Is a Mathematical Definition of Observation Possible?

Daniel Wagner Fonteles Alves

Centro Brasileiro de Pesquisas Físicas

dw-alves@hotmail.com

Abstract. The measurement problem is one of the central puzzles of Quantum Mechanics. By exploring the foundations of General Relativity from the 3-space perspective and the concept of meaningful information we find what may be a hint towards a mathematical definition of observation.

1.Introduction

Several experiments made at the beginning of the last century have shocked mankind forever. Phenomena such as wave-particle duality, discrete energy levels, non-local interaction were discovered, predicted and explained by the most enigmatic theory of all times: quantum mechanics (QM). Since its proposal, however, there has been a general feeling that something was missing. Even though the theory provides us with an infallible mathematical algorithm for predicting the outcomes of experiments, some of its fundamental concepts remain unclear.

One of the most important quantum phenomena is the observer effect. Traditional QM divides all possible physical interactions in two: those that are observations and those that are not. A system subject to a Hamiltonian H will evolve unitarily according to Schroëdinger´s equation:

$$ih\frac{d}{dt}\varphi = H\varphi$$

To any observable quantity there corresponds a linear operator A. When subject to an observation, the evolution changes drastically and the ket φ evolves to one of the eigenkets of the operator A with a probability defined by the theory. The problem then immediately comes: what is an observation? How can we know if an interaction is an observation or not? In principle, the whole process of measurement can be reduced to, say, an electromagnetic interaction, which has an associated Hamiltonian. So what makes observation special? Why the evolution ceases to be unitary?

Numerous answers have been proposed. They range from splitting the world into classical and quantum (Copenhagen interpretation), to assigning a distinctive physical character to consciousness. Others have built a conceptual apparatus that actually deny any observer effect, such as the many-worlds interpretation. In the heart of the problem lies the definition of observation. If we could, by any means, define observation mathematically and unambiguously we´d be able to understand QM much better. The task however seems formidable: at first sight, there´s nothing special to observation. It is just another interaction which in principle should have its associated Hamiltonian. A mathematical definition seems impossible. Could there be a solution to this puzzle?

2.Space and Time

Let´s forget for now all the fundamental problems in quantum mechanics. We will now focus on a subject that seems completely independent: the nature of space and time. Let´s go back to where everything begun. Consider Newton´s law:

$$m\ddot{q}_i = -(\nabla V)_i$$

Where V is a function that could in principle depend on (x,y,z,t).

Positions, by their own *definition*, are always expressed in relation to something and the $q_i(t)$ are expressed in relation to absolute space. But since absolute space cannot be observed, how could this theory be tested?

We could search and find a particular observable object which, when building the functions $q_i(t)$ in relation to it, the second order differential equations mentioned before would hold. This particular object chosen in this way is said to constitute an inertial frame of reference. It was shown in the 19th century how to define an inertial frame of reference operationally. According to the differential equation itself applied to the inertial frame object, it must either be at rest or moving with constant velocity in relation to absolute space. So we could ‘’almost’’ see absolute space, except that there would still be no operational way to discover absolute velocity and absolute position. It is a strange coincidence the fact that it turns out that no physical process actually depends on absolute position and absolute velocity.

But, is the notion of absolute space really indispensable? Newton knew that all he really measured were particle separations r_{ij}. Could we make physics dispensing absolute positions? This is the proposal of the physicist Ernst Mach. According to him, motion is better described by the observable relations between physical objects, all the rest being metaphysical excess. This is the relational conception of motion.

Another similar objection was raised against the conception of time. In Newton´s classical mechanics, time is a monotonic label attached to configurations of the universe that gives them an order. But, as with absolute space, there seems to be excessive structure: we only have access to time by looking at clocks, which are physical objects themselves. So time is a derived concept from that of motion of physical objects. In a relational picture, to say that time runs in an empty universe is both useless (for there is no motion to be analyzed) and meaningless (for the flow of time is unobservable). No change equals no time. If everything speeded up in absolute time, including clocks, could you tell the difference? In Mach´s own words:

‘’It is utterly beyond our power to measure the changes of things by time. Quite the contrary, time is an abstraction at which we arrive by means of the changes of things.’’

Mach´s thoughts were very similar to Wheeler´s it from bit proposal in the sense of being an inspiring idea, but hard to be defined and put forward in a complete theory. One of the scientists that tried to translate to Mach´s thoughts to rigorous physics was Einstein when he formulated GR [2]. However, instead of creating a theory where only observable data (relative distances) appear, Einstein still used “invisible” coordinate systems by presupposing a space-time manifold. GR is, thus, an indirect solution to the Machian problem of the origin of inertia and history tells that the reason why Einstein chose this path was the supposed impracticability of the direct one, which would use only relative distances from the start [12].

But what would we find if we tried to do what Mach suggested and that Einstein thought was impractical, that is, develop physics without absolute space and external time from the start?

3. Barbour´s Relational Physics

Julian Barbour is a British scientist who has dedicated much of his life to the study of Machian principles and their translation to physics. With the help of collaborators, he has successfully implemented Mach´s thoughts [3],[4]. I will give a quick exposition of his work and those interested on the details should look at [5] [6] [9] [10]

For setting up rigorously what a Machian theory should achieve we have to first understand the important concepts of Relative Configuration Spaces and Shape Spaces, used extensively by Barbour. When we look at N point-particles living in Newton´s Euclidean absolute space, the observable data upon that is actually the set of $N(N-1)/2$ numbers r_{ij} of particle separations. These numbers are not independent: they satisfy constraints analogous to the familiar triangular inequalities (when N=3). By performing absolute rotations and translations, the set r_{ij} does not change. Thus, in principle, they can be expressed in a set of $3N-6$ intrinsic coordinates which eliminates the Cartesian redundancy. We should also turn our attention to the set of $3N-7$ numbers $\tilde{r}_{ij}$ which are also invariant to absolute contractions and dilatations. This is because when we say a distance between two objects is “x “ we actually mean that the ratio between this distance and a particular reference length is x, so that only ratio of distances should be significant. For each N, there are many possible sets $r_{ij}, \tilde{r}_{ij}$, and each of these sets is a point in the r*elative configuration space* (RCS) and in *shape space* (SS), respectively.

Information in SS is exactly all that is directly seen-every time we state that a physical system is represented as a point in the configuration space (for instance the Cartesian coordinates of n particles $\{x_i\}_j$)all that is observed is the projection of this point on SS. Barbour states that for a theory to be perfectly relational, *a point and a direction in SS should uniquely determine a curve in SS* [7] . A direction, rather than a velocity, is specified because there should be no unobservable time parameter t. This makes explicit ‘’the no change equals no time’’ aspect of Machian philosophy.

But we know that Newtonian mechanics has good empirical results, for instance, when we substitute the absolute space for the distant stars and the external time with the rotation of the earth. So, by applying the requirements for a relational theory, could we recover Newtonian mechanics-that is, derive the concepts of duration and inertial frame-from data in SS?

The derivation of duration is very simple. Newtonian mechanics can be cast as a timeless action principle known as Jacobi´s principle [8]:

$$\delta A_j = 0, \qquad A_j = 2\int \sqrt{\bar{T}(E-V)}\, d\gamma \;\therefore\; \bar{T} \equiv \frac{1}{2}\sum_{i=1}^{n} m_i \frac{dx_i}{d\gamma}\frac{dx_i}{d\gamma}$$

This can also be written as

$$A_j = \int \sqrt{2(E-V)} \sqrt{\sum_i m_i \delta \boldsymbol{x}_i \delta \boldsymbol{x}_i}$$

Where γ is an arbitrary parameter and E is the total energy of the system. The resulting Euler-Lagrange equations are:

$$\frac{d\boldsymbol{p}_i}{d\gamma} = -\sqrt{\frac{\bar{T}}{E-V}} \frac{\partial V}{\partial \boldsymbol{x}_i} \therefore \boldsymbol{p}_i \equiv \frac{\partial L}{\partial \boldsymbol{x}_i} = \sqrt{\frac{E-V}{\bar{T}}} m_i \frac{d\boldsymbol{x}_i}{d\gamma}$$

But the parameter γ is arbitrary. Choosing it in such a way that $\sqrt{\frac{\bar{T}}{E-V}}=1$, we get

$$\frac{d\boldsymbol{p}_i}{d\gamma} = -\frac{\partial V}{\partial \boldsymbol{x}_i}$$

This is exactly Newton´s second law, so that γ may be identified with t. The condition $\sqrt{\frac{\bar{T}}{E-V}}=1$, which is just conservation of energy for the universe in the Newtonian view actually *defines* time in the Machian view as

$$\delta t = \sqrt{\frac{\sum_i m_i \delta \boldsymbol{x}_i \, \delta \boldsymbol{x}_i}{2(E-V)}}$$

It is usually interpreted that Jacobi´s principle gives us just the *path* of the system through configuration space, not the motion in time, which would be found after imposing energy conservation. But in the Machian view, this is the *definition of time upon motion*. The $\delta \boldsymbol{x}_i$ however are defined in a Cartesian frame which comes with redundancy as said before. To obtain a completely Machian theory we still need to eliminate absolute space.

Barbour and Bertotti proposed the method of *"best-matching"* to resolve that. It works as follows: suppose you have two configurations of the universe, described in absolute space with coordinates $\boldsymbol{x}_i$ and $\bar{\boldsymbol{x}}_i$. Since there is no absolute space, we cannot specify $\delta \boldsymbol{x}_i$ because of the Cartesian redundancy. Instead, hold one configuration fixed and apply Euclidean translations and rotations to the other until the incongruence measure

$$\sum_i \sqrt{m_i \delta \boldsymbol{x}_i \delta \boldsymbol{x}_i}$$

gets minimized. Now the $\delta \boldsymbol{x}_i$ calculated with that specific rotation and translation is unique. This is defined as the best-matched distance between the two configurations. Notice that only data from RCS or SS is needed to determine it. We then recover Newton´s laws exactly by using a Jacobi-type action with best-matched distances:

$$A_j = \int \sqrt{2(E-V)} \sqrt{\sum_i m_i \delta \boldsymbol{x}_i{}^{bm} \, \delta \boldsymbol{x}_i{}^{bm}}$$

Actually the procedure of best-matching imposes constraints such as zero total angular and linear momentum, zero total energy for the whole universe and also forces the potential V to depend only on r_{ij}.

Those ideas can be extended to field theories, and when applied to 3-D metric field, General Relativity is recovered in a sense [9],[10] ! This is something truly remarkable. General Relativity (actually shape dynamics, which is GR with a few modifications) can be seen as a direct consequence of imposing a relational conception of motion to a 3-D metric field theory. The bending of light in a gravitational field, the perihelion precession of Mercury, time dilatation… all of this can be seen as a result of simply imposing a relational conception of motion. How can such a philosophical idea be so powerful? Why does it work so well? We should spend more time thinking about this.

4.Observation and Meaningful Information

When Mach proposed his relational conception of motion, he was worried about the meaning of some statements in Newtonian Mechanics. Take a configuration of n-point particles described by Newton as $\vec{r_i} = \vec{a_i}$ and $t = t_0$.

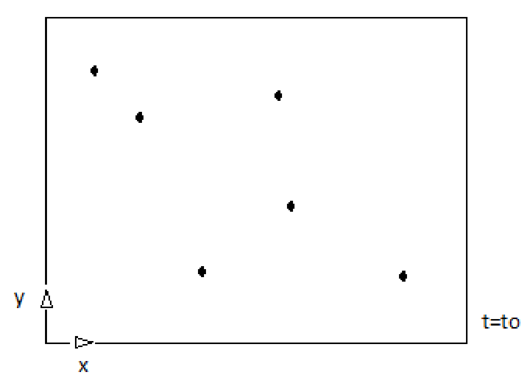

Fig.1

If we now translate the whole system rigidly, that is, if we apply a transformation $\vec{r_i} \rightarrow \vec{r_i} + \vec{c}$ we could never tell the difference. This transformation is empirically unverifiable. For Newton this is just an accident but for Mach the two descriptions mean exactly the same: empirically indiscernible configurations should be physically identical. He was proposing a deep link between the *meaning* of statements produced by our theoretical apparatus and *observable* facts. Using a coordinate system to identify configurations of the universe in our theories brings more information than what we have access to: only the particle separations r_{ij} are visible.

The moral is that sometimes we can describe a physical system with different mathematical structures. In the case of classical mechanics, we can use, for instance, the Cartesian coordinate system (x, y, z) or simply the r_{ij}, as shown above. These two structures carry the same amount of meaningful information, which for Mach resides on the empirically accessible r_{ij}, but (x, y, z) comes also with a little bit of redundancy.

Now consider a general situation. Let x and y be two configurations of the universe described by an arbitrary mathematical structure. Let´s imagine a process between these two configurations, which is also described within this mathematical structure.

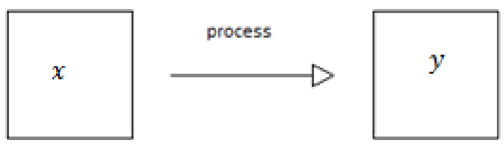

Fig.2

We have just seen that it may happen that we can do the same physics with very different mathematical structures (for instance, we could do classical mechanics on Configuration Space or Relative Configuration Space). Let the same process and configurations above be described in this different structure:

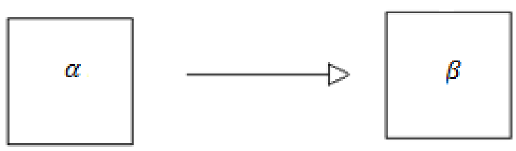

Fig.3

These objects and arrows are said to constitute "categories". For a quick exposition with the essential definitions of category theory take a look at the appendix. Now it is reasonable to assume that we can somehow have a translation between these two languages, that is, we can map configurations and process in one structure to configurations and process in the other. The mathematical object that does that in category theory is called a "functor". Let´s call the functor that maps mathematical descriptions of configurations and process that have the same meaning the "semantic functor". Then the following square should commute:

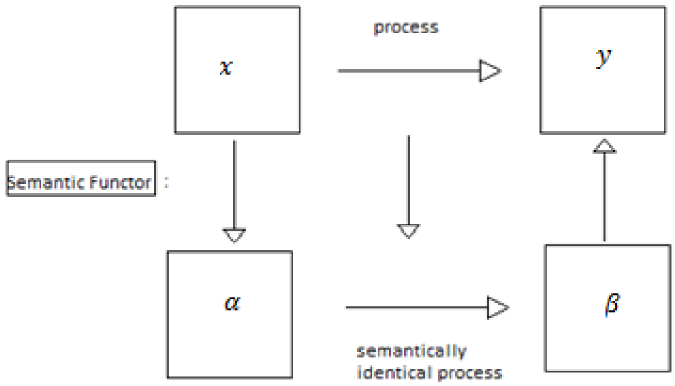

Fig. 4

Ok, now how can we tell, in this general setting, whether two configurations have the same meaning? We could in principle think in a lot of ways to make the square commute but there is one very special choice. It is that criterion of meaning that, speaking roughly, takes "no difference to mean no "observational" difference". That is, if there is no observational difference between the descriptions x and α then they mean the same. This is the Machian criteria of meaning that generates GR in the sense of Barbour´s argument, as we just saw above.

Now look what we have done: it seems we´re very close to a mathematical definition of observation! We can simply forget our common usage of this word and state that observation is "the criteria of meaning that generates General Relativity". That is, *somehow, the mathematical definition of observation lies in the semantic functor*. We have associated what seemed an obscure concept to an unambiguous mathematical object!

Notice that in this context we don´t need to invoke any special role to consciousness or the human mind to define observation. The human being does not enter at all in such a formal definition. Equipped with it, we can start to investigate what it means to "observe" something and why there is an observer effect at all. This does not seem to be a trivial task, but without a mathematical definition of observation it is clearly impossible.

Furthermore, we are having trouble to put QM and GR in the same conceptual scheme. The theory of Quantum Gravity is the most longstanding problem ever: never in the history of physics it took so much to unify two incompatible theories [1] (and there is no irrefutable success yet). A theory of QG would lead us to a completely new level of understanding of space, time and matter, finally completing the revolution started by QM and GR in the beginning of the last century. Could observation be at the core not only of QM, but also of GR?

5.Conclusion

Starting by questioning the concept of space and time in classical mechanics, we were led to replace Newtonian absolute structures with relative configurations. By extending these ideas to a 3-d metric field theory, GR is recovered. And, upon that and the recognition that different mathematical structures may carry the same meaningful information about the physical world, we were able to associate the seemingly imprecise concept of observation to a mathematically unambiguous object. Of course there is still a long way to go before we have a really precise definition of observation. What is the rigorous relation between diagram commutativity on figure 4 and GR? Can Barbour´s 3-space approach to GR really be cast in category theoretical language? Once we associate "observation" with the semantic functor, what does it mean to observe a physical system? And there are a lot of other questions. We don´t have, yet, a mathematical definition of observation, but maybe we could find it right in the heart of general relativity.

Acknowledgements

I´d like to thank the financial help of Brazilian's National Council for Scientific and Technological Development. I´d also like to thank the excellent discussions with the Foundations and Frontiers group at the Brazilian Center for Research in Physics.

Appendix: Category Theory

A category is a collection of objects A, B, C… together with a set hom(A, B) for each of pair of objects A, B of the category (If $f \ni$ hom(A, B) then we write $f: A \to B$ or $A \xrightarrow{f} B$) such that :

- For every object X, there is an identity morphism: $Id_x: X \to X$
- Morphisms are composable, that is, given $A \xrightarrow{f} B$ and B $\xrightarrow{g} C$ there is $A \xrightarrow{gf} C$ and the composition is associative.
- Any identity morphism satisfies, for any morphism f: $A \xrightarrow{f} B$, $fId_A = f = Id_B f$

These definitions make Category theory a highly abstract, generic formalism for mathematics. It has been shown by Baez[11] that the formalism of category theory makes a number of analogies between physics, logic, topology and computation become natural. For instance, in quantum theory, objects are *Hilbert spaces* and the morphisms are *linear operators*. In logic, objects are *propositions* and morphisms are proofs. Baez argues that these analogies suggest that category theory is a language suitable for developing *a general science of systems and process.*

Given two categories A, B, a functor $F: A \to B$ is a map sending each A-object a to a B-object $F(a)$, and each A-morphism f, to a B-morphism $F(f)$ such that F preserves identity and composition ($F(Id_a) = Id_{F(a)}$, $F(gf) = F(g)F(f)$).

References

[1] L. Smolin, *The trouble with Physics*, Houghton Mifflin Harcourt (2006)

[2] Huggett, Nick and Hoefer, Carl, *Absolute and Relational Theories of Space and Motion*, The Stanford Encyclopedia of Philosophy (Fall 2009 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/fall2009/entries/spacetime-theories

[3] J. B. Barbour and B. Bertotti, *Mach's Principle and the Structure of Dynamical Theories* Proc. R. Soc. A 382 (1982) no. 1783, 295–306.

[4] E. Anderson, *Foundations of Relational Particle Dynamics*, Class. Quant. Grav. 25 (2008) 025003 [arXiv:0706.3934 [gr-qc]]

[5] J. Barbour, *Shape Dynamics: An Introduction*, To be published in the refereed proceedings of the conference Quantum Field Theory and Gravity (Regensburg, 2010)(arXiv:11.05.0183v1[gr-qc])

[6] E. Anderson, J. Barbour, B. Foster, N. O'Murchadha, *Scale invariant gravity: Geometrodynamics*, Class. Quant. Grav. 20 (2003) 1571. [gr-qc/0211022

[7] J.Barbour, *The Definition of Mach´s Principle*, (2010) to be published in Foundations of Physics as invited contribution to Peter Mittelstaedt´s 80th birthday Festschrift

[8] C. Lanczos, *The Variational Principle of Mechanics*, Toronto University Press (1949)

[9] Barbour J, Foster B, and O Murchadha N, *Relativity without Relativity*, Class. Quantum Grav. 19, 1571 (2002) (arXiv: gr-qc/0012089).

[10] H. Gomes, T.Koslowski, *The link between General Relativity and Shape Dynamics*, Class.Quant.Grav. 29 (2012) 075009 (arXiv:1101.5974v3[gr-qc])

[11] J.C.Baez, *Physics, Topology, Logic and Computation: A Rosetta Stone*, Physics, topology, logic and computation.In New Structures for Physics, *ed. Bob Coecke, Lecture Notes in Physics vol. 813, Springer, Berlin, 2011, pp. 95-174* (2009) (arXiv:0903.0340v3 [quant-ph])

[12] J.Barbour, *The Discovery of Dynamics: A Study From the Machian Point of View of the Discovery and the Structure of Dynamical Theories*, Oxford University Press (2001)