}.$$

Each such chain is a k-simplex; the face relation is given by inclusion of subchains.

The geometric realization $|\Delta(P)|$ is the topological space obtained by assigning to each k-simplex $\{x_0 < \cdots < x_k\}$ a standard geometric k-simplex in $\mathbb{R}^{k+1}$ and gluing them along common faces. This yields a finite CW-complex, canonically associated to P.

Remark 1. *The construction is functorial: an order-preserving map $f: P \to Q$ induces a simplicial map $\Delta(f): \Delta(P) \to \Delta(Q)$, and hence a continuous map $|\Delta(f)|: |\Delta(P)| \to |\Delta(Q)|$.*

The homotopy type of $|\Delta(P)|$ is an invariant of P; it captures global connectivity properties (e.g., presence of "holes") while ignoring local combinatorial details.

Example 1. *Let us consider some easy cases:*

1. *If $P = \{a\}$ is a singleton, $\Delta(P) = \{\{a\}\}$ and $|\Delta(P)|$ is a point.*
2. *If $P = \{a, b\}$ with $a \parallel b$ (incomparable), then $\Delta(P) = \{\{a\}, \{b\}\}$ (no 1-simplex), so $|\Delta(P)|$ is a discrete 2-point space, homotopy equivalent to S^0.*
3. *If $P = \{a < b\}$, then $\Delta(P) = \{\{a\}, \{b\}, \{a, b\}\}$, and $|\Delta(P)|$ is a closed interval and then it is contractible.*

A powerful simplification technique for finite posets is *core reduction*, introduced by Stong 1966 and later refined in the homotopy-theoretic context by Raptis 2010.

Definition 4 (Upbeat and Downbeat Points)**.** *Let P be a finite poset and $x \in P$.*

- *x is an* upbeat point *if the set $U(x) := \{y \in P : y > x\}$ is nonempty and has a* least element*, denoted $x^{\uparrow}$.*
- *x is a* downbeat point *if the set $L(x) := \{y \in P : y < x\}$ is nonempty and has a* greatest element*, denoted $x^{\downarrow}$.*

Equivalently, x is upbeat iff there exists $x^{\uparrow} > x$ such that for all $y > x$, $y \geq x^{\uparrow}$; similarly for downbeat.

Intuitively, an upbeat point x is "redundant": it sits directly below a unique minimal element above it, so removing x and identifying it with $x^{\uparrow}$ does not change the homotopy type.

Theorem 1 (Core Reduction: Stong 1966; Raptis 2010)**.** *Let P be a finite T_0-poset (i.e. the specialization preorder is antisymmetric – it is automatically true for posets). Then*

1. *There exists a unique subposet* $\mathrm{core}(P) \subseteq P$*, called the* core *of P, such that:*
   - $\mathrm{core}(P)$ *contains no upbeat or downbeat points;*
   - $\mathrm{core}(P)$ *is a* strong deformation retract *of P in the Alexandrov topology (hence $|\Delta(P)| \simeq |\Delta(\mathrm{core}(P))|$).*
2. $\mathrm{core}(P)$ *is obtained recursively: repeatedly delete any upbeat or downbeat point until none remain; the result is independent of deletion order.*
3. *For finite posets P and Q, we have*

$$|\Delta(P)| \simeq |\Delta(Q)| \iff \mathrm{core}(P) \cong \mathrm{core}(Q) \textit{ as posets.}$$

Example 2. *Consider the following 4-element poset:*

$$P = \{a, b, c, d\}, \quad a < c,\ b < d,\ a \parallel b,\ c \parallel d,\ a \parallel d,\ b \parallel c.$$

No element has a unique cover above or below, e.g., $U(a) = \{c\}$ has least element c, but $U(a) = \{c\}$ is a singleton, so c is its minimum. Thus a is upbeat, with $a^{\uparrow} = c$. Similarly b is upbeat ($b^{\uparrow} = d$), c is downbeat ($c^{\downarrow} = a$), d is downbeat ($d^{\downarrow} = b$). But core reduction forbids removing both endpoints of a covering pair simultaneously. In fact, the core is obtained by removing one pair, e.g., delete a and b; the remaining poset $\{c, d\}$ with $c \parallel d$ is discrete 2-point. Thus, its order complex is S^0.

Example 3. *Consider another 4-element poset:*

$$P = \{0, a, b, 1\}, \quad 0 < a < 1,\ 0 < b < 1,\ a \parallel b.$$

This is the Boolean lattice $\mathbf{2}^2$. Now:

- *a: $U(a) = \{1\}$, then it is upbeat ($a^{\uparrow} = 1$);*
- *b: $U(b) = \{1\}$, then it is upbeat;*
- *1: $L(1) = \{a, b\}$ has no greatest element (since $a \parallel b$); therefore, it is not downbeat;*
- *0: $U(0) = \{a, b\}$ has no least element; therefore, it is not upbeat.*

Removing a and b leaves $\{0, 1\}$ with $0 < 1$, but it is a contractible chain.

Example 4. *A new 4-element poset:*

$$P = \{-1, a, b, +1\}, \quad -1 < a < +1,\ -1 < b < +1,\ a \parallel b.$$

Let its order complex be as follows:

$$\Delta(P) = \{\{-1\}, \{a\}, \{b\}, \{+1\}, \{-1, a\}, \{-1, b\}, \{a, +1\}, \{b, +1\}\}.$$

Then, geometrically, this is a 1-dimensional simplicial complex consisting of two edges sharing endpoints, but it is a circle S^1. Indeed,

$$|\Delta(P)| \cong S^1.$$

Beyond core reduction, several other powerful techniques exist:

1. **Cross-cuts** (Lakser 1971). Let L be a bounded lattice ($0 = \min L$, $1 = \max L$). A *cross-cut* is a subset $X \subseteq L \setminus \{0, 1\}$ such that:
   (i) X is an antichain (no two elements comparable);
   (ii) Every maximal chain in L contains exactly one element of X.

   Then the inclusion-induced map $|\Delta(X)| \hookrightarrow |\Delta(L)|$ is a homotopy equivalence. In particular, if X is discrete with k elements, $|\Delta(L)| \simeq \bigvee^{k-1} S^0$.
2. **Contractible Carriers** (Walker 1981). Let K be a simplicial complex and Y a topological space. A *carrier* is a function C: {simplices of K} $\to$ {subspaces of Y} such that $\sigma \subseteq \tau \Rightarrow C(\sigma) \subseteq C(\tau)$. It is *contractible* if each $C(\sigma)$ is contractible. Then:
   (a) There exists a continuous map $f : |K| \to Y$ with $f(|\sigma|) \subseteq C(\sigma)$ for all σ (*f is carried by C*);
   (b) Any two such maps are homotopic.

   This is used to construct homotopy equivalences combinatorially.
3. **Order Homology**. For $n \geq 0$, define the *order homology groups* of P by:

   $$H_n(P; \mathbb{Z}) := H_n^{\text{sing}}(|\Delta(P)|; \mathbb{Z}),$$

   the singular homology of the geometric realization. Equivalently, one may use simplicial homology of $\Delta(P)$. These are homotopy invariants:

   $$P \simeq Q \Rightarrow H_n(P) \cong H_n(Q)\ \forall n.$$

   For instance:

   $$H_0(P) \cong \mathbb{Z}^c, \quad c \text{ is a number of connected components of } |\Delta(P)|,$$

   and $H_1(P) \cong \mathbb{Z}^r$ detects the number r of independent 1-dimensional loops.

These tools enable a systematic classification of logeme semantics. Given a logeme F, one proceeds step by step as follows:

1. Compute its associated poset presentation $P_F := A_F/\equiv$.

2. Reduce P_F to its core, core(P_F), by iteratively stripping away all dismantlable or contractible elements. This yields a canonical representative in the homotopy class of P_F.
3. Compute the (co)homology groups $H_*(P_F)$, or alternatively recognize P_F (or core(P_F)) as a known combinatorial poset or simplicial shape (e.g. a diamond $\mathbf{2}^2$, a Boolean lattice $\mathbf{2}^n$, a sphere, or a bouquet of circles).
4. From these invariants determine the homotopy type of the semantic space carried by F.

Two logemes F and G are identified precisely when their associated invariants coincide; that is,

$$\mathrm{core}(P_F) \simeq \mathrm{core}(P_G) \quad \text{and} \quad H_*(P_F) \cong H_*(P_G),$$

ensuring that they define the same semantic homotopy type. This procedure yields a complete and structurally transparent classification of logemes.

4. Homotopy Identification of Logemes

To systematize logemes in a manner compatible with both historical reasoning and modern mathematics, we work within the framework of *univalent foundations* (Univalent Foundations Program 2013). At its core lies the *Univalence Axiom*, which reinterprets equality of mathematical objects as equivalence:

Axiom 1 (Univalence Axiom). *For any two types (spaces) X and Y, the identity type $X = Y$ is equivalent to the type of homotopy equivalences between them:*

$$(X = Y) \simeq (X \simeq Y).$$

Let F be a meaningful logeme (cf. definition 2). Its Lindenbaum quotient $A_F/{\equiv}$ carries a natural partial order $\leq$, making it a finite poset, denoted

$$P_F := A_F/{\equiv}.$$

As explained in section 3, every finite poset gives rise to a *simplicial complex* $\Delta(P_F)$, or its *order complex*, whose geometric realization $|\Delta(P_F)|$ is a finite CW-complex. This space encodes the global "shape" of the logeme: connected components correspond to independent reasoning fragments, loops to cyclic dependencies, and higher holes to unresolvable contradictions.

We thus treat P_F not merely as a combinatorial object, but as a space via the composite functor

$$F \longmapsto P_F \longmapsto \Delta(P_F) \longmapsto |\Delta(P_F)|.$$

The Univalence Axiom motivates two related but distinct relations on logemes:

Definition 5 (Strict Identity of Logemes). *Two logemes F and F' are* strictly identical*, written $F = F'$, iff their order complexes are the same:*

$$F = F' \iff \Delta(P_F) = \Delta(P_{F'}).$$

Strict identity means that combinatorically both logemes are identical. More useful is the univalent notion:

Definition 6 (Homotopy Identity of Logemes). *Two logemes F and F' are* homotopy identical *(in the univalent sense), written $F \simeq F'$, iff their realizations are homotopy equivalent:*

$$F \simeq F' \iff |\Delta(P_F)| \simeq |\Delta(P_{F'})|.$$

Note that under the Univalence Axiom (axiom 1), this is not merely a definition, since it is a justified identification: the identity type $F \simeq F'$ is the type of homotopy equivalences between $|\Delta(P_F)|$ and $|\Delta(P_{F'})|$.

The notation $|\Delta(P_F)|$ denotes the geometric realization of the order complex of the poset P_F. In the present framework, this space admits a natural interpretation as a *logic diagram* for the formulas generated by the logeme F.

More explicitly, the simplices of $\Delta(P_F)$ correspond to chains of entailments or dependency relations among the formulas encoded in F, while their geometric realization $|\Delta(P_F)|$ provides a continuous topological space in which these logical relations are represented as higher-dimensional cells. Thus, $|\Delta(P_F)|$ serves as a faithful topological model of the inferential structure of F, capturing not only the individual formulas but also the ways in which they combine, branch, or form cycles within the logical system. It provides a rigorous mathematical framework for *logic diagrams* as a bona fide logical discipline in its own right, distinct from syntactic proof theory or model-theoretic semantics (Anger et al. 2022, Schumann and Lemanski 2022) – by formalizing their structural invariants through homotopy-theoretic and order-theoretic methods. Specifically, diagrams are no longer treated as heuristic illustrations, but as *combinatorial models* whose inferential content is encoded in the homotopy type of their associated posets (or, equivalently, the geometric realization of their order complexes). This approach validates the five "dogmas" challenged in Anger et al. 2022, e.g., the belief that diagrams are inherently ambiguous, non-compositional, or incapable of expressing generality, by demonstrating that:

(i) *Syntax* can be captured via formulas from F;
(ii) *Semantics* arises from the simplicial complex $\Delta(P_F)$ for the poset P_F obtained based on the formulas of F;
(iii) *Inference rules* correspond to homotopy-preserving transformations (e.g., core reduction, barycentric subdivision, or strong collapses);
(iv) *Soundness and completeness* can be formulated as homotopy equivalences between diagram spaces and algebraic models (e.g., Boolean or Heyting algebras);
(v) *Diagrammatic equivalence* coincides with univalent identity: two diagrams represent the same reasoning pattern iff their realizations are homotopy equivalent.

Thus, the study of logical diagrams is elevated from a pedagogical aid to a formal branch of logic, what one may call *homotopy theory of logic diagrams* with its own syntax, semantics, and proof theory grounded in univalent foundations. As we can see, univalent foundations provide not only a formal framework, but a *conceptual clarification*, for which logical reasoning is not about manipulating symbols in a fixed syntax, but about navigating a space of homotopy types, where identity is sameness of shape, and inference is continuous transformation.

5. Case Study: Stoic and Yogācāra Logemes

We now apply the univalent framework to two historically and geographically distinct traditions: Hellenistic Stoicism (3rd c. BCE) and Indian Yogācāra Buddhism (5th c. CE). Despite a millennium of separation and no known direct textual transmission, their core inferential systems exhibit striking structural parallels. We show that this is homotopy equivalence.

The Stoics codified five *anapodeiktoi sullogismoi* (indemonstrable syllogisms), see Bobzien 1999, captured by the logeme:

$$F_{\text{Stoic}} = \left\{ \begin{array}{ll} (\text{MP}) & ((p \Rightarrow q) \,\&\, p) \Rightarrow q, \\ (\text{MT}) & ((p \Rightarrow q) \,\&\, \neg q) \Rightarrow \neg p, \\ (\text{MPT}) & (\neg(p \,\&\, \neg q) \,\&\, p) \Rightarrow \neg\neg q, \\ (\text{MTP}_1) & ((\neg p \oplus q) \,\&\, \neg p) \Rightarrow \neg q, \\ (\text{MTP}_2) & ((\neg p \oplus q) \,\&\, \neg\neg p) \Rightarrow q. \end{array} \right\}. \tag{1}$$

These correspond respectively to *modus ponens* (MP), *modus tollens* (MT), *modus ponendo tollens* (MPT), and two variants of *modus tollendo ponens* (MTP_1 and MTP_2).

Their semantics is given by the square of opposition in Figure 1, interpreted as a poset P_{Stoic}, in which $\|p\|$:= "day", $\|q\|$:= "light", $\|\neg p\|$:= "dark", $\|\neg q\|$:= "night".

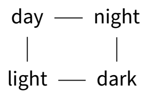


Figure 1. The Stoic square of opposition. Bottom edge is *contradictory* (not subcontrary).

In this poset P_{Stoic}, we know that $\|q \oplus \neg p\| = 1$. Then

$$\Delta(P_{\text{Stoic}}) = [\{\|p\|, \|q\|, \|q \oplus \neg p\|\}, \{\|\neg q\|, \|\neg p\|, \|q \oplus \neg p\|\}],$$

i.e., $\Delta(P_{\text{Stoic}})$ consists of two 2-simplices, see Figure 2. It is worth noting that $|\Delta(P_{\text{Stoic}})|$ is contractible (collapsed to a point).

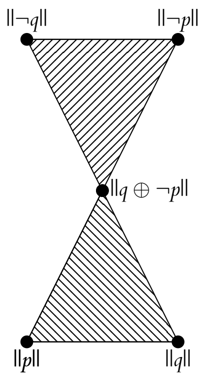


Figure 2. The logic diagram $|\Delta(P_{\text{Stoic}})|$ for the Stoic square of opposition, see Figure 1.

Dharmakīrtī's system (Dreyfus 1997), systematized in the *Nyāyabindu*, employs 13 context-sensitive inference rules, categorized by token type (analytic, causal, genus, etc.) and compositionally extended (e.g., *composite modus tollendo ponens*). So, the Yogācāra logeme F_{Yogacara} is defined as follows:

(i) (*MP*, analytic token) $((p \Rightarrow q) \& p) \Rightarrow q$,
(ii) (*MP*, causal token) $((p \Rightarrow q) \& p) \Rightarrow q$,
(iii) (*MT*, analytic token) $((p \Rightarrow q) \& \neg q) \Rightarrow \neg p$,
(iv) (*MT*, causal token) $((p \Rightarrow q) \& \neg q) \Rightarrow \neg p$,
(v) (*MT*, genus) $((p \Rightarrow q) \& \neg q) \Rightarrow \neg p$,

(vi) (composite MTP_1) $((p \Rightarrow q)\&(q \oplus \neg p)\&p) \Rightarrow \neg\neg p$,
(vii) (composite MTP_1) $((q \oplus \neg p)\&p\&((p \Rightarrow r)\&(r \Rightarrow q))) \Rightarrow \neg\neg p$,
(viii) (composite MPT) $(\neg(q\&\neg p)\&p\&(p \Rightarrow q)) \Rightarrow \neg\neg p$,
(ix) (MPT, cause) $(\neg(q\&\neg p)\&q) \Rightarrow \neg\neg p$,
(x) (composite MTP_1) $((q \oplus \neg p)\&p\&(p \Rightarrow q)) \Rightarrow \neg\neg p$,
(xi) (MT, cause) $((p \Rightarrow q)\&\neg q) \Rightarrow \neg p$,
(xii) (composite MTP_2) $((q \oplus \neg p)\&p\&(p \Rightarrow \neg q)) \Rightarrow \neg p$,
(xiii) (MPT, token) $(\neg(q\&\neg p)\&q) \Rightarrow \neg\neg p$.

These rules are contextually enriched variants of the Stoic indemonstrables: *modus ponens* (MP), *modus tollens* (MT), *modus ponendo tollens* (MPT), and two versions *modus tollendo ponens* (MTP_1 and MTP_2).

Both logemes (Stoic and Yogācāra) are verified on the same poset. Indeed, the Yogācāra semantics is given by the square of opposition in Figure 3, interpreted as a poset P_{Yogacara}, in which $\|p\|$:= "fire", $\|q\|$:= "warmth", $\|\neg p\|$:= "cold", $\|\neg q\|$:= "ice".

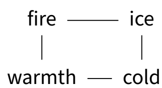


Figure 3. The Yogācāra square of opposition. Bottom edge is also *contradictory* (not subcontrary).

Yogācāra logic, like its Stoic counterpart, is grounded in a theory of *sign-inference* (*anumāna*). A phenomenon q (the *signified*, *sādhya*) is inferred from its *sign* or *token* (*liṅga*, p), provided p is invariably connected with q (the *vyāpti* condition).

Proposition 2. *Under the above interpretation,* $P_{\text{Yogacara}} \cong P_{\text{Stoic}}$ *as posets. Consequently:*

$$\Delta(P_{\text{Yogacara}}) \cong \Delta(P_{\text{Stoic}}).$$

□

Hence, we conclude:

Theorem 2 (Stoic–Yogācāra Logeme Identity). *The Stoic logeme* F_{Stoic} *and the Yogācāra logeme* F_{Yogacara} *are identical:*

$$F_{\text{Stoic}} = F_{\text{Yogacara}}.$$

□

In contrast, the *Aristotelian* syllogistic logeme $F_{\mathsf{Aristotelian}}$, modeled on the classical square of opposition (with subcontrary bottom) is not homotopy equivalent to Stoic or Yogācāra logemes: $F_{\mathsf{Aristotelian}} \not\simeq F_{\mathsf{Stoic}}$.

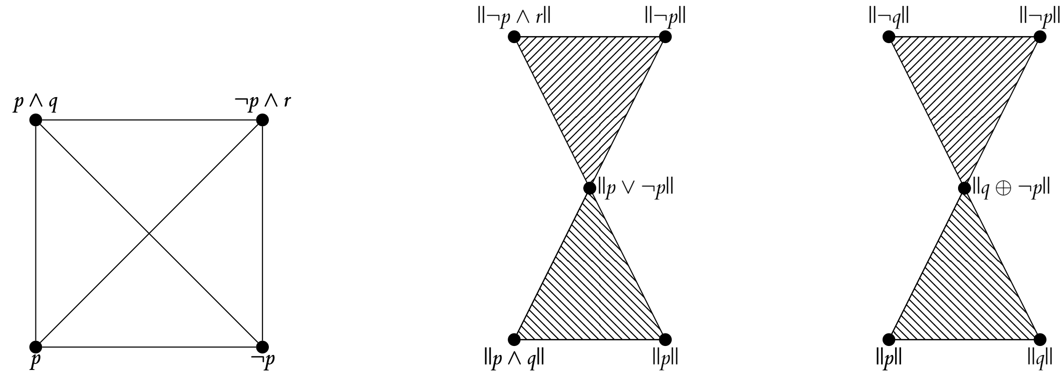
Historically, the earliest attested logemes appear in Mesopotamian legal and divinatory texts (2nd millennium BCE), explicitly deploying MP and MT (Schumann 2023). No such reasoning fragments are found in contemporaneous cultures lacking Sumerian–Akkadian influence. The homotopy identity $F_{\text{Stoic}} = F_{\text{Yogacara}}$ thus supports a diffusionist hypothesis: Hellenistic logical doctrine likely influenced early Indian Buddhist epistemology along Silk Road intellectual networks. This case exemplifies how univalent foundations can be applied to two historically distinct traditions.

Remarkably, the same homotopy type arises even within some fragments of *classical propositional logic* (CPL):

Example 5 (Classical Realization). *Consider the* CPL-*logeme:*

$$F_{\mathsf{CPL}} = \Big\{(p \wedge q) \Rightarrow p,\ (\neg p \wedge r) \Rightarrow \neg p,\ p \oplus \neg p,\ \neg\big((p \wedge q) \wedge (\neg p \wedge r)\big)\Big\}.$$

Its semantic poset P_{CPL} is isomorphic to P_{Stoic}, and $\Delta(P_{\mathsf{CPL}}) \cong \Delta(P_{\mathsf{Stoic}}) \cong \Delta(P_{\mathsf{Yogacara}})$, see Figure 4. Thus, the Stoic, Yogācāra, and even a fragment of classical logic converge on the same order complex – *a two 2-simplices connected at one point.*



(a) The Stoic square of opposition for F_{CPL}

(b) $|\Delta(P_{\mathsf{CPL}})|$

(c) $|\Delta(P_{\mathsf{Stoic}})| \cong |\Delta(P_{\mathsf{Yogacara}})|$

Figure 4. (a) The Stoic square of opposition for the logeme F_{CPL}; (b) the simplicial complex of the Stoic square of opposition for the logeme F_{CPL}; (c) the simplicial complex of the Stoic square of opposition for the logeme F_{Stoic} of the Stoics as well as for the logeme F_{Yogacara} of the Yogācārins.

6. Conclusion

We have formalized *logemes* as the elementary units of cross-cultural logical practice and provided a homotopy-theoretic criterion for their identity. This framework:

- Explains "logical monism" in history (Lemanski 2025) by using different combinations of logemes;
- Mathematically validates claims about "common cores" in diagrammatic reasoning (Anger et al. 2022);
- Explains the structural coincidence of Stoic and Yogācāra inference systems as homotopy invariance;
- Grounds the "logic of diagrams" in pure mathematics, due to univalence.

The earliest attested logemes (*modus ponens* and *modus tollens* in Mesopotamian divination texts) suggest a diffusionist hypothesis: the shared homotopy type of Hellenistic and Buddhist logemes likely reflects historical transmission along the Silk Road, not independent convergence.

References

Anger, C., Th. Berwe, A. Olszok, A. Reichenberger, and J. Lemanski. 2022. Five dogmas of logic diagrams and how to escape them. *Language & Communication* 87:258–270. https://doi.org/10.1016/j.langcom.2022.09.001.

Bobzien, S. 1999. Logic: the stoics (part one). In *The cambridge history of hellenistic philosophy,* edited by K. Algra, J. Barnes, J. Mansfeld, and M. Schofield, 92–125. Cambridge: Cambridge University Press.

Dreyfus, G. 1997. *Recognizing reality: dharmakırti's philosophy and its tibetan interpretations.* Albany: State University of New York Press.

Lakser, H. 1971. The homology of a lattice. *Discrete Mathematics* 1 (2): 187–192. https://doi.org/10.1016/0012-365X(71)90024-0.

Lemanski, J. 2025. Does logic have a history at all? Published online 14 November 2023, *Foundations of Science* 30:227–249. https://doi.org/10.1007/s10699-023-09933-w.

Raptis, G. E. 2010. Homotopy theory of posets. *Homology, Homotopy and Applications* 12 (2): 211–230. https://doi.org/10.4310/HHA.2010.v12.n2.a7.

Schumann, A. 2023. *Archaeology of logic.* 1st ed. Boca Raton: CRC Press/Taylor & Francis Group.

Schumann, A., and J. Lemanski. 2022. Logic, spatial algorithms and visual reasoning. *Logica Universalis* 16:535–543. https://doi.org/10.1007/s11787-022-00311-x.

Stong, R. E. 1966. Finite topological spaces. *Transactions of the American Mathematical Society* 123:325–340. https://doi.org/10.2307/1994465.

Univalent Foundations Program. 2013. *Homotopy type theory: univalent foundations of mathematics.* Princeton, NJ: Institute for Advanced Study. https://homotopytypetheory.org/book/.

Walker, J. W. 1981. Homotopy type and euler characteristic of partially ordered sets. *European Journal of Combinatorics* 2 (4): 373–384. https://doi.org/10.1016/S0195-6698(81)80045-5.

6. Gravitation

Hubert Antoine

Kielnarowa Review (2025), 2, 68–95
doi:https://doi.org/10.57656/kr-2025-0006

Kielnarowa
REVIEW

ARTICLE

Gravitation

Hubert Antoine*

Sagita SA, Esneux, Belgium
*Corresponding author. Email: h.antoine@sagita.be

Abstract

Gravitation is introduced as a gauge field that couples to the impulsion in the Lagrangian approach to the classical mechanics. The equations of motion in a central static field yield an amazing result: a repulsive gravitational potential appears at short distance. As a consequence black holes would have a stable and finite size, thus eliminating gravitational singularities. The expansion of the universe and its acceleration could be explained without recourse to a hypothetical dark energy.

1. Introduction

A short historical review

After Einstein produced the famous General Relativity (GR) theory for gravitation, several theories were then also proposed to either improve, reinterpret or supplement GR. The aim of these efforts was to quantify gravitation and explain later discoveries like the expansion of the universe. Several of these theories are based on the non-zero torsion of the space time. Elie Cartan did develop a model where torsion would be generated by angular momentum Scholz 2018; Hehl and Obukhov 2007 Blagojevis and Hehl 2012: Cartan associates to each closed infinitesimal contour a rotation (which expresses curvature) and a translation (which expresses torsion). In Cartan's mind the rotation can be represented by a vector and the translation by a torque. Tetrad formalism, teleparallel gravity, Weitzenböck and Möller theories are shown to be equivalent to GR in reference Arcos and J G Pereira 2005, and reference Schucking 2008 shows that the Schwartzschild metric can have an interpretation of teleparallelism in the Pound-Rebka experiment. An Introduction to teleparallel gravity is given in reference Lurie 2013 where curvature =0 and torsion is the gravitational field strength. In consequence, there are no geodesics in Teleparallel Gravity, only force equations quite analogous to the Lorentz force equation of electrodynamics. The authors expected this result by because, like electrodynamics, Teleparallel Gravity is also a gauge theory. Gauge theories for gravitation are treated in Hayashi and Nakano 1967 in a formal mathematical frame, and in Arcos and J G Pereira 2005; Kleinert 2010 where a mathematical formalism in which torsion and curvature can be exchanged via a supergauge symmetry leads to the GR equations. Translation gauge potentials Ivanenko and Sardanashvily 1987 meet Cartan's idea of the spin of matter being the source of torsion: The gauge gravitation theory based on the relativity and equivalence principles reformulated in fibre bundle terms is the gravitation theory with torsion whose source is the spin of matter. "Since translation gauge potentials fail to be utilized for describing a gravitational field, a question on their physical meaning arises." "Therefore, translation gauge potentials may be responsible for weaker forces than gravity as discussed by some of the authors." Goldstonic supergravity, supergroups, supertransformations, superspace, supersymmetries, superfunctions etc. are examined in a mathematical frame in this paper. The expansion of the universe and its acceleration are explained in the framework of GR by the action of a phantom dark energy Peebles and Ratra 2002; Dutta and Scherrer 2009 in the ΛCDM model.

The present work

This is a new approach on gravitation. In the opposition to GR it is defined as a gauge theory with the gauge group of translations. The source is the energy impulsion tensor that is conserved which results from the invariance of the Lagrangian under the group of translations. Gauge theories introduced by Hermann Weyl postdate GR and are now a successful mathematical formalism in providing a unified framework to describe the quantum field theories of electromagnetism, the weak force and the strong force. Here our gauge symmetry group of translations is abelian as is the U(1) group for electromagnetism and this makes the mathematical treatment simpler. The graviton is defined as a massless spin 2 boson. At this stage one can notice that abelian gauge theories have massless gauge bosons with the infinite range.

We will follow the classical Lagrangian approach where space-time is distorted by the presence of a gravitational field.This distortion will induce a change into the metric. At infinite distance from massive bodies where the gravitational field tends to zero, the metric is Minkovskian $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$. We will begin with the introduction of the gravitational potential in space-time and into the metric. With the Pound-Rebka effect as a first result. Then we will introduce the gravitational potential into the Lagrangian that leads to the classical equations of motion. This will lead to the equations of motion of a body in a gravitational field. The action change under gauge transforms of the gravitational field is evaluated, as well as the symmetry properties of the gravitational potential. The equations of the field are thus derived. We can then apply the equations of motion in a central static field, beginning with the deflection of light by a massive body, the sun in this case. Newton's law is then demonstrated, as is the principle of equivalence. The non-relativistic and relativistic equations of motion in a central static field are expressed in polar coordinates. The precession of Mercury's perihelion is calculated which fits well to the measured value. The equations also show that black holes must have a finite size, when the attractive and repulsive gravitational forces are in equilibrium. Some considerations on black hole properties follow, and also on the accelerated expansion of the Universe.

2. The gravitational field Γ^μ_ν

Time-space, vectors, metric, scalars and so on

In our 4-dimensional space-time, we can meet vectors (ex: x^μ), covectors (ex: p_μ), scalars that are the contraction of a vector on a covector: $x^\mu a_\mu = (x|a)$. A scalar can also be produced from two vectors x^μ and y^ν with the help of the metric $g_{\mu\nu}$: $x^\mu g_{\mu\nu} y^\nu$ and conversely $x_\mu g^{\mu\nu} y_\nu$ with two covectors x_μ and y_ν , $g^{\mu\nu}$ being the inverse matrix of $g_{\mu\nu}$. We can consider that $x_\mu g^{\mu\nu}$ is a vector x^ν that contracts with y_ν In the free (free means that there is no gravitation) space-time, the metric is $\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ the Minkovskian metric where $\eta_{\mu\nu} = \eta^{\mu\nu}$ is its own inverse.

Now a word about the contraction operator that creates a scalar from a vector and a covector: Let the vector be $x^\mu = \begin{pmatrix} x^0 & x^1 & x^2 & x^3 \end{pmatrix}$ and the covector be $a_\nu = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 \end{pmatrix}$. The scalar $(x|a)$ is made by multiplying x^0 with a_0, x^1 with a_1 and so on and then taking the sum. It could have been $x^0 a_1 + x^1 a_2 + x^2 a_3 + x^3 a_0$ or any other mix of indices but it is not. The operator governing this contraction is the Kronecker δ^ν_μ that will make that x^0 matches with a_0, x^1 with a_1 and so on. Thus, we write: $(x|a) = x^\mu \delta^\nu_\mu a_\nu$ Also $x_\mu g^{\mu\nu} y_\nu = x_\mu \delta^\mu_\lambda g^{\lambda\rho} \delta^\nu_\rho y_\nu$: here we have a double contraction on μ and ν. All this may seem trivial, but we will demonstrate that the gravitational field acts on the contraction operator δ^ν_μ in the following way: δ^ν_μ becomes $\delta^\nu_\mu + \Gamma^\nu_\mu$ when acting on the impulsion

p_ν. For instance, this modifies the differential of the action

$$dS = p_\mu dx^\mu \text{ to } dS' = p_\mu(\delta^\mu_\nu + \Gamma^\mu_\nu)dx^\nu = p_\mu dx^\mu + p_\mu\Gamma^\mu_\nu dx^\nu$$

. This can be compared to the action that includes an electromagnetic field A_μ: $dS' = p_\mu dx^\mu + eA_\mu dx^\mu$ Mass m will also be a subject to the change in a gravitational field. One had $m^2 = p_\mu \eta_{\mu\nu} p_\nu$ or $m^2 = \frac{E^2}{c^4} - p^2$ in the free space. In a gravitational field $p_\mu = \delta^\nu_\mu p_\nu$ becomes

$$p_\mu = (\delta^\nu_\mu + \Gamma^\nu_\mu)p_\nu \text{ and } m'^2 = p_\lambda(\delta^\lambda_\mu + \Gamma^\lambda_\mu)\eta^{\mu\nu}(\delta^\rho_\nu + \Gamma^\rho_\nu)p_\rho \simeq m^2 + 2p_\mu\Gamma^\rho_\nu\eta^{\rho\nu}p_\nu + \Gamma\Gamma \text{ terms.}$$

This explains the slight mass gain of a body moving in a central static potential. We can also write: $m^2 = p_\lambda g^{\lambda\rho} p_\rho$ with the metric in a gravitational field becoming $g^{\lambda\rho} = (\delta^\lambda_\mu + \Gamma^\lambda_\mu)\eta^{\mu\nu}(\delta^\rho_\nu + \Gamma^\rho_\nu)$ a metric for covectors. And $g_{\lambda\rho} = (g^{\lambda\rho})^{-1} = (\delta^\mu_\lambda - \Gamma^\mu_\lambda)\eta_{\mu\nu}(\delta^\nu_\rho - \Gamma^\nu_\rho)$ would be the metric for vectors. Developing to the first order in Γ we get: $g^{\lambda\rho} = \eta^{\lambda\rho} + 2\Gamma^\lambda_\mu\eta^{\mu\rho}$ and $g_{\lambda\rho} = \eta_{\lambda\rho} - 2\Gamma^\mu_\lambda\eta_{\mu\rho}$

It will be shown later that Γ^μ_ν is symmetrical in μ and ν and can be made diagonal by a suitable change of coordinates. This change leaves δ^μ_ν invariant. Then $g^{\lambda\rho}$ can be written, to the first order in Γ:

$$g^{\lambda\rho} = \begin{pmatrix} 1 + 2\Gamma^0_0 & 0 & 0 & 0 \\ 0 & -1 - 2\Gamma^1_1 & 0 & 0 \\ 0 & 0 & -1 - 2\Gamma^2_2 & 0 \\ 0 & 0 & 0 & -1 - 2\Gamma^3_3 \end{pmatrix}$$

And

$$g_{\lambda\rho} = \begin{pmatrix} 1 - 2\Gamma^0_0 & 0 & 0 & 0 \\ 0 & -1 + 2\Gamma^1_1 & 0 & 0 \\ 0 & 0 & -1 + 2\Gamma^2_2 & 0 \\ 0 & 0 & 0 & -1 + 2\Gamma^3_3 \end{pmatrix}$$

The Pound and Rebka experiment

The Pound and Rebka experiment shows how the time is modified by the gravitation: thus only Γ^0_0 is acting and dt^2 becomes $dt^2(1 - 2\Gamma^0_0)$. The scalar time lapse $|dt|$ becomes $|dt|(1 - \Gamma^0_0)$. Let us write Γ for Γ^0_0. The Figure 1 shows the layout of the experiment:

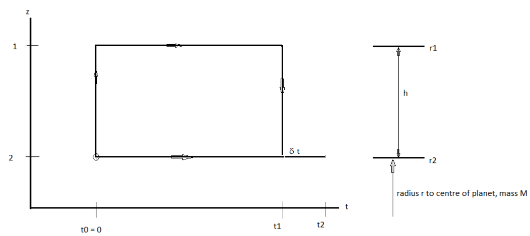


Figure 1. Pound-Rebka experiment

At point 1, $dt_1 = (1 - \Gamma_1)dt$. At point 2, $dt_2 = (1 - \Gamma_2)dt$, so that:

$$\frac{dt_1}{dt_2} = \frac{1 - \Gamma_1}{1 - \Gamma_2} \simeq 1 - \Gamma_1 + \Gamma_2 \qquad (1)$$

$\frac{dt_1}{dt_2}$ was measured by Pound and Rebka in 1960 to be equal to: $1 - \frac{GM}{c^2}(\frac{1}{r_1} - \frac{1}{r_2})$. Which allows us to identify Γ^0_0 to $\frac{GM}{rc^2}$. Here M is the earth mass, G the gravitational constant, c the speed of light and $r_{1,2}$ the distance of points 1 and 2 to the center of Earth.

3. Equations of motion in the gravitational field Γ^μ_ν

Classical mechanics tells us that the trajectory of a body between two points A and B is the path where the integral of the action dS is minimal, with $dS = p_\mu dx^\mu$ and p_μ being the impulsion of the body. The occurrence of a field Γ^μ_ν will influence dx^μ and generate the equations of motion of a body in a gravitational field.

Let us consider the Lagrangian equation of motion, seen from a geometrical point of view. A small deviation from the trajectory between the two points A and B will not change the action to the first order:

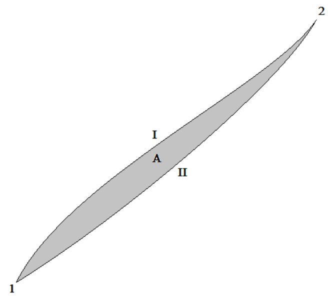


Figure 2. Action variation

$$\int_I p_\mu \, dx^\mu = \int_{II} p_\mu \, dx^\mu$$

or: $\oint_{I,-II} p_\mu \, dx^\mu = 0$ which is equivalent to: $\iint_A (\partial_\mu p_\nu - \partial_\nu p_\mu) dS = 0$. Thus, $\partial_\mu p_\nu - \partial_\nu p_\mu = 0$

Multiplying by $v^\mu = \frac{dx^\mu}{dt}$, one gets: $\frac{dx^\mu}{dt}\frac{\partial p_\nu}{\partial x^\mu} - v^\mu \partial_\nu p_\mu = 0$. Since $\partial_\nu v^\mu = 0$ and $\frac{dx^\mu}{dt}\frac{\partial p_\nu}{\partial x^\mu} = \frac{d}{dt} p_\nu$ we get: $\frac{d}{dt} p_\nu - \frac{\partial}{\partial x^\nu}(v^\mu p_\mu) = 0$

But $v^\mu p_\mu$ is the Lagrangian $\mathcal{L}$, defined as the time derivative of the action:

$$S = \int p_\mu \, dx^\mu = \int p_\mu v^\mu \, dt = \int \mathcal{L} \, dt.$$

We can write the equation of motion as:

$$\frac{d}{dt} p_\nu - \frac{\partial}{\partial x^\nu} \mathcal{L} = 0$$

or $\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial v^\nu} - \frac{\partial}{\partial x^\nu}\mathcal{L} = 0$ with $p_\nu = \frac{\partial \mathcal{L}}{\partial v^\nu}$

When the gravitational field Γ^μ_ν is introduced we get the following action: $S = \int p_\mu (\delta^\mu_\nu + \Gamma^\mu_\nu) \, dx^\nu$. Thus $S = \int p^\mu (\delta^\nu_\mu + \Gamma^\nu_\mu) v^\nu \, dt = \int \mathcal{L} \, dt$ which implies

$$\mathcal{L} = p_\mu (\delta^\mu_\nu + \Gamma^\mu_\nu) v^\nu.$$

Equations of motion

1st term:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial v^\nu} = \frac{d}{dt}(p_\nu + \Gamma^\mu_\nu p_\mu) = \frac{d}{dt} p_\nu + p_\mu \frac{d}{dt}\Gamma^\mu_\nu + \Gamma^\mu_\nu \frac{d}{dt} p_\mu \tag{2}$$

$$= \frac{d}{dt} p_\nu + p_\mu v^\rho \frac{\partial \Gamma^\mu_\nu}{\partial x^\rho} + \Gamma^\mu_\nu v^\rho \frac{\partial p_\mu}{\partial x^\rho} \tag{3}$$

2d term:

$$\frac{\partial \mathcal{L}}{\partial x^\nu} = \frac{\partial}{\partial x^\nu}(p_\rho v^\rho) + \frac{\partial}{\partial x^\nu}(p_\mu \Gamma^\mu_\rho v^\rho) \tag{4}$$

We simplify the notation by: $\partial_\nu \equiv \frac{\partial}{\partial x^\nu}$. Then, since $\partial_\nu v^\rho = 0$

$$\partial_\nu \mathcal{L} = v^\rho \partial_\nu p_\rho + p_\mu v^\rho \partial_\nu \Gamma^\mu_\rho + v^\rho \Gamma^\mu_\rho \partial_\nu p_\mu \tag{5}$$

and assembling the simplified terms in the Lagrange equation, we get:

$$v^\mu \partial_\mu p_\nu + \Gamma^\mu_\nu v^\rho \partial_\rho p_\mu = p_\mu v^\rho (\partial_\nu \Gamma^\mu_\rho - \partial_\rho \Gamma^\mu_\nu) + v^\rho \Gamma^\mu_\rho \partial_\nu p_\mu + v^\mu \partial_\nu p_\mu \tag{6}$$

Or:

$$v^\mu(\partial_\mu p_\nu + \partial_\nu p_\mu) + v^\rho(\Gamma^\mu_\nu \partial_\rho p_\mu + \Gamma^\mu_\rho \partial_\nu p_\mu) = p_\mu v^\rho(\partial_\nu \Gamma^\mu_\rho - \partial_\rho \Gamma^\mu_\nu) \tag{7}$$

and finally

$$\begin{aligned} v^\mu(\partial_\mu p_\nu - \partial_\nu p_\mu) &= v^\rho[\partial_\nu(p_\mu \Gamma^\mu_\rho) - \partial_\rho(p_\mu \Gamma^\mu_\nu)] \\ \Leftrightarrow v^\rho(\partial_\rho p_\nu - \partial_\nu p_\rho) &= v^\rho[\partial_\nu(p_\mu \Gamma^\mu_\rho) - \partial_\rho(p_\mu \Gamma^\mu_\nu)]. \end{aligned} \tag{8}$$

The equation (8) is satisfied if

$$\partial_\rho(p_\nu + p_\mu \Gamma^\mu_\nu) - \partial_\nu(p_\rho + p_\mu \Gamma^\mu_\rho) = 0. \tag{9}$$

Temporal and spatial fields

Let us define $\gamma_\nu \equiv p_\mu \Gamma^\mu_\nu$ then, the equation (8) can be rewritten as:

$$v^\rho \partial_\rho p_\nu - v^\rho \partial_\nu p_\rho = v^\rho(\partial_\nu \gamma_\rho - \partial_\rho \gamma_\nu). \tag{10}$$

Or, with $v^\rho \partial_\rho = \frac{d}{dt}$ and $v^\rho \partial_\nu p_\rho = \partial_\nu v^\rho p_\rho = \partial_\nu \mathcal{L}$ as:

$$\frac{d}{dt} p_\nu = \frac{\partial \mathcal{L}}{\partial x^\nu} + v^\rho(\partial_\nu \gamma_\rho - \partial_\rho \gamma_\nu). \tag{11}$$

If no other extra fields are present, $\frac{\partial \mathcal{L}}{\partial x^\nu}$ can be ignored and for $\nu = 0$ one gets:

$$\frac{d}{dt} p_0 = v^i(\partial_t \gamma_i - \partial_i \gamma_0) \; ; \; i = 1, 2, 3. \tag{12}$$

Or, if we define $E_i \equiv \partial_i \gamma_0 - \partial_t \gamma_i$ we get:

$$\frac{d}{dt} p_0 = -v^i E_i. \tag{13}$$

For $\nu = 1, 2, 3$ noted as $\nu = i$:

$$\frac{d}{dt} p_i = v^\rho(\partial_i \gamma_\rho - \partial_\rho \gamma_i) \; ; \; \rho = 0, 1, 2, 3. \tag{14}$$

For $\rho = 0$, $v^0 \equiv 1$ and we get: $1(\partial_i \gamma_0 - \partial_t \gamma_i) = E_i$ again. For $\rho = 1, 2, 3$ noted as $\rho = j$ one has:

$$v^j(\partial_i \gamma_j - \partial_j \gamma_i) = \vec{v} \times \overrightarrow{rot \vec{\gamma}}. \tag{15}$$

If we define $\vec{H} \equiv \overrightarrow{rot \vec{\gamma}}$ we finally get:

$$\frac{d}{dt} \vec{p} = \vec{E} + (\vec{v} \times \vec{H}). \tag{16}$$

4. Symmetry of Γ^{ν}_{μ}

Γ^{ν}_{μ} is a mixed co and contravariant tensor (its product with $p_{\nu}v^{\mu}$ contributes to the Lagrangian as a scalar). This 4×4 tensor can have a symmetric and antisymmetric parts. We show that its antisymmetric part would correspond to a transform where the 4-length is preserved and there is no distortion.

If $dx^{\mu}dx^{\nu}\eta_{\mu\nu} = dx^2$ is the 4-length of the vector dx, then after Γ operation the length of the vector will become:

$$(dx^{\mu} + \Gamma^{\mu}_{\lambda}dx^{\lambda})(dx^{\nu} + \Gamma^{\nu}_{\rho}dx^{\rho})\eta_{\mu\nu} =$$
$$= dx^{\mu}dx^{\nu}\eta_{\mu\nu} + dx^{\mu}\eta_{\mu\nu}\Gamma^{\nu}_{\rho}dx^{\rho} + dx^{\nu}\eta_{\mu\nu}\Gamma^{\mu}_{\lambda}dx^{\lambda} + \eta_{\mu\nu}\Gamma^{\mu}_{\lambda}\Gamma^{\nu}_{\rho}dx^{\lambda}dx^{\rho}$$

To the first order in Γ, and if the length is preserved, we would get: $dx^{\mu}dx^{\nu}\eta_{\mu\nu} = dx^{\mu}dx^{\nu}\eta_{\mu\nu} + dx^{\nu}\Gamma_{\nu\rho}dx^{\rho} + dx^{\mu}\Gamma_{\mu\rho}dx^{\rho}$. Switching the dummy index ν to μ in the second term $2\Gamma_{\mu\rho}dx^{\rho}dx^{\mu} = 0$ and by symmetry of $dx^{\rho}dx^{\mu}$, $\Gamma_{\mu\rho}$ should be antisymmetric. The symmetrical part of Γ^{ν}_{μ} is $\eta_{\mu\lambda}\Gamma^{\nu\lambda}$ and also $\eta_{\mu\lambda}$ and $\Gamma^{\nu\lambda}$ both are symmetrical. Thus $\Gamma^{\mu}_{\nu} = \Gamma^{\nu}_{\mu}$ The symmetry of μ and ν in Γ^{ν}_{μ} ensures that it will produce no rotation (or more generally conservation of the 4-length) but it will produce only the distortion of space time. As a consequence Γ^{ν}_{μ} will have 10 components and Γ^{ν}_{μ} could possibly correspond to a spin 2 graviton. Feynman 1995

To make it visual, consider the following Figures 3 and 4. On Fig. 3, the symmetrical $\Gamma^{2}_{1} = \Gamma^{1}_{2}$ creates a distortion in the 1, 2 plane. $\vec{1'} + \vec{2'} = \vec{1} + \vec{2}\Gamma^{2}_{1} + \vec{2} + \vec{1}\Gamma^{1}_{2} = (\vec{1} + \vec{2})(1 + \Gamma^{2}_{1})$

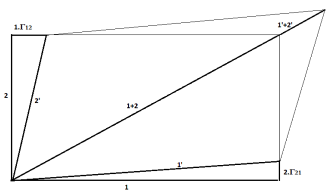


Figure 3. distortion

If $\Gamma^{2}_{1} = -\Gamma^{1}_{2}$ then we would have a rotation instead of distortion, and as a consequence Γ^{0}_{0} would also be = 0 instead of $\frac{GM}{rc^2}$. $\vec{1'} + \vec{2'} = \vec{1} + \vec{2}\Gamma^{2}_{1} + \vec{2} + \vec{1}\Gamma^{1}_{2} = \vec{1} + \vec{2}\Gamma^{2}_{1} + \vec{2} - \vec{1}\Gamma^{2}_{1}$.

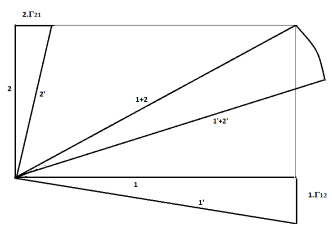


Figure 4. rotation

5. The equations of field

We look for an action that is a scalar, gauge invariant and that includes only the Γ^{ν}_{μ} terms. Such an action will be noted as S_f. Let us start from 11: $\frac{d}{dt}p_{\nu} = \frac{\partial \mathcal{L}}{\partial x^{\nu}} + v^{\rho}(\partial_{\nu}\gamma_{\rho} - \partial_{\rho}\gamma_{\nu})$. The term

$(\partial_\nu \gamma_\rho - \partial_\rho \gamma_\nu)$ is invariant under gauge transform $\Gamma^\mu_\nu \rightarrow \Gamma^\mu_\nu + \frac{\partial_\nu \Phi}{p_\mu}$. The action term $\int p_\mu \Gamma^\mu_\nu \, dx^\nu$ is also invariant under the same gauge transform.

Expanding and neglecting $\frac{\partial \mathcal{L}}{\partial x^\nu}$ (no other potential but gravitational is present), we get:

$$\frac{d}{dt} p_\nu = v^\rho p_\mu \underbrace{(\partial_\nu \Gamma^\mu_\rho - \partial_\rho \Gamma^\mu_\nu)}_{I} + v^\rho \underbrace{(\Gamma^\mu_\rho \partial_\nu p_\mu - \Gamma^\mu_\nu \partial_\rho p_\mu)}_{II} \tag{17}$$

The term I is invariant under the gauge transform $\Gamma^\mu_\nu \rightarrow \Gamma^\mu_\nu + \partial_\nu \, G^\mu$. The gauge transform leaving invariant γ_ν and I must satisfy $\frac{\partial_\nu \Phi}{p_\mu} = \partial_\nu \, G^\mu$. The scalar term is $(\partial_\nu \Gamma^\mu_\rho - \partial_\rho \Gamma^\mu_\nu)^2$ and it is defined as:

$$(\partial_\nu \Gamma^\mu_\rho - \partial_\rho \Gamma^\mu_\nu) g^{\nu\alpha} g^{\rho\beta} g_{\mu\gamma} (\partial_\alpha \Gamma^\gamma_\beta - \partial_\beta \Gamma^\gamma_\alpha). \tag{18}$$

We define the action of the field as:

$$S_f = \alpha \int (\partial_\nu \Gamma^\mu_\rho - \partial_\rho \Gamma^\mu_\nu)^2 \, d\Omega. \tag{19}$$

With α an arbitrary constant, $d\Omega = \sqrt{-g} \, dVdt$, V is the spatial volume and $g = det(g_{ij}) = -1 + 2Tr\Gamma$, Γ is the Γ^μ_ν matrix. The action is now completed with the field – matter interaction term:

$$S = \int p_\mu \Gamma^\mu_\nu v^\nu \, d\Omega + \alpha \int (\partial_\nu \Gamma^\mu_\rho - \partial_\rho \Gamma^\mu_\nu)^2 \, d\Omega \tag{20}$$

where p_μ stands for the density of impulsion.

How does the entire action vary under a variation of the potential Γ?

$$\delta S = \int \delta [p_\mu v^\nu \, \Gamma^\mu_\nu + \alpha \, (\partial_\nu \Gamma^\mu_\rho - \partial_\rho \Gamma^\mu_\nu)^2 \, (1 - Tr\Gamma)] dV \, dt. \tag{21}$$

The term in α is equal to: $([\partial_\nu \Gamma^\mu_\rho - \partial_\rho \Gamma^\mu_\nu](\eta^{\nu\alpha} + 2\eta^{\nu\lambda}\Gamma^\alpha_\lambda)(\eta^{\rho\beta} + 2\eta^{\rho\kappa}\Gamma^\beta_\kappa)(\eta_{\mu\gamma} - 2\eta_{\mu\lambda}\Gamma^\lambda_\gamma)[\partial_\alpha \Gamma^\gamma_\beta - \partial_\beta \Gamma^\gamma_\alpha] \, (1 - Tr\Gamma))$.

The variation of this product of six terms can be much simplified if we consider only the first order terms in Γ in the product of $g^{\nu\alpha} = \eta^{\nu\alpha}$ and $1 - Tr\Gamma = 1$. Thus, we get:

$$\delta S = \int p_\mu v^\nu \, \delta \Gamma^\mu_\nu + \alpha \delta [(\partial_\nu \Gamma^\mu_\rho - \partial_\rho \Gamma^\mu_\nu) \, (\partial^\nu \Gamma^\rho_\mu - \partial^\rho \Gamma^\nu_\mu)] \, dVdt \tag{22}$$

or:

$$\delta S = \int p_\mu v^\nu \, \delta \Gamma^\mu_\nu + 2\alpha (\partial_\nu \Gamma^\mu_\rho - \partial_\rho \Gamma^\mu_\nu) \, \delta (\partial^\nu \Gamma^\rho_\mu - \partial^\rho \Gamma^\nu_\mu) \, dVdt \tag{23}$$

where $\partial^\nu = \eta^{\nu\mu} \partial_\mu$.

$$\delta S = \int p_\mu v^\nu \, \delta \Gamma^\mu_\nu + 2\alpha [(\partial_\nu \Gamma^\mu_\rho - \partial_\rho \Gamma^\mu_\nu) \, \partial^\nu \delta \Gamma^\rho_\mu \tag{24}$$

$$- (\partial_\nu \Gamma^\mu_\rho - \partial_\rho \Gamma^\mu_\nu) \, \partial^\rho \delta \Gamma^\nu_\mu] \, dVdt. \tag{25}$$

We swapped ∂ and δ, and by swapping ρ and ν we get:

$$\delta S = \int p_\mu v^\nu \, \delta \Gamma^\mu_\nu + 4\alpha (\partial_\rho \Gamma^\mu_\nu - \partial_\nu \Gamma^\mu_\rho) \, \partial^\rho \delta \Gamma^\nu_\mu \, dVdt. \tag{26}$$

The term in 4α is integrated by parts:

$$\int (\partial_\rho \Gamma^\mu_\nu - \partial_\nu \Gamma^\mu_\rho)\, \partial^\rho \delta \Gamma^\nu_\mu \, dVdt \tag{27}$$

$$= - \int \partial^\rho (\partial_\rho \Gamma^\mu_\nu - \partial_\nu \Gamma^\mu_\rho)\, \delta \Gamma^\nu_\mu \, dVdt$$

$$+ \int (\partial_\rho \Gamma^\mu_\nu - \partial_\nu \Gamma^\mu_\rho)\, \delta \Gamma^\nu_\mu \, dS^\rho. \tag{28}$$

The term $\int (\partial_\rho \Gamma^\mu_\nu - \partial_\nu \Gamma^\mu_\rho)\, \delta \Gamma^\nu_\mu \, dS^\rho = 0$ since $\delta \Gamma^\nu_\mu = 0$ in the time limits and $(\partial_\rho \Gamma^\mu_\nu - \partial_\nu \Gamma^\mu_\rho) = 0$ at ∞. The field strength is 0 on the boundary at ∞.

Thus we obtain:

$$\delta S = \int p_\mu v^\nu \, \delta \Gamma^\mu_\nu - 4\alpha \partial^\rho (\partial_\rho \Gamma^\mu_\nu - \partial_\nu \Gamma^\mu_\rho)\, \delta \Gamma^\nu_\mu \, dVdt. \tag{29}$$

By cancelling the variation of S and swapping μ and ν in the first term we have:

$$p_\nu v^\mu - 4\alpha \partial^\rho (\partial_\rho \Gamma^\mu_\nu - \partial_\nu \Gamma^\mu_\rho) = 0. \tag{30}$$

One corollary of eq. 30 is that the divergence of the energy impulsion tensor $(p_\nu v^\mu)$ is equal to zero. Indeed, swapping the dummy indices ν and ρ, we get:

$$\partial^\nu (p_\nu v^\mu) = 4\alpha(\partial^\nu \partial^\rho \partial_\rho \Gamma^\mu_\nu - \partial^\rho \partial^\nu \partial_\nu \Gamma^\mu_\rho)$$

$$= 4\alpha(\partial^\nu \partial^\rho \partial_\rho \Gamma^\mu_\nu - \partial^\nu \partial^\rho \partial_\rho \Gamma^\mu_\nu) = 0. \tag{31}$$

Now let us evaluate 4α. If the source current $p_\mu v^\nu$ is generated by a mass with rest density ρ $p_\mu v^\nu = \rho\, \delta^\nu_0 \delta^0_\mu c^2$ and from (30) we get:

$$\rho\, c^2 = 4\alpha[\partial_\nu \partial^\nu \Gamma^0_0 - \partial^\nu \partial_t \Gamma^0_\nu]. \tag{32}$$

With the mass density ρ at rest, the field Γ^0_ν must be static: $\partial_t \Gamma^0_\nu = 0$ and $\rho\, c^2 = 4\alpha \Delta \Gamma^0_0$ with Δ the Laplacian. A solution is: $\Gamma^0_0 = \frac{1}{4\pi} \int \frac{c^2 \rho}{4\alpha r}\, dV$ which for a mass $M = \int \rho\, dV$ and since $\Gamma^0_0 = \frac{GM}{rc^2}$, we get:

$$16\pi\alpha = \frac{c^4}{G}\, [m\frac{kg}{s^2}] \rightarrow 4\alpha = \frac{c^4}{4\pi G}. \tag{33}$$

Eq. 30 can be rewritten:

$$\partial^\rho (\partial_\rho \Gamma^\mu_\nu - \partial_\nu \Gamma^\mu_\rho) = \frac{4\pi G}{c^4} p_\nu v^\mu \, [m^{-2}] \tag{34}$$

where p_ν is the density of impulsion.

Depending on the values of μ, ν we have the following field equations:

1) μ, ν =0; $p_0 v^0 = \rho c^2$

$$\frac{4\pi G}{c^2}\rho = \partial_\lambda \partial^\lambda \Gamma^0_0 - \partial^\lambda \partial_t \Gamma^0_\lambda \, [m^{-2}] \tag{35}$$

and if the field is static, we get:

$$\Delta \Gamma^0_0 = \frac{4\pi G}{c^2}\rho \rightarrow \Gamma^0_0 = \frac{GM}{rc^2} \tag{36}$$

with $M = \int \rho \, dV$. The equations of motion in a central static field will be considered in the next chapter.

2) Let $\mu, \nu \neq 0$ are denoted by i, j, then for $v \ll c$, $p \simeq \rho v$ we get:

$$\frac{4\pi G}{c^4} \rho v_i v^j = \partial^\lambda (\partial_\lambda \Gamma_i^j - \partial_i \Gamma_\lambda^j) \, [m^{-2}]. \tag{37}$$

This is symmetric in i and j on the left-hand-side and can be made symmetrical in i and j on the right-hand-side because $\Gamma_i^j = \Gamma_j^i$ and by a kind of "Lorentz" condition: $\partial^\lambda \Gamma_\lambda^j = 0$.

Gravitational waves

Equation (34) can be rewritten assuming the above "Lorentz condition": $\partial^\lambda \Gamma_\lambda^j = 0$ as:

$$\partial^\rho \partial_\rho \Gamma_\nu^\mu = \frac{4\pi G}{c^4} p_\nu v^\mu. \tag{38}$$

In the vacuum:

$$\partial^\rho \partial_\rho \Gamma_\nu^\mu = 0 = (\frac{1}{c^2} \partial_t \partial_t - \sum_i \partial_i \partial_i) \Gamma_\nu^\mu$$

the field Γ_ν^μ (the massless graviton) is a wave propagating at speed of light. It can be seen as a massless particle that propagates a gravitational field.

6. The test of deflection of light by the sun

At this point we have enough results to perform the first test of this new theory, it is a kind of "stop and go" procedure as in the following scheme:

The gravitational potential of the sun is considered as central and static:

$$\Gamma_\nu^\mu = \delta_0^\mu \, \delta_\nu^0 \, \Gamma_0^0 \equiv \Gamma = \frac{GM}{rc^2}. \tag{39}$$

The impulsion of the photon is $\vec{p} = \hbar \vec{k}$ and $\vec{k} = (\omega, k_0 \infty, 0, 0) = (\omega, \omega, 0, 0)$ in the x, y plane shown

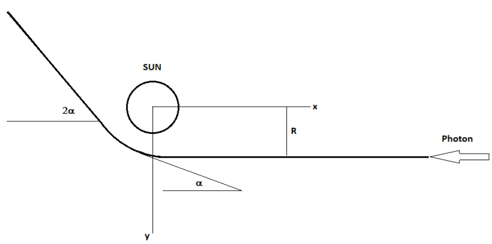


Figure 5. Deflection of light

on Fig. 5 and $\vec{v} = (1, -1, 0, 0)$ with $c = 1$; in the following calculations we also take $c = 1 = \hbar$ for clarity.

With the symbol $\dot{k} = \frac{d}{dt} k$ eq. 11 gives:

$$\dot{k_\nu} = v^\rho \, [\partial_\nu (k_\mu \Gamma_\rho^\mu) - \partial_\rho (k_\mu \Gamma_\nu^\mu)] \tag{40}$$

With Γ^{μ}_{ν} defined by eq. 39, it gives rise to:

$$\dot{k_\nu} = v^0 \partial_\nu (k_0 \Gamma) - v^\rho \partial_\rho (k_\mu \Gamma^\mu_\nu) \tag{41}$$

For $\dot{k_0}$: $\dot{k_0} = v_0 \partial_t (k_0 \Gamma) - v_\rho \partial_\rho (k_0 \Gamma)$; $\rho = (1, 2, 3)$

$$\dot{k_0} = \partial_t (k_0 \Gamma) + \partial_1 (k_0 \Gamma) = \Gamma \partial_t k_0 + k_0 \partial_t \Gamma + \Gamma \partial_1 k_0 + k_0 \partial_1 \Gamma$$

since $\partial_t \Gamma = 0$, $\dot{k_0} = \Gamma(\partial_t k_0 + \partial_1 k_0) + k_0 \partial_1 \Gamma$.

With $x = x_\infty - ct = x_\infty - t$ (with c=1), $\rightarrow \partial_t = -\partial_1$ (where $\partial_1 \equiv \partial_x$). Thus,

$$\dot{k_0} = k_0 \partial_1 \Gamma \tag{42}$$

For $\dot{k_i}$: $\dot{k_i} = v^0 \partial_i (k_0 \Gamma) - v^\rho \partial_\rho (k_\mu \Gamma^\mu_i) = \partial_i (k_0 \Gamma)$. Remembering that for $v_0 = 1$, $\Gamma^i_\mu = 0 \,; i \neq 0$.

Along the coordinate $y = 2$, $\dot{k_2} = \partial_2 (k_0 \Gamma) = \partial_y (k_0 \Gamma) = k_0 \partial_2 \Gamma$, since $\partial_2 k_0 = 0$. Thus:

$$\dot{k_2} = k_0 \partial_2 \Gamma \tag{43}$$

With $\dot{k_0} = k_0 \partial_1 \Gamma$ the time derivation of $\dot{k_2}$ reads

$$\ddot{k_2} = \partial_2 (\dot{k_0} \Gamma) = (k_0 \dot{\partial_2} \Gamma) = \dot{k_0} \partial_2 \Gamma + k_0 \partial_2 \dot{\Gamma} \rightarrow \ddot{k_2} = k_0 \partial_1 \Gamma \partial_2 \Gamma + k_0 \partial_2 (v^j \partial_j \Gamma).$$

The first term is in $\partial\Gamma^2$ and can be neglected versus $\partial\Gamma$ ($\Gamma \approx 10^{-6}$ at surface of the sun). Thus, $\ddot{k_2} = k_0 \partial_2 (v^0 \partial_t \Gamma + v^1 \partial_1 \Gamma)$, with $v^0 = 1$, $v^1 = -1$ and $\partial_t = -\partial_1$. This leads to:

$$\ddot{k_2} = k_0 \partial_2 (-\partial_1 \Gamma - \partial_1 \Gamma) = -2 k_0 \partial_1 \partial_2 \Gamma. \tag{44}$$

Since $\Gamma = \frac{GM}{r} = \frac{GM}{\sqrt{x^2+y^2}}$ with $c = 1$

$$\partial_2 \Gamma = \frac{-yGM}{(x^2+y^2)^{3/2}}$$

$$\partial_1 \partial_2 \Gamma = \frac{3xyGM}{(x^2+y^2)^{5/2}}$$

then

$$\ddot{k_2} = -6 k_0 \frac{xyGM}{(x^2+y^2)^{5/2}}. \tag{45}$$

With $\ddot{k_2} = \frac{d}{dt} \dot{k_2} = \frac{dx}{dt} \frac{d\dot{k_2}}{dx}$ and $\frac{dx}{dt} = -1$ we get $\dot{k_2} = -\int \ddot{k_2}\, dx$. Thus

$$\dot{k_2} = 6 k_0 GM \int \frac{xy}{(x^2+y^2)^{5/2}}\, dx. \tag{46}$$

Integrating again with $dx = -dt$ we obtain:

$$k_2 = 2 k_0 GMy \int_\infty^0 \frac{1}{(x^2+y^2)^{3/2}}\, dx = -2 k_0 \frac{GMy}{y^2}. \tag{47}$$

Thus, $k_2 = -2 k_0 \frac{GM}{y}$ which for $y = R$ leads to

$$\frac{k_2}{k_0} = \frac{-2GM}{R} = \tan(\alpha) \simeq \alpha \tag{48}$$

still with $c = 1$. Reintroducing c we get a total deviation as $2\alpha = \frac{4GM}{Rc^2}$. This corresponds to the measured value and is the first succesful test of the validity of this new gravitational model.

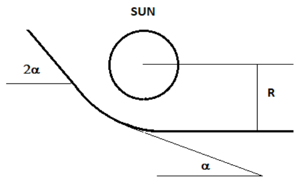


Figure 6. Total deviation

7. Equation of motion in a central static field

The basic assumptions of a central static field are: The field $\Gamma_i^j = \Gamma_0^0\,\delta_0^j\,\delta_i^0$, $\Gamma_0^0 \equiv \Gamma$ $\partial_t\Gamma = 0$ $\dot{\Gamma} = v^j\partial_j\Gamma$, and $\Gamma = \frac{GM}{rc^2}$.

Starting from eq. 11

$$\dot{p_\nu} = v^\rho[\partial_\nu(p_\mu\Gamma^\mu_\rho) - \partial_\rho(p_\mu\Gamma^\mu_\nu)].$$

For $i = 1, 2, 3$ one has:

$$\dot{p_i} = v^0\partial_i(p_0\Gamma) - v^j\partial_j(p_0\Gamma_i^0) = \partial_i(p_0\Gamma). \tag{49}$$

For i=0:

$$\dot{p_0} = v^0\partial_t(p_0\Gamma) - v^i\partial_i(p_0\Gamma). \tag{50}$$

Thus:

$$\dot{p_i} = \partial_i(p_0\Gamma) = p_0\partial_i\Gamma + \Gamma\partial_i p_0 \tag{51}$$

and

$$\dot{p_0} = -v^i\partial_i(p_0\Gamma) + \Gamma\partial_t p_0. \tag{52}$$

We now calculate $\ddot{p_i}$:

$$\ddot{p_i} = \overset{\cdot}{(p_0\partial_i\Gamma + \Gamma\partial_i p_0)} = \dot{p_0}\partial_i\Gamma + p_0\dot{\partial_i}\Gamma + \dot{\Gamma}\partial_i p_0 + \Gamma\partial_i\dot{p_0}. \tag{53}$$

Then replacing $\dot{p_0}$ in eq. 53 we have:

$$\ddot{p_i} = \underbrace{(-v^j\partial_j(p_0\Gamma) + \Gamma\partial_t p_0)\partial_i\Gamma}_{I} + \tag{54}$$

$$\underbrace{(p_0\partial_i(v^j\partial_j\Gamma)}_{II} + \tag{55}$$

$$\underbrace{v^j\partial_j\Gamma\partial_i p_0}_{III} + \tag{56}$$

$$\underbrace{\Gamma\partial_i(-v^j\partial_j(p_0\Gamma) + \Gamma\partial_t p_0)}_{IV}. \tag{57}$$

The term $I + IV$ gives: $\partial_i[\Gamma(-v^j\partial_j(p_0\Gamma) + \Gamma^2\partial_t p_0]$, and the term $II + III$ gives: $\partial_i[p_0 v^j \partial_j \Gamma]$, so thus,

$$\ddot{p}_i = \partial_i[p_0 v^j \partial_j \Gamma + \Gamma^2 \partial_t p_0 - \Gamma v^j \partial_j(p_0\Gamma)] \tag{58}$$

$$= \partial_i[p_0 v^j \partial_j \Gamma + \Gamma^2 \partial_t p_0 - \Gamma v^j (p_0 \partial_j \Gamma + \Gamma \partial_j p_0)] \tag{59}$$

$$= \partial_i[(p_0 v^j \partial_j \Gamma)(1-\Gamma) + \Gamma^2(\partial_t p_0 - v^j \partial_j p_0)] \tag{60}$$

Also $\partial_t p_0 - v^j \partial_j p_0 = \dot{p_0}$ and eq. (52) gives:

$$\dot{p_0} = \Gamma \partial_t p_0 - \Gamma v^i \partial_i p_0 - v^i p_0 \partial_i \Gamma = \Gamma \dot{p_0} - v^i p_0 \partial_i \Gamma. \tag{61}$$

Then $\dot{p_0}(1-\Gamma) = -p_0 v^i \partial_i \Gamma$ or $\dot{p_0} = \frac{-p_0 v^i \partial_i \Gamma}{1-\Gamma}$.

Replacing in eq. (60) we get:

$$\ddot{p}_i = \partial_i \left[(p_0 v^j \partial_j \Gamma)(1-\Gamma) + \frac{\Gamma^2}{1-\Gamma}(-p_0 v^j \partial_j \Gamma) \right] \tag{62}$$

Or:

$$\ddot{p}_i = \partial_i \left[(p_0 v^j \partial_j \Gamma) \frac{1-2\Gamma}{1-\Gamma} \right] ; i,j = 1,2,3 \tag{63}$$

Non relativistic equations of motion :

For $v \ll c$ and no external field, $p_0 \simeq mc^2$ and $p_i = \delta_{ik} v^k m = mv_i$ still for $i,j = 1,2,3$
Equation 63 becomes :

$$\ddot{p}_i = mc^2 v^j \partial_i \left[\frac{1-2\Gamma}{1-\Gamma} \partial_j \Gamma \right] + mc^2 \partial_i v^j \left[\frac{1-2\Gamma}{1-\Gamma} \partial_j \Gamma \right] \tag{64}$$

or

$$m\ddot{v}_i = mc^2 v^j \left[\frac{1-2\Gamma}{1-\Gamma} \partial_i \partial_j \Gamma - \frac{\partial_i \Gamma \partial_j \Gamma}{(1-\Gamma)^2} \right] + mc^2 \partial_i v^j \left[\frac{1-2\Gamma}{1-\Gamma} \partial_j \Gamma \right] \tag{65}$$

This results in:

$$\frac{\ddot{v}_i}{c^2} = v^j \frac{1}{(1-\Gamma)^2}[(1 - 3\Gamma + 2\Gamma^2)\partial_i \partial_j \Gamma - \partial_i \Gamma \partial_j \Gamma] + \partial_i v^j \left[\frac{1-2\Gamma}{1-\Gamma} \partial_j \Gamma \right] \tag{66}$$

Equations of motion in polar coordinates r, φ

Let us now evaluate the equation of motion (66) of a body in the central static gravitational field Γ , with $\Gamma = \frac{GM}{c^2 r}$ and $\frac{\partial \Gamma}{\partial \varphi} = 0$ in the polar coordinates r, φ.

In the reference frame $\vec{e}$, $\vec{n}$, $\vec{n} = \frac{d\vec{e}}{d\varphi}$ and $\vec{e} = -\frac{d\vec{n}}{d\varphi}$. Thus, $\dot{\vec{e}} = \vec{n}\dot{\varphi}$ and $\dot{\vec{n}} = -\vec{e}\dot{\varphi}$ The velocity is given by $\vec{v} = \dot{r}\vec{e} + r\dot{\varphi}\vec{n}$. The gradient of a scalar f is $\vec{\partial} f = \frac{\partial f}{\partial r}\vec{e} + 1/r \frac{\partial f}{\partial \varphi}\vec{n}$.

Let us first evaluate the last term $\partial_i v^j \left[\frac{1-2\Gamma}{1-\Gamma} \partial_j \Gamma \right]$ in the equation 66, so that we have

$$\partial_i v^j = \partial_r(\dot{r}\vec{e_j} + r\dot{\varphi}\vec{n_j})\vec{e_i} + \frac{1}{r}\partial_\varphi(\dot{r}\vec{e_j} + r\dot{\varphi}\vec{n_j})\vec{n_i} = \dot{\varphi}\vec{n_j}\vec{e_i} \tag{67}$$

and

$$\partial_j \Gamma = \partial_r \Gamma \vec{e_j} + \frac{1}{r}\partial_\varphi \Gamma \vec{n_j} = -\frac{GM}{c^2 r^2}\vec{e_j}. \tag{68}$$

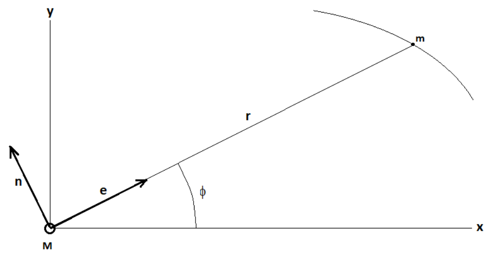


Figure 7. polar coordinates

This term contains the factor $\vec{e_j}\vec{n_j}\vec{e_i} = 0$ since $(\vec{e}|\vec{n}) = 0$. Thus, when developing eq. 66 we are left with two terms: Term I $= \underbrace{(1 - 3\Gamma + 2\Gamma^2)v^j\partial_j\partial_i\Gamma)}_{I}$ and term II $= \underbrace{(-v^j\partial_j\Gamma\partial_i\Gamma)}_{II}$.

Developing them we have (we leave the arrow on top of e and n for easier identification)

Term I: The calculation of $v^j\partial_j\partial_i\Gamma$ gives rise to

$$v^j\partial_j\partial_i\Gamma = (\dot{r}\vec{e_j} + r\dot{\varphi}\vec{n_j} \mid \frac{2GM}{c^2r^3}\vec{e_j}\vec{e_i} - \frac{GM}{c^2r^3}\vec{n_j}\vec{n_i}). \tag{69}$$

Indeed:

$$\partial_j\partial_i\Gamma = \frac{\partial(\partial_i\Gamma)}{\partial r}\vec{e_j} + 1/r\frac{\partial(\partial_i\Gamma)}{\partial\varphi}\vec{n_j} = \partial_r(\frac{-GM}{c^2r^2}\vec{e_i})\vec{e_j} + 1/r\,\partial_\varphi(\frac{-GM}{c^2r^2}\vec{e_i})\vec{n_j} \tag{70}$$

$$= \frac{2GM}{c^2r^3}\vec{e_i}\vec{e_j} - \frac{GM}{c^2r^2}(\partial_r\vec{e_i})\vec{e_j} + 1/r\left(\partial_\varphi(\frac{-GM}{c^2r^2})\vec{e_i}\vec{n_j} - \frac{GM}{c^2r^2}(\partial_\varphi\vec{e_i})\vec{n_j}\right)$$

$$= \frac{2GM}{c^2r^3}\vec{e_i}\vec{e_j} - \frac{GM}{c^2r^3}\vec{n_i}\vec{n_j} \tag{71}$$

since $\partial_i\Gamma = \partial_r\Gamma\vec{e_i} + \frac{1}{r}\partial_\varphi\Gamma\vec{n_i}$, $\partial_r\Gamma = \frac{-GM}{c^2r^2}$ and $\partial_\varphi\Gamma = 0$.

Then eq. 69 becomes:

$$v^j\partial_j\partial_i\Gamma = \dot{r}\frac{2GM}{c^2r^3}(\vec{e_j}|\vec{e_j})\vec{e_i} - \dot{r}\frac{GM}{c^2r^3}(\vec{e_j}|\vec{n_j})\vec{n_i} + r\dot{\varphi}\frac{2GM}{c^2r^3}(\vec{n_j}|\vec{e_j})\vec{e_i} - r\dot{\varphi}\frac{GM}{c^2r^3}(\vec{n_j}|\vec{n_j})\vec{n_i} \tag{72}$$

Since $(\vec{e}|\vec{e}) = (\vec{n}|\vec{n}) = 1$ and $(\vec{e}|\vec{n}) = 0$ we get:

$$v^j\partial_j\partial_i\Gamma = \frac{2\dot{r}GM}{c^2r^3}\vec{e_i} - \dot{\varphi}\frac{GM}{c^2r^2}\vec{n_i}$$

And term I becomes (with $\Gamma = \frac{GM}{c^2r}$)

$$(1 - 3\Gamma + 2\Gamma^2)\,v^j\partial_j\partial_i\Gamma$$

$$= (2\dot{r}\frac{GM}{c^2r^3} - 3a\frac{GM}{c^2r}\frac{2\dot{r}GM}{c^2r^3} + 2\frac{G^2M^2}{c^4r^2}\frac{2\dot{r}GM}{c^2r^3})\vec{e_i}$$

$$+ (-\dot{\varphi}\frac{GM}{c^2r^2} + 3\frac{GM}{c^2r}\dot{\varphi}\frac{GM}{c^2r^2} - 2\frac{G^2M^2}{c^4r^2}\dot{\varphi}\frac{GM}{c^2r^2})\vec{n_i}. \tag{73}$$

Term II:

$$(-v^j\partial_j\Gamma\partial_i\Gamma) = -(\dot{r}\vec{e_j} + r\dot{\varphi}\vec{n_j} \mid \frac{-GM}{c^2r^2}\vec{e_j})\frac{-GM}{c^2r^2}\vec{e_i} = -\dot{r}\frac{G^2M^2}{c^4r^4}\vec{e_i}. \tag{74}$$

Collecting both terms $I + II$ we get:

$$\frac{\ddot{\vec{v}}}{c^2} = \frac{1}{(1-\frac{GM}{c^2r})^2}\left[\left(2\dot{r}\frac{GM}{c^2r^3} - 7\dot{r}\frac{G^2M^2}{c^4r^4} + 4\dot{r}\frac{G^3M^3}{c^6r^5}\right)\vec{e} + \left(-\dot{\varphi}\frac{GM}{c^2r^2} + 3\dot{\varphi}\frac{G^2M^2}{c^4r^3} - 2\dot{\varphi}\frac{G^3M^3}{c^6r^4}\right)\vec{n}\right] \tag{75}$$

$$\frac{\ddot{\vec{v}}}{c^2} = \frac{1}{(1-\frac{GM}{c^2r})^2}\left[\left(-\frac{\dot{GM}}{c^2r^2} + 7/3\frac{\dot{G^2M^2}}{c^4r^3} - \frac{\dot{G^3M^3}}{c^6r^4}\right)\vec{e}\right. \tag{76}$$

$$\left. + \left(-\frac{GM}{c^2r^2} + 3\frac{G^2M^2}{c^4r^3} - 2\frac{G^3M^3}{c^6r^4}\right)\dot{\varphi}\vec{n}\right]. \tag{77}$$

Since $\dot{\varphi}\vec{n} = \dot{\vec{e}}$ we get (we postpone terms in $\vec{e}$ for the subsequent steps)

$$\frac{\ddot{\vec{v}}}{c^2} = \frac{1}{\left(1-\frac{GM}{c^2r}\right)^2}\left[\left(\frac{-\dot{GM}}{c^2r^2} + \frac{9}{3}\frac{\dot{G^2M^2}}{c^4r^3} - \frac{2}{3}\frac{\dot{G^2M^2}}{c^4r^3} - \frac{\dot{2G^3M^3}}{c^6r^4} + \frac{\dot{G^3M^3}}{c^6r^4}\right)\vec{e}\right. \tag{78}$$

$$\left. + \left(\frac{-GM}{c^2r^2} + 3\frac{G^2M^2}{c^4r^3} - 2\frac{G^3M^3}{c^6r^4}\right)\dot{\vec{e}}\right]. \tag{79}$$

$$\frac{\ddot{\vec{v}}}{c^2} = \frac{1}{\left(1-\frac{GM}{c^2r}\right)^2}\frac{d}{dt}\left[\left(\frac{-GM}{c^2r^2} + 3\frac{G^2M^2}{c^4r^3} - 2\frac{G^3M^3}{c^6r^4}\right)\vec{e}\right] \tag{80}$$

$$+\frac{1}{\left(1-\frac{GM}{c^2r}\right)^2}(-2/3\frac{\dot{G^2M^2}}{c^4r^3} + \frac{\dot{G^3M^3}}{c^6r^4})\vec{e}. \tag{81}$$

The last term in 81 can be shown to be equal to:

$$\left(-\frac{GM}{c^2r^2} + 3\frac{G^2M^2}{c^4r^3} - 2\frac{G^3M^3}{c^6r^4}\right)\vec{e}\frac{d}{dt}\frac{1}{\left(1-\frac{GM}{c^2r}\right)^2}$$

Thus we finally get:

$$\frac{\ddot{\vec{v}}}{c^2} = \frac{d}{dt}\left[\frac{1}{\left(1-\frac{GM}{c^2r}\right)^2}\left(\frac{-GM}{c^2r^2} + 3\frac{G^2M^2}{c^4r^3} - 2\frac{G^3M^3}{c^6r^4}\right)\vec{e}\right]. \tag{82}$$

Integrating on time we get the gravitational acceleration (the integration constant can be set to 0 in a suitable reference frame)

$$\frac{\dot{\vec{v}}}{c^2} = \frac{1}{\left(1-\frac{GM}{c^2r}\right)^2}\left(-\frac{GM}{c^2r^2} + 3\frac{G^2M^2}{c^4r^3} - 2\frac{G^3M^3}{c^6r^4}\right)\vec{e}. \tag{83}$$

To the first order in G we find **Newton's law**:

$$\vec{F} = m\dot{\vec{v}} = \frac{-GmM}{r^2}\vec{e} \tag{84}$$

and also the **Equivalence principle** stating that the effects of a gravitational field are identical to an acceleration given by eq. (83) .

It is interesting to note that the second term in (83) is a **repulsive acceleration** $3\frac{G^2M^2}{c^4r^3}$. It is usually very small compared to the first term. For instance at the surface of earth, with: $M = 6\,10^{24}\,kg, r = 6,37\,10^6 m, G = 6,67\,10^{-11}\,m^3/kgs^2, c = 3\,10^8 m/s$, the ratio of the repulsive force to the main attractive force is $\frac{3GM}{c^2r} \simeq 2\,10^{-9}$. Could this repulsive force be measured ? Possibly by comparing the velocities of satellites at different altitudes. However, it remains open right now.

Solution in the polar coordinates r, φ

The radial $\vec{e}$ component of the acceleration $\dot{\vec{v}}$ is:

$$\frac{d^2r}{dt^2} - r(\frac{d\varphi}{dt})^2$$

And its tangential $\vec{n}$ component is:

$$r\frac{d^2\varphi}{dt^2} + 2\frac{dr}{dt}\frac{d\varphi}{dt} = \frac{1}{r}\frac{d}{dt}(r^2\dot{\varphi}).$$

We thus get the following set of equations:

$$\frac{d^2r}{dt^2} - r\left(\frac{d\varphi}{dt}\right)^2 =$$

$$c^2\left[\left(1 - \frac{GM}{c^2r}\right)^{-2}\left(-\frac{GM}{c^2r^2} + 3\frac{G^2M^2}{c^4r^3} - 2\frac{G^3M^3}{c^6r^4}\right)\right] \tag{85}$$

and

$$\frac{1}{r}\frac{d}{dt}(r^2\dot{\varphi}) = 0. \tag{86}$$

Eq. (86) expresses the conservation of angular momentum and gives $r^2\dot{\varphi} = h$ with h a constant (m^2/s).

Regarding eq. (85), we must first develop $\frac{d^2r}{dt^2}$ as follows

$$\frac{dr}{dt} = \frac{dr}{d\varphi}\frac{d\varphi}{dt} \rightarrow \dot{r} = \frac{h}{r^2}\frac{dr}{d\varphi}$$

$$\ddot{r} = \frac{d\dot{r}}{d\varphi}\frac{d\varphi}{dt} = \frac{h^2}{r^4}\frac{d^2r}{d\varphi^2} - 2\frac{h^2}{r^5}(\frac{dr}{d\varphi})^2. \tag{87}$$

Eq (85) becomes

$$\frac{h^2}{r^4}\frac{d^2r}{d\varphi^2} - 2\frac{h^2}{r^5}(\frac{dr}{d\varphi})^2 - \frac{h^2}{r^3} - \left(1 - \frac{GM}{c^2r}\right)^{-2}\left(-\frac{GM}{r^2} + 3\frac{G^2M^2}{c^2r^3} - 2\frac{G^3M^3}{c^4r^4}\right) = 0. \tag{88}$$

And after multiplying by r^2:

$$\frac{h^2}{r^2}\frac{d^2r}{d\varphi^2} - 2\frac{h^2}{r^3}(\frac{dr}{d\varphi})^2 - \frac{h^2}{r} - \left(1 - \frac{GM}{c^2r}\right)^{-2}\left(-GM + 3\frac{G^2M^2}{c^2r} - 2\frac{G^3M^3}{c^4r^2}\right) = 0. \tag{89}$$

Let $u = \frac{1}{r}$ then $\frac{du}{d\varphi} = -\frac{1}{r^2}\frac{dr}{d\varphi}$ and $\frac{d^2u}{d\varphi^2} = -\frac{1}{r^2}\frac{d^2r}{d\varphi^2} + \frac{2}{r^3}(\frac{dr}{d\varphi})^2$. So that we get:

$$-h^2\frac{d^2u}{d\varphi^2} - h^2u + \left(1 - \frac{GMu}{c^2}\right)^{-2}\left(GM - \frac{3G^2M^2u}{c^2} + \frac{2G^3M^3u^2}{c^4}\right) = 0. \tag{90}$$

Or

$$-h^2\frac{d^2u}{d\varphi^2} - h^2u + GM\left(\frac{1 - \frac{2GMu}{c^2} + \frac{G^2M^2u^2}{c^4} - \frac{GMu}{c^2} + \frac{G^2M^2u^2}{c^4}}{1 - \frac{2GMu}{c^2} + \frac{G^2M^2u^2}{c^4}}\right) = 0$$

Dividing by h^2 and writing $\frac{du}{d\varphi} \equiv u'$ and rearranging, we get

$$u'' + u - \frac{GM}{h^2}\left(1 - \frac{\frac{GMu}{c^2}}{1 - \frac{GMu}{c^2}}\right) = 0$$

Or

$$u'' + u\left(1 + \frac{G^2M^2}{c^2h^2}\left(1 - \frac{GMu}{c^2}\right)^{-1}\right) = \frac{GM}{h^2}. \tag{91}$$

Eq. (91) is the equation of the orbit of a body around a stationary body of a mass M for non-relativistic speed.

Relativistic equations of motion of a body in a central static field

We rewrite eq. (63) where we will use the relativistic impulsion $\vec{p}$ and p_0:

$$\ddot{p}_i = \partial_i\left[p_0v^j\frac{1-2\Gamma}{1-\Gamma}\partial_j\Gamma\right]. \tag{92}$$

Which with the relativistic impulsion of a body according to Landau and Lifschitz 1964, gives

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{93}$$

and

$$p_0 = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{94}$$

Then, one gets:

$$\ddot{p}_i = p_0v^j\partial_i\left[\frac{1-2\Gamma}{1-\Gamma}\partial_j\Gamma\right] + (\frac{1-2\Gamma}{1-\Gamma}\partial_j\Gamma)v^j\partial_ip_0 \tag{95}$$

The first term of eq. (95) is:

$$\frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}} v^j \partial_i \left[\frac{1-2\Gamma}{1-\Gamma}\partial_j \Gamma\right] = \frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}} v^j \frac{1}{(1-\Gamma)^2}\left[(1-3\Gamma+2\Gamma^2)\partial_i\partial_j\Gamma - \partial_i\Gamma\partial_j\Gamma\right] \tag{96}$$

$$= \frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}} \frac{d}{dt}\left[\frac{1}{\left(1-\frac{GM}{c^2 r}\right)^2}\left(\frac{-GM}{c^2r^2} + 3\frac{G^2M^2}{c^4r^3} - 2\frac{G^3M^3}{c^6r^4}\right)\vec{e}\right] \tag{97}$$

which is obtained in the same way as eq. (82) was obtained from eq. (66).

The second term of eq. (95) is

$$(\frac{1-2\Gamma}{1-\Gamma}\partial_j\Gamma)v^j\partial_i p_0 = (\frac{1-2\Gamma}{1-\Gamma})\partial_j\Gamma\, v^j mc^2\, \partial_i \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$$

Which in polar coordinates reads

$$\partial_j \Gamma\, v^j = (\dot{r}\vec{e_j} + r\dot{\varphi}\vec{n_j} | \partial_r(\frac{GM}{c^2 r})\vec{e_j} + \frac{1}{r}\partial_\varphi \Gamma \vec{n_j}) = -\dot{r}\frac{GM}{c^2r^2}$$

since $\partial_\varphi \Gamma = 0$ and $\partial_i \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{1}{2c^2}(1-\frac{v^2}{c^2})^{-3/2}\partial_i(-v^2)$.

But $\partial_i(v^2) = \partial_r(\dot{r}^2 + r^2\dot{\varphi}^2)\vec{e_i} + \frac{1}{r}\partial_\varphi(\dot{r}^2 + r^2\dot{\varphi}^2)\vec{n_i}$ and $= 2r\dot{\varphi}^2\vec{e_i}$ since $\partial_r(\dot{r}^2) = 0 = \partial_\varphi(\dot{\varphi}^2)$ which after combining gives the second term as

$$-mc^2(1-\frac{v^2}{c^2})^{-3/2}\frac{1-2\Gamma}{1-\Gamma}\frac{1}{c^2}r\dot{r}\dot{\varphi}^2\frac{GM}{c^2r^2}\vec{e_i}$$

$$= -mc^2(1-\frac{v^2}{c^2})^{-3/2}\frac{1-3\Gamma+2\Gamma^2}{(1-\Gamma)^2}\frac{1}{c^2}\dot{r}\dot{\varphi}^2\Gamma\vec{e_i}$$

$$= -\frac{1}{(1-\frac{GM}{c^2r})^2}(\frac{GM}{c^2r} - 3\frac{G^2M^2}{c^4r^2} + 2\frac{G^3M^3}{c^6r^3})\frac{1}{c^2}mc^2(1-\frac{v^2}{c^2})^{-3/2}\dot{r}\dot{\varphi}^2\vec{e_i}$$

$$= \frac{1}{(1-\frac{GM}{c^2r})^2}(\frac{-GM}{c^2r^2} + 3\frac{G^2M^2}{c^4r^3} - 2\frac{G^3M^3}{c^6r^4})\frac{1}{c^2}mc^2(1-\frac{v^2}{c^2})^{-3/2}r\dot{r}\dot{\varphi}^2\vec{e_i}.$$

and since $\partial_i(v^2) = 2r\dot{\varphi}^2\vec{e_i}$, we have: $\dot{r}2r\dot{\varphi}^2\vec{e_i} = \dot{r}\partial_i(v^2)\vec{e_i} = \dot{r}\partial_r(v^2)\vec{e_i} = \frac{dv^2}{dt}\vec{e_i}$ from which it follows that $\frac{1}{c^2}mc^2(1-\frac{v^2}{c^2})^{-3/2}r\dot{r}\dot{\varphi}^2\vec{e_i} = \frac{d}{dt}\frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}}\vec{e_i}$. Combining the two terms, we get

$$\ddot{\vec{p}} = \frac{d}{dt}\left[\frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}}\frac{1}{\left(1-\frac{GM}{c^2r}\right)^2}\left(\frac{-GM}{c^2r^2} + 3\frac{G^2M^2}{c^4r^3} - 2\frac{G^3M^3}{c^6r^4}\right)\vec{e}\right].$$

Then integrating on time gives

$$\dot{\vec{p}} = \frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}}\frac{1}{\left(1-\frac{GM}{c^2r}\right)^2}\left(\frac{-GM}{c^2r^2} + 3\frac{G^2M^2}{c^4r^3} - 2\frac{G^3M^3}{c^6r^4}\right)\vec{e}. \tag{98}$$

This can be compared with eq. (83). Let us look at the left hand term of that equation where there appears the time derivative of the relativistic impulsion. From the definition (93) we evaluate the time derivative of the relativistic impulsion as follows.

$$\frac{d}{dt}\vec{p} = m\left(\frac{\dot{\vec{v}}}{\sqrt{1-\frac{v^2}{c^2}}} + \frac{(\vec{v}|\dot{\vec{v}})\,\vec{v}}{c^2(1-\frac{v^2}{c^2})^{3/2}}\right). \tag{99}$$

Note that according to Landau and Lifschitz 1964, when the force is normal to speed we would have

$$\frac{d}{dt}\vec{p} = m\left(\frac{\dot{\vec{v}}}{\sqrt{1-\frac{v^2}{c^2}}}\right) \tag{100}$$

and when the force and speed are co-linear we would get

$$\frac{d}{dt}\vec{p} = m\left(\frac{\dot{\vec{v}}}{(1-\frac{v^2}{c^2})^{3/2}}\right). \tag{101}$$

In the following, since forces and speeds can have independent orientations we have to use the general formula as in eq. (99) for the time derivative of impulsion. Consequently in polar coordinates we obtain

$$\dot{\vec{v}} = (\ddot{r} - r\dot{\varphi}^2)\vec{e} + \frac{1}{r}(r^2\dot{\varphi})^{\cdot}\,\vec{n}$$

and

$$(\vec{v}|\dot{\vec{v}}) = \frac{1}{2}\dot{v^2} = \frac{1}{2}(\dot{r}^2 + r^2\dot{\varphi}^2)^{\cdot}.$$

Then eqs. 98 and 99 result in the following cases.

Along $\vec{n}$

$$\frac{\frac{1}{r}(r^2\dot{\varphi})^{\cdot}}{\sqrt{1-\frac{v^2}{c^2}}} + \frac{r\dot{\varphi}(\dot{v^2})}{2c^2(1-\frac{v^2}{c^2})^{3/2}} = 0 \tag{102}$$

and along $\vec{e}$ (dividing both by m)

$$\frac{\ddot{r} - r\dot{\varphi}^2}{\sqrt{1-\frac{v^2}{c^2}}} + \frac{\dot{r}\,\dot{v^2}}{2c^2(1-\frac{v^2}{c^2})^{3/2}} = \frac{c^2}{\sqrt{1-\frac{v^2}{c^2}}}\frac{1}{\left(1-\frac{GM}{c^2r}\right)^2}\left(\frac{-GM}{c^2r^2} + 3\frac{G^2M^2}{c^4r^3} - 2\frac{G^3M^3}{c^6r^4}\right). \tag{103}$$

Eq. (102) represents the relativistic conservation of angular momentum, after the following development. Multiply both by $\sqrt{1-\frac{v^2}{c^2}}$ so that $\Rightarrow (r^2\dot{\varphi})^{\cdot} = -r^2\,\dot{\varphi}\,\frac{\dot{v^2}}{2(c^2-v^2)}$. Thus,

$$\frac{(r^2\dot{\varphi})^{\cdot}}{r^2\dot{\varphi}} = -\frac{\dot{v^2}}{2(c^2-v^2)} = \frac{1}{2}\frac{(c^2-v^2)^{\cdot}}{(c^2-v^2)}. \tag{104}$$

Integrating on time gives (the units are again more physical in this calculation)

$$\ln(r^2\dot{\varphi}) = \frac{1}{2}\ln(c^2-v^2) + \kappa \tag{105}$$

with κ a constant, and after exponentiation and defining $e^{\kappa} \equiv h/c$ we obtain

$$r^2\dot{\varphi} = \sqrt{(c^2 - v^2)}\, e^{\kappa} = \sqrt{(c^2 - v^2)}\, h/c = h\sqrt{1 - \frac{v^2}{c^2}}. \tag{106}$$

For $v \ll c$ one gets the classical form $r^2\dot{\varphi} = h$ however for the calculation of the precession of the perihelion of Mercury we will use the relativistic equation eq. 106.

Let us now develop eq. 103 such that it will give us the relativistic relation for the orbit of a body around a stationary mass M. After multiplication of both terms of (103) by $\sqrt{1 - \frac{v^2}{c^2}}$ we get:

$$\ddot{r} - r\dot{\varphi}^2 + \frac{\dot{r}(\dot{v^2})}{2(c^2 - v^2)} = c^2 \frac{1}{\left(1 - \frac{GM}{c^2 r}\right)^2} \left(\frac{-GM}{c^2 r^2} + 3\frac{G^2M^2}{c^4 r^3} - 2\frac{G^3M^3}{c^6 r^4}\right). \tag{107}$$

We now expand the first term of (107) so that r depends on φ, using the relations $\dot{\varphi} = \frac{h}{r^2}\sqrt{1 - \frac{v^2}{c^2}}$ and $\dot{r} = \frac{dr}{d\varphi}\dot{\varphi} = r'\frac{h}{r^2}\sqrt{1 - \frac{v^2}{c^2}}$ Then, let us calculate $\ddot{r}$.

$$\ddot{r} = \frac{d\dot{r}}{d\varphi}\dot{\varphi} = \frac{h}{r^2}\sqrt{1 - \frac{v^2}{c^2}} \left(\frac{d}{d\varphi}(\frac{h}{r^2}\sqrt{1 - \frac{v^2}{c^2}}\frac{dr}{d\varphi})\right) =$$

$$\frac{h^2}{r^2}\sqrt{1 - \frac{v^2}{c^2}} \left(\frac{-2}{r^3} r'^2 \sqrt{1 - \frac{v^2}{c^2}} + r'\frac{1}{r^2}\frac{d}{d\varphi}\sqrt{1 - \frac{v^2}{c^2}} + \frac{\sqrt{1 - \frac{v^2}{c^2}}}{r^2} r''\right)$$

$$\ddot{r} = \frac{h^2}{r^2}\left(\frac{-2}{r^3} r'^2 (1 - \frac{v^2}{c^2}) + \frac{r'}{r^2}\sqrt{1 - \frac{v^2}{c^2}}\frac{d}{d\varphi}\sqrt{1 - \frac{v^2}{c^2}} + \frac{1 - \frac{v^2}{c^2}}{r^2} r''\right) \tag{108}$$

Now $\sqrt{1 - \frac{v^2}{c^2}}\frac{d}{d\varphi}\sqrt{1 - \frac{v^2}{c^2}} = 1/2\frac{d}{d\varphi}(1 - \frac{v^2}{c^2}) = 1/2\frac{d}{d\varphi}(-\frac{v^2}{c^2})$ and $\frac{d}{d\varphi}v^2 = \dot{v^2}\frac{dt}{d\varphi} = \dot{v^2}\frac{r^2}{h\sqrt{1 - \frac{v^2}{c^2}}}$.

After replacing by the above in (108) we get:

$$\ddot{r} = (1 - \frac{v^2}{c^2})(\frac{-2h^2 r'^2}{r^5}) - \frac{hr'\dot{v^2}}{2r^2c^2\sqrt{1 - \frac{v^2}{c^2}}} + (1 - \frac{v^2}{c^2})\frac{h^2 r''}{r^4}. \tag{109}$$

The two other terms in (107) are

$$-r\dot{\varphi^2} = \frac{-h^2}{r^3}(1 - \frac{v^2}{c^2}) \tag{110}$$

and

$$\frac{\dot{r}\dot{v^2}}{2c^2(1 - \frac{v^2}{c^2})} = \frac{hr'\dot{v^2}}{2r^2c^2\sqrt{1 - \frac{v^2}{c^2}}} \tag{111}$$

using $\dot{r} = \frac{dr}{d\varphi}\frac{d\varphi}{dt}$ and $\dot{\varphi} = \frac{h}{r^2\sqrt{1 - \frac{v^2}{c^2}}}$.

So the first term of (107) $\ddot{r} - r\dot{\varphi}^2 + \frac{\dot{r}\dot{(v^2)}}{2(c^2-v^2)}$ becomes:

$$(1-\frac{v^2}{c^2})(\frac{-2h^2r'^2}{r^5}+\frac{h^2r''}{r^4}-\frac{h^2}{r^3}). \tag{112}$$

Dividing both terms of (107) by $(1-\frac{v^2}{c^2})$ we get

$$\frac{-2h^2r'^2}{r^5}+\frac{h^2r''}{r^4}-\frac{h^2}{r^3}=\frac{1}{(1-\frac{v^2}{c^2})\left(1-\frac{GM}{c^2r}\right)^2}\left(\frac{-GM}{r^2}+3\frac{G^2M^2}{c^2r^3}-2\frac{G^3M^3}{c^4r^4}\right). \tag{113}$$

We now apply the same analysis to (113) as we did for eq. 89. After multiplying in (113) by r^2 we have

$$\frac{-2h^2r'^2}{r^3}+\frac{h^2r''}{r^2}-\frac{h^2}{r}-\frac{1}{(1-\frac{v^2}{c^2})\left(1-\frac{GM}{c^2r}\right)^2}\left(-GM+3\frac{G^2M^2}{c^2r}-2\frac{G^3M^3}{c^4r^2}\right)=0. \tag{114}$$

Let $u=\frac{1}{r}$ then $\frac{du}{d\varphi}=-\frac{1}{r^2}\frac{dr}{d\varphi}$ and $\frac{d^2u}{d\varphi^2}=-\frac{1}{r^2}\frac{d^2r}{d\varphi^2}+\frac{2}{r^3}(\frac{dr}{d\varphi})^2$. So, we get the following differential equation in u

$$-h^2\frac{d^2u}{d\varphi^2}-h^2u+\left(1-\frac{GMu}{c^2}\right)^{-2}(1-\frac{v^2}{c^2})^{-1}\left(GM-\frac{3G^2M^2u}{c^2}+\frac{2G^3M^3u^2}{c^4}\right)=0 \tag{115}$$

or

$$-h^2\frac{d^2u}{d\varphi^2}-h^2u+(1-\frac{v^2}{c^2})^{-1}GM\left(\frac{1-\frac{2GMu}{c^2}+\frac{G^2M^2u^2}{c^4}-\frac{GMu}{c^2}+\frac{G^2M^2u^2}{c^4}}{1-\frac{2GMu}{c^2}+\frac{G^2M^2u^2}{c^4}}\right)=0.$$

Dividing by h^2, writing $\frac{du}{d\varphi}\equiv u'$ and rearranging we get

$$u''+u-\frac{GM}{h^2(1-\frac{v^2}{c^2})}\left(1-\frac{\frac{GMu}{c^2}}{1-\frac{GMu}{c^2}}\right)=0 \tag{116}$$

or else

$$u''+u\left(1+\frac{G^2M^2}{c^2h^2(1-\frac{v^2}{c^2})(1-\frac{GMu}{c^2})}\right)=\frac{GM}{h^2(1-\frac{v^2}{c^2})}. \tag{117}$$

This is the polar equation of the orbit of a body (a planet) around a stationary body of mass M for relativistic speed. Compared with eq. (91), it contains a relativistic factor $(1-\frac{v^2}{c^2})$ which has some dependance on u. We try to solve this differential equation.

Let us find a relation between $(1-\frac{v^2}{c^2})$ and u. Observe that both $1/(1-\frac{v^2}{c^2})$ and u are periodic in φ on a classical non relativistic orbit, see Figs. 8 and 10. We define the relation between $1/(1-\frac{v^2}{c^2})$ and u as follows.

$$\frac{1}{(1-\frac{v^2}{c^2})}=A\frac{u}{(\frac{GM}{h^2})}+B \tag{118}$$

where u is dimensionless via the factor $1/\frac{GM}{h^2}$. A and B can be derived from the values of $1/(1-\frac{v^2}{c^2})$ and u at the perihelion and aphelion of the planet, the result is, with e being the eccentricity A= $\frac{2G^2M^2}{c^2h^2}$ and B= $1-\frac{G^2M^2(1-e^2)}{c^2h^2}$.

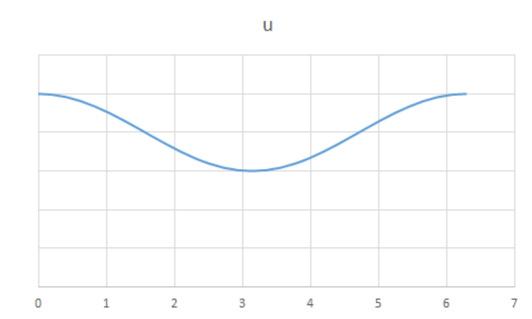


Figure 8. $u(\varphi)$

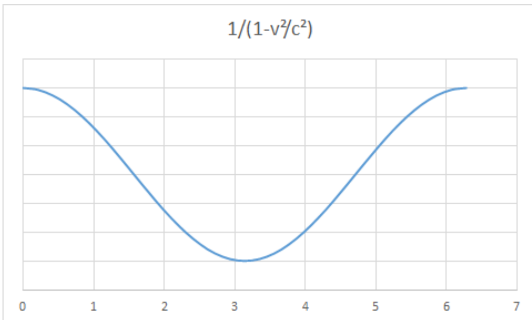


Figure 9. $1/(1-\frac{v^2}{c^2})(\varphi)$

$1/(1-\frac{v^2}{c^2})$ can be written as

$$\frac{1}{(1-\frac{v^2}{c^2})} = (\frac{2GMu}{c^2} + 1 - \frac{G^2M^2(1-e^2)}{c^2h^2}) \tag{119}$$

For non relativistic speed $v << c$ it holds $\frac{1}{(1-\frac{v^2}{c^2})} \approx (1+\frac{v^2}{c^2})$ and we get the conservation of energy: $\frac{mv^2}{2} = \frac{mGM}{r} - \frac{G^2M^2(1-e^2)m}{2h^2}$, i.e., the sum of kinetic and potentail energy is a constant.

Then, the eq. 117 becomes

$$u'' + u\left(1 + \frac{G^2M^2}{c^2h^2(1-\frac{GMu}{c^2})}(\frac{2GMu}{c^2} + 1 - \frac{G^2M^2(1-e^2)}{c^2h^2})\right) = \frac{2G^2M^2u}{c^2h^2} + \frac{GM}{h^2} - \frac{G^3M^3(1-e^2)}{c^2h^4} \tag{120}$$

8. The Schwarzschild metric

In order to calculate the precession of the perihelion of Mercury, we need to express eq. 120 in the Schwarzschild metric of the proper coordinates of Mercury and not in the Minkowskian metric used up to now that corresponds to an observer situated at infinite distance from the massive body and where Γ tends to zero.

In the spherical coordinates a gravitational field can be written to the first order as $\Gamma = \frac{GM}{rc^2}$ (36) and the Schwartszchild metric is expressed as $d\tau^2 = (1-2\Gamma)dt^2 - (1-2\Gamma)^{-1}dr^2 - r^2d\theta^2 - r^2\sin\theta^2 d\phi^2$.

Which gives for $\Gamma << 1$ in a Minkowskian metric: $d\tau^2 = dt'^2 - dr'^2 - r^2 d\theta^2 - r^2 \sin\theta^2 d\phi^2$ with the change of variables $dt' = (1 - \Gamma)dt$, $dr' = (1 + \Gamma)dr$.

Then $\ddot{r}$ in eq. 107 becomes $\frac{1-\Gamma}{1+2\Gamma}\frac{d^2r'}{dt'^2} = (1 - 3\Gamma)\frac{d^2r'}{dt'^2}$. The $(1 - 3\Gamma)$ factor applies to the first term of (107) and eq. 115 becomes

$$(1 - 3\Gamma)(u'' + u) - (1 + 4\Gamma)\frac{GM}{h^2}(1 - 3\Gamma + 2\Gamma^2) = 0. \tag{121}$$

Recalling that $u = 1/r$, u'' is the second derivative of u by ϕ and $\Gamma = GMu/c^2$. We also need to consider the modification of the Laplacian in (36) that generates Γ; Δ becomes $(1 + 2\Gamma)\Delta$ and the term $GM(1 - 3\Gamma + 2\Gamma^2)$ must be multiplied by $(1 + 2\Gamma)\Delta$. Thus, the final form of eq. 115 is

$$(1 - 3\Gamma)(u'' + u) - (1 + 4\Gamma)\frac{GM}{h^2}(1 + 2\Gamma)(1 - 3\Gamma + 2\Gamma^2) = 0 \tag{122}$$

which can be simplified by eliminating the higher order terms

$$(u'' + u) - (1 + 6\Gamma)\frac{GM}{h^2} = 0. \tag{123}$$

Precession of the perihelion of Mercury

We now look for a solution of eq. 123 in the form $u = \mu + \alpha\cos(\beta\varphi)$ with $\alpha = \frac{eGM}{h^2} = \frac{(u_1 - u_2)}{2}$ and $\mu = \frac{(u_1+u_2)}{2}$. Eq. 123 can be written

$$u'' + u = (1 + 6\,GMu/c^2)\frac{GM}{h^2} \tag{124}$$

or else

$$u'' + u(1 - 6\,G^2M^2/h^2c^2) = \frac{GM}{h^2}. \tag{125}$$

Replacing u and solving for the term $\cos(\beta\varphi)$ we get

$$-\beta^2\alpha\cos(\beta\varphi) + (1 - 6\frac{G^2M^2}{c^2h^2})\alpha\cos(\beta\varphi) = 0. \tag{126}$$

Thus $\beta^2 = \left[1 - 6\frac{G^2M^2}{c^2h^2}\right]$ and $\beta \simeq 1 - 3\frac{G^2M^2}{c^2h^2}$ we get for the homogeneous solution

$$u = \alpha\cos\left[\left(1 - 3\frac{G^2M^2}{c^2h^2}\right)\varphi\right]. \tag{127}$$

What is the resulting advance of the perihelion? With $M = 2\,10^{30}\,kg$ the mass of the sun, $G = 6.67\,10^{-11} m^3/kgs^2$, $r = 1/u = 57.9\,10^9 m$ is the average distance of Mercury to the Sun, $e = 0.204$ for Mercury, $h = 2.7\,10^{15} m^2/s$ for Mercury, $c = 3\,10^8 m/s$. The period of revolution of Mercury = 88 days. We get a phase shift due to the $3\frac{G^2M^2}{c^2h^2}$ term. This shift is $\frac{6\pi G^2M^2}{c^2h^2} = 5.11\,10^{-7}$ radians per revolution. This corresponds to 43.2″ per century. The corresponding change in the position of the perihelion moves forward to the orbit of Mercury, the accepted value up to now is indeed 43″ in the same direction.

9. Black holes

Objects whose gravitational fields are too strong for light to escape were already considered in the 18th century by John Michell and Pierre-Simon Laplace. When described by general relativity, the black hole contains a gravitational singularity at the origin, a region where the spacetime curvature becomes infinite and contains all the mass of the black hole.

First we will limit ourselves to the study of black holes having mass M with no electric charge and no angular momentum. What happens to a body in the vicinity r of a black hole? Will it be swallowed and disappear forever?

Let r be the distance of this object to the centre of the black hole, which is supposed to be located at r=0 and to contain the mass M. We don't know yet if the large body of mass M is a black hole or not, at this stage it is just a homogeneous compact body with mass M "concentrated" at the origin. The radial extension of the massive body does not matter as long as it is smaller than or equal to r. This is so, because if r is smaller than the radial extension of the large body of mass M, then some amount of mass will not be taken into account when calculating the force of attraction at radius r. When assuming r to be the radial extension of the body of mass M, the following equations express the force acting on the surface of the body, i.e. at radius r. Eqs. 85 and 86 describe the motion of a body in a central static field. The gravitational acceleration is as in (83), so that

$$c^2\left(1-\frac{GM}{rc^2}\right)^{-2}\left(-\frac{GM}{r^2c^2}+3\frac{G^2M^2}{c^4r^3}-2\frac{G^3M^3}{c^6r^4}\right).$$

$F(r)$ is the radial force acting on a unit mass (unit $[N]$). Let us evaluate that force. The central mass M is supposed to be concentrated at the origin or at least on a radius smaller than r. We choose the unit system $\frac{GM}{c^2}=1$ for brevity and clarity. Then,

$$F(r)=(1-\frac{1}{r})^{-2}(-\frac{1}{r^2}+\frac{3}{r^3}-\frac{2}{r^4}) \tag{128}$$

has two singularities 1 at $r=0$ and 0 at r=2. Between r=0 and r= $\frac{GM}{c^2}$ the force is always attractive ($F<0$, Fig. 10). The region from r=0 to $r=\frac{GM}{c^2}$ is attractive, with infinite attraction for $r=0$ and

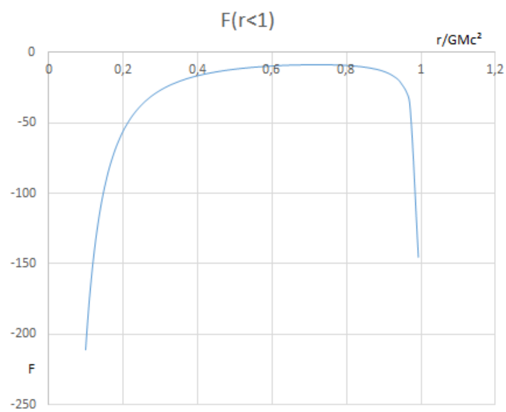


Figure 10. F for $r<GM/c^2$

$r=\frac{GM}{c^2}$. Between $r=\frac{GM}{c^2}$ and $r=2\frac{GM}{c^2}$ the force is repulsive ($F>0$) and is attractive again for $r>2\frac{GM}{c^2}$. The force then follows a $\frac{1}{r^2}$ law for large r. The point $2\frac{GM}{c^2}$ is stable in equilibrium with

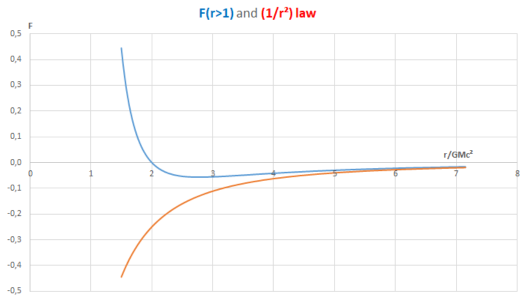


Figure 11. F for $r > GM/c^2$

zero force, and the corresponding radius is called the limit radius R_L. Note that it is equal to the Schwarzschild horizon in GR. *A black hole can never shrink to a null radius*: An infinite repulsive barrier at $r = \frac{GM}{c^2}$ prevents this collapse to happen and R_L represents the stable size of a non-rotating black hole. At R_L, the surface of the black hole is in the equilibrium and no force is acting on the surface. The force (83) derives from the following potential $\Phi(r)$, for $r > \frac{GM}{c^2}$ and set to zero at r_∞

$$\Phi(r) = c^2 \left(\frac{2GM}{rc^2} + \ln \left(1 - \frac{GM}{rc^2} \right) \right) [\frac{m^2}{s^2}]. \tag{129}$$

Which can be developed in the series:

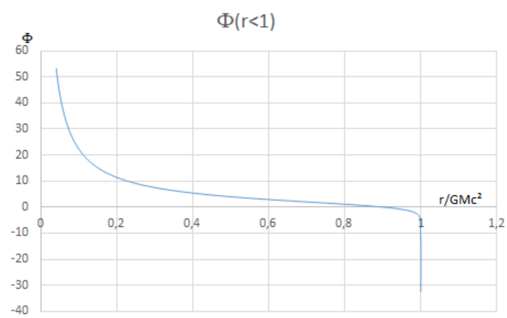


Figure 12. Φ for $r < GM/c^2$

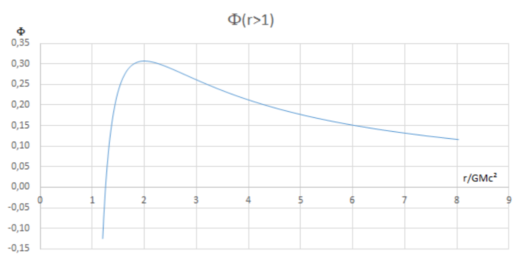


Figure 13. Φ for $r > GM/c^2$

$$\Phi(r) = \frac{GM}{r} - \frac{G^2M^2}{2r^2c^2} - \frac{G^3M^3}{3r^3c^4} - \ldots. \tag{130}$$

This series is the sum of an attractive potential $\frac{GM}{r}$ and a repulsive potential equal to $-\frac{G^2M^2}{2r^2c^2} - \frac{G^3M^3}{3r^3c^4} - \ldots$. The latter repulsive term could be seen as a graviton-graviton interaction term:Iit is negligible for large distance r but predominant at short distances corresponding to the Schwarzschild radius. It has a maximum of $c^2(1 + \ln(1/2)) \simeq 0.307c^2$ at R_L.

The escape of light from a "black hole"

In Fig. 14 we plot $\check{}\Phi$ for a more intuitive understanding. We follow a classical approach equaling potential and kinetic energies. We have for a unit mass m

$$1/2\, mv^2 = mc^2\left(\frac{2GM}{rc^2} + \ln\left(1 - \frac{GM}{rc^2}\right)\right).$$

Dividing by m and equalling v to c gives rise to

$$1/2 = \left(\frac{2GM}{rc^2} + \ln\left(1 - \frac{GM}{rc^2}\right)\right).$$

There is no solution to this equation, the kinetic energy is always higher than the potential well, so

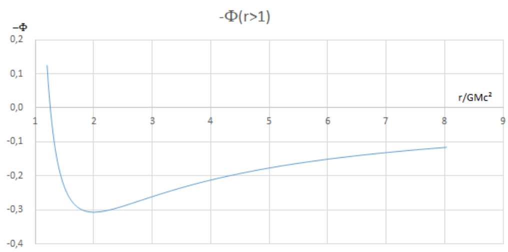


Figure 14. $-\Phi$ for $r > GM/c^2$

light can always escape from the "black hole". We have to find a new name for this class of bodies having a radius close or equal to R_L: They're not holes and they are not black either. Let us call them CORE. Quasars could be such cores presenting red shifts and pulsations, as will be shown in the next sections. The core radius of stability R_L is equal to the Schwarzschild radius of GR. The core sits there in stable equilibrium between expansion and contraction forces. If there was no repulsive terms, $-\frac{G^2M^2}{2r^2c^2} - \frac{G^3M^3}{3r^3c^4} - \ldots$, in the potential, then a solution would be possible where no light can escape at a horizon radius equal to the Schwarzschild radius and a gravitational singularity would appear.

10. Red shift of a core

The core is still considered as a non-rotating mass M. From (1) we obtain (by taking the potential Φ into account)

$$\nu_1\left(1 - \frac{\Phi_1}{c^2}\right) = \nu_2\left(1 - \frac{\Phi_2}{c^2}\right).$$

If the position 1 is at infinite distance $r = \infty$, then $\Phi_1 = 0$ At position 2, $r = R_L$ and $\Phi_2 = 0.307c^2$ and then $\nu_1 = \nu_2(1 - 0.307) = 0.693\,\nu_2$. The frequency of a spectral line emitted from a point at the surface R_L of a core will be perceived from an infinite distance at 0.693 times the frequency due to time slowdown at the surface of the core. But if we take into account the potential well of $0.307\,c^2$, the emitted frequency would decrease further by a factor of 0.614 since $E = h\nu$ and E is reduced by a factor $\frac{0.307c^2}{((c^2)/2)}$. The total frequency shift factor is the 0.693 times $(1 - 0.614)$ which is equal to 0.268.

11. Pulsation of a core

On the surface of the core, a mass is at equilibrium but can also oscillate radially around the equilibrium point R_L. Let us look at its first mode of oscillation: For a unit mass $m = 1$ on the surface of the core, the oscillation frequency ω is $\sqrt{\frac{k}{m}} = \sqrt{k}$, and

$$k = \frac{dF}{dr}|_{R_L} = -\frac{c^6}{4G^2M^2} \tag{131}$$

$$\sqrt{|k|} = \frac{c^3}{2GM}$$

Thus, $\omega = \frac{c^3}{2GM}$ for a core of mass M. For instance a non-rotating core of mass = 1000 times the sun mass would pulsate on its first mode at $\omega = 101\ rad/s = 16Hz$.

12. Expansion of the universe

The estimated mass of the known universe is in a range $1.7\,10^{52}$ to $1.7\,10^{54}\,kg$. Let us calculate the R_L of the universe.

$$R_L = \frac{2GM}{c^2} = 2.5\,10^{25} \text{ to } 2.5\,10^{27}\ m.$$

The estimated radius of the universe according to the standard cosmological model is $46\,10^9$ light years = $4.2\,10^{26}\ m$. So the estimated radius of the universe is in the same range of magnitude as its R_L radius and it could even be very close to its R_L radius! And the universe could then have some properties of a core. Figure 15 represents $-\Phi$ of the universe between 0.7 and $2\,R_L$.

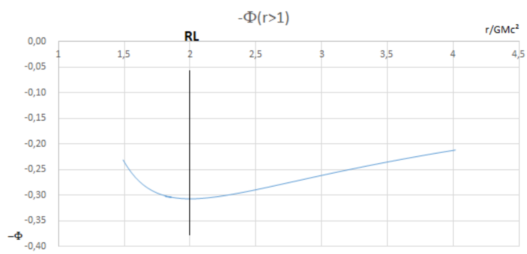


Figure 15. $-\Phi$ for $r > GM/c^2$

If considered as a non rotating core, the universe would pulsate around its R_L size, at a frequency $\frac{c^3}{2GM}$ or a period $\frac{4\pi c^3}{2GM} = 278\,10^9$ years, considering $R_L = 4.2\,10^{26}m = \frac{2GM}{c^2}$ (point 2 of Fig 16).

If the universe is now in an expansion phase, this would mean that its size is presently lower than its horizon R_L.

If the universe had begun at a size $< 1.27\,R_L/2$, it would have enough potential energy to expand to an infinite radius if there is no energy loss during that expansion. Otherwise the universe will oscillate or fluctuate around its R_L radius, where Φ is maximum. This resembles the A. D. Sakharov's concept of the fluctuating or oscillating universe (Al'tshuler 1991).

13. Acceleration of the expansion of the universe

Let us consider a body of mass m situated on the rim R_L of the universe, the force acting on this body is (83)

$$F = mc^2 \frac{1}{\left(1 - \frac{GM}{c^2 r}\right)^2} \left(\frac{-GM}{c^2 r^2} + 3\frac{G^2M^2}{c^4 r^3} - 2\frac{G^3M^3}{c^6 r^4} \right).$$

The equation of motion is

$$m\ddot{r} - c^2 \frac{1}{\left(1 - \frac{GM}{c^2 r}\right)^2} \left(\frac{-GM}{c^2 r^2} + 3\frac{G^2M^2}{c^4 r^3} - 2\frac{G^3M^3}{c^6 r^4} \right) m = 0$$

with the potential given by

$$\Phi(r) = c^2 \left(\frac{2GM}{rc^2} + \ln\left(1 - \frac{GM}{rc^2}\right)\right).$$

Again, we represent $-\Phi$ on the graphs, to make the presentation more intuitive. For small motion

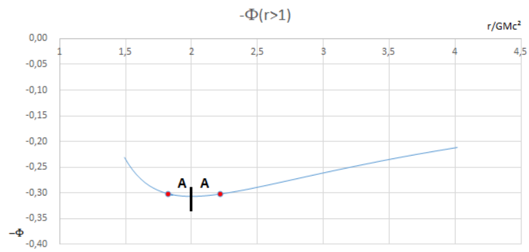


Figure 16. Amplitude A

around the value R_L, we approximate $F(r)$ by an harmonic force $F(r) = k(r - RL)$ with $k = \frac{c^6}{4G^2M^2}$ (eq. 131) and with $r = R_L$ at $t = 0$ we get:

$$(r - R_L) = A \sin\left(\frac{c^3}{GM} t\right)$$

A being the maximal fluctuation in size of the universe, and

$$(r - \dot{R_L}) = A\frac{c^3}{GM} \cos\left(\frac{c^3}{GM} t\right)$$

is the expansion rate of the universe, and

$$(r - \ddot{R_L}) = -A\frac{c^6}{G^2M^2} \sin\left(\frac{c^3}{GM} t\right)$$

is the acceleration of the expansion of the universe. When $r < R_L$, the expansion and the acceleration are both positive, but with a negative rate of the acceleration. This is a common feature of periodic motion. When the radius of the universe will reach its horizon R_L the expansion will continue but at a decelerating rate until the universe reaches its maximum size $R_L + A$ and stops expanding. Then the inverse movement will take place.

A and t are two unknowns which can be determined by the values of the expansion rate (Hubble-Lemaître constant = 70 km/s/Mps) and the value of the acceleration of the universe expansion.

14. Conclusion

The present model for gravitation is not equivalent to general relativity in that the field Γ determines the geometry and plays a central role. The existence of a short range repulsive potential leads to very interesting results, especially with regard to black holes which should not exist but instead would be "cores" or "compact bodies" that have a finite size. The current that generates the gravitational forces is the Energy-Impulsion tensor and this is a natural consequence of the invariance of the action under the group of translations in space-time. The repulsive $-\frac{G^2M^2}{r^2}$ potential term is absent in general relativity. For that reason and in order to explain the acceleration of the expansion of the universe, the influence of a hypothetical 'dark energy' was invoked in GR. *Our model does not need 'dark energy' to explain the acceleration of the expansion of the universe.* The repulsive $-\frac{G^2M^2}{r^2}$ potential tends to open the orbit of Mercury and this contributes to fix the advance the perihelion to $42.3''$ per century. This is in good agreement with the measured value of $43''$. The expansion of the universe could be a consequence of the universe being considered as a core with a natural pulsation frequency of one cycle per $278\ 10^9$ years. As such the universe would radially oscillate around an equilibrium point instead of being originated from a 'Big Bang". The speed distribution in rotating galaxies arms could also be calculated in the new theoretical model, possibly taking into account a "Lorentz" force. What's more, we can show that this "Lorentz" force acts in the right centripetal direction without maybe having to rely on "dark matter" to do the job. Quantum gravity should have the massless spin 2 graviton for the propagator of the interaction and this quantification could be the subject of further work. Rotating cores are also a topic for further study.

Acknowledgement

I wish to thank Martin Poty for his precious help in the typesetting of this paper and a reviewer of Kielnarowa Review for his patient work in proofreading this paper.

Competing Interests None.

References

Al'tshuler, B. L. 1991. The scientific works of A. D. Sakharov. *Usp. Fiz. Nauk* 161:3–27.

Arcos, H. I., and J. G. J G Pereira. 2005. *Torsion gravity: A reappraisal.* arXiv: 0501017v1.

Blagojevis, M., and F. Hehl. 2012. *Gauges theories of gravitation, a reader with commentaries.* arXiv: 1210.3775.

Dutta, S., and R. Scherrer. 2009. Dark energy from a phantom field near a local potential minimum. *Phys. Lett. B* 676 (2): 12–15.

Feynman, R. 1995. *Lectures on gravitation.* Addison-Wesley.

Hayashi, K., and T. Nakano. 1967. Extended translation invariance and associated gauge fields. *Progress of Theoretical Physics* 38 (2).

Hehl, F., and Y. Obukhov. 2007. *E. Cartan torsion.* arXiv: 0711.1535.

Ivanenko, D., and D. Sardanashvily. 1987. On the Goldstonic gravitation theory. *Pramana J. Phys.* 29 (1): 21–37.

Kleinert, H. 2010. New gauge symmetry in gravity and the evanescent role of torsion. *EJTP* 7 (24): 287–298.

Landau and Lifschitz. 1964. *Physique Théorique II: Théorie du champ.* Moscow, USSR: MIR.

Lurie, J. 2013. An introduction to teleparallel gravity, edited by R. Aldrovandi and J. G. Pereira, vol. 173. Fundamental Theories of Physics. Springer.

Peebles, P., and Ratra. 2002. *The cosmological constant and dark energy.* arXiv: astro-ph/0207347v2.

Scholz, E. 2018. *E. Cartan attempt at bridge-building between Einstein and the cosserats.* arXiv: 1810.03872v1.

Schucking, E. 2008. *Gravitation is torsion.* arXiv: 0803.4128v1.