Article:
Pusey, Matthew F. orcid.org/0000-0002-6189-7144 (2014) Anomalous Weak Values Are Proofs of Contextuality. Physical Review Letters. 200401. ISSN 1079-7114

https://doi.org/10.1103/PhysRevLett.113.200401

Anomalous Weak Values Are Proofs of Contextuality

Matthew F. Pusey*

Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario N2L 2Y5, Canada

(Received 5 September 2014; published 12 November 2014)

The average result of a weak measurement of some observable A can, under postselection of the measured quantum system, exceed the largest eigenvalue of A. The nature of weak measurements, as well as the presence of postselection and hence possible contribution of measurement disturbance, has led to a long-running debate about whether or not this is surprising. Here, it is shown that such "anomalous weak values" are nonclassical in a precise sense: a sufficiently weak measurement of one constitutes a proof of contextuality. This clarifies, for example, which features must be present (and in an experiment, verified) to demonstrate an effect with no satisfying classical explanation.

DOI: 10.1103/PhysRevLett.113.200401 PACS numbers: 03.65.Ta, 03.67.-a

In 1988 Aharonov, Albert and Vaidman explained "how the result of a measurement of a component of the spin of a spin-$\frac{1}{2}$ particle can turn out to be 100." [1] Defining the weak value of an observable A for a quantum system prepared in state $|\psi\rangle$ and postselected on giving the first outcome of $\{|\phi\rangle\langle\phi|, I - |\phi\rangle\langle\phi|\}$,

$$A_w = \frac{\langle\phi|A|\psi\rangle}{\langle\phi|\psi\rangle}, \tag{1}$$

they exhibited a $|\psi\rangle$ and $|\phi\rangle$ on a qubit for which $Z_w = 100$. The motivation for weak values starts by considering a von Neumann model [2] of the measurement of A. The strength of the interaction between the system and "pointer" is then drastically reduced, such that the pointer reading is correlated only slightly with A. The weak value then arises as an approximation of the average pointer reading to first order in the interaction strength.

Weak values outside the eigenvalue range of A are termed anomalous. Aside from possible practical applications (see Ref. [3] and references therein), it has been suggested that such values have foundational significance. For example, both their theoretical prediction and experimental observation are said to shed light on "quantum paradoxes" [4–9] and even the nature of time [10].

However, there is still no consensus on the most basic question about anomalous weak values: to what extent do they represent a genuinely nonclassical effect? The lesser the extent, the more severe are the limitations on their practical and foundational significance.

The arguments that anomalous weak values are nonclassical have often been somewhat heuristic, appearing to depend on issues such as the extent to which weak measurements should be called measurements at all [11,12]. Perhaps the most rigorous evidence provided so far is a connection between anomalous weak values and the failure of a notion of classicality called "macroscopic realism" [13–15]. On the other hand, classical models have been given that reproduce various aspects of the phenomena [16–18].

The question can be made precise by asking if anomalous weak values constitute proofs of the incompatibility of quantum theory with noncontextual ontological models [19], or equivalently [20] if anomalous weak values require negativity in all quasiprobability representations. This was conjectured to be the case in Ref. [21]. Here I will prove it. Interestingly, the proof hinges on two issues already identified in the literature: what do weak measurements measure, and how much do they disturb the system? It transpires that both questions have clear answers in the setting of a noncontextual ontological model, but the particular information-disturbance tradeoff of the weak measurements in quantum theory makes these answers irreconcilable with the anomaly.

Let us begin by specifying exactly what is meant by an anomalous weak value. Inspection of Eq. (1) shows that A_w need not be real even though A is Hermitian. A complex number will certainly not be a convex combination of the eigenvalues of A, and so this might be seen as surprising. However, the imaginary part of A_w is manifested very differently from the real part [22]. Indeed, complex weak values are easily obtained even in the Gaussian subset of quantum mechanics, which has weak measurements (with the same information-tradeoff disturbance utilized here) and yet admits a very natural noncontextual model [23]. Hence, I will call a weak value A_w anomalous only when $\mathrm{Re}(A_w)$ is smaller than the smallest eigenvalue of A, or larger than the largest eigenvalue of A.

A simplification can be obtained by substituting the spectral decomposition $A = \sum_a a\Pi^{(a)}$ into the rhs of Eq. (1) and taking the real part:

$$\mathrm{Re}(A_w) = \sum_a a\mathrm{Re}\left(\frac{\langle\phi|\Pi^{(a)}|\psi\rangle}{\langle\phi|\psi\rangle}\right) = \sum_a a\mathrm{Re}(\Pi_w^{(a)}).$$

If we had $0 \leq \mathrm{Re}(\Pi_w^{(a)}) \leq 1$ for all a then A_w could not be anomalous. Hence, an anomalous weak value for any observable always implies an anomalous weak value for a projector. Since $\sum_a \Pi_w^{(a)} = I_w = 1$, if one projector has $\mathrm{Re}(\Pi_w^{(a)}) > 1$ then another must have $\mathrm{Re}(\Pi_w^{(a')}) < 0$. In conclusion, without loss of generality we can always take the anomalous weak value to be associated with projector Π having $\mathrm{Re}(\Pi_w) < 0$.

I will now briefly review the relevant notion of noncontextuality, following Ref. [19] (where the definitions are motivated and compared to the traditional definition of noncontextuality due to Kochen and Specker [24]). Assumptions of noncontextuality are constraints on an ontological model. I will only need two notions: measurement noncontextuality, and outcome determinism for sharp measurements. (The latter can be shown to itself follow from the assumption of preparation noncontextuality together with some simple facts about quantum theory, see Refs. [19,25] for details.)

Suppose we prepare a quantum system in some way, represented in quantum theory by a state $|\psi\rangle$. In an ontological model the preparation is represented by a probability distribution $p(\lambda)$ over a set of ontic states Λ. Suppose we now implement the positive operator valued measure (POVM) $\{E_k\}$. In a measurement noncontextual model, this is represented by a conditional probability distribution $\{p(E_k|\lambda)\}$. The assumption of measurement noncontextuality is what allows us to write $p(E_k|\lambda)$ as a function of the effect E_k and the ontic state λ only, with no dependence on other things ("contexts"), such as the other elements of the POVM or details of how the POVM was implemented. Outcome determinism for sharp measurements is the assumption that $p(\Pi|\lambda) \in \{0, 1\}$ for all projectors Π and ontic states λ, so that any inability to predict the outcome of a projective measurement is due purely to ignorance of λ.

The final requirement, for any ontological model, is that when we marginalize over the ontic states, the model must reproduce the predictions of quantum theory:

$$\langle\psi|E_k|\psi\rangle = \int_\Lambda p(E_k|\lambda)p(\lambda)d\lambda. \qquad (2)$$

We can now state the main result, identifying certain features in the measurement of anomalous weak values that, taken together, defy noncontextual explanation.

Theorem 1: Suppose we have states $|\phi\rangle$ and $|\psi\rangle$, and a generalized measurement [26] $\{M_x\}_{x\in\mathbb{R}}$, such that

(a) the pre-and postselection are nonorthogonal, i.e.,

$$p_\phi := |\langle\phi|\psi\rangle|^2 > 0, \qquad (3)$$

(b) the POVM is a projector plus unbiased noise, i.e.,

$$E_x := M_x^\dagger M_x = p_n(x-1)\Pi + p_n(x)\tilde{\Pi} \qquad (4)$$

for some projector Π, $\tilde{\Pi} = I - \Pi$, and probability distribution $p_n(x)$ with median $x = 0$,

(c) we can define a probability p_d (the "probability of disturbance") such that

$$S := \int_{-\infty}^{\infty} M_x^\dagger|\phi\rangle\langle\phi|M_x dx = (1-p_d)|\phi\rangle\langle\phi| + p_d E_d \qquad (5)$$

for some POVM $\{E_d, I - E_d\}$, and

(d) the values of x under the pre-and postselection have a negative bias that "outweighs" p_d, i.e., [27]

$$p_- := \frac{1}{p_\phi}\int_{-\infty}^{0} |\langle\phi|M_x|\psi\rangle|^2 dx > \frac{1}{2} + \frac{p_d}{p_\phi}. \qquad (6)$$

Then there is no measurement noncontextual ontological model for the preparation of $|\psi\rangle$, measurement of $\{M_x\}$, and postselection of $|\phi\rangle$ satisfying outcome determinism for sharp measurements.

Showing that operators $\{M_x\}$ with these properties actually exist whenever we have a $|\psi\rangle$, $|\phi\rangle$, and Π with $\mathrm{Re}(\Pi_w) < 0$ is a routine calculation in the theory of weak measurement [1,22,28], postponed until later. Loosely speaking, if $g \ll 1$ is the strength of the measurement then to leading order $(p_- - \frac{1}{2}) \sim g$ whereas $p_d \sim g^2$.

Proof.—Suppose such an ontological model exists. We can consider the weak measurement $\{M_x\}$ followed by the projective measurement $\{|\phi\rangle\langle\phi|, I - |\phi\rangle\langle\phi|\}$ as one "consolidated measurement," represented by the POVM $\{S_x\}\cup\{F_x\}$, where $S_x = M_x^\dagger|\phi\rangle\langle\phi|M_x$ and $F_x = M_x^\dagger(I - |\phi\rangle\langle\phi|)M_x$. The key question is how the $\{S_x\}$ are represented in the model, because Eq. (2) gives

$$|\langle\phi|M_x|\psi\rangle|^2 = \langle\psi|S_x|\psi\rangle = \int_\Lambda p(S_x|\lambda)p(\lambda)d\lambda. \qquad (7)$$

Let us consider two methods for implementing the POVM $\{E_x\}$. By the assumption of measurement noncontextuality they must both lead to the same $p(E_x|\lambda)$. The first method is to implement the consolidated measurement and then ignore the result of the postselection, giving $p(E_x|\lambda) = p(S_x|\lambda) + p(F_x|\lambda)$. The second method, according to Eq. (4), is to measure $\{\Pi, \tilde{\Pi}\}$ and then classically sample from $p_n(x-1)$ or $p_n(x)$ as appropriate. Hence, we also have $p(E_x|\lambda) = p_n(x-1)p(\Pi|\lambda) + p_n(x)p(\tilde{\Pi}|\lambda)$. Since the median of $p_n(x)$ is 0 we have $\int_{-\infty}^{0} p_n(x-1)dx \leq \int_{-\infty}^{0} p_n(x)dx = \frac{1}{2}$. Combining this with $p(S_x|\lambda) \leq p(E_x|\lambda)$ from the first method, we have

$$\int_{-\infty}^{0} p(S_x|\lambda)dx \leq \int_{-\infty}^{0} p(E_x|\lambda)dx \leq \frac{1}{2}. \qquad (8)$$

Next, we apply the assumption of measurement noncontextuality to the POVM $\{S, I - S\}$. One way to implement this is to use the consolidated measurement and ignore x; hence, $p(S|\lambda) = \int_{-\infty}^{\infty} p(S_x|\lambda)dx$. A second way, according to Eq. (5), is to measure $\{|\phi\rangle\langle\phi|, I - |\phi\rangle\langle\phi|\}$

with probability $1 - p_d$ and $\{E_d, I - E_d\}$ with probability p_d. Hence, $p(S|\lambda) = (1 - p_d)p(|\phi\rangle\langle\phi||\lambda) + p_d p(E_d|\lambda)$.

Finally, we calculate the model's prediction for p_-. Using outcome determinism for the sharp measurement $\{|\phi\rangle\langle\phi|, I - |\phi\rangle\langle\phi|\}$ we can partition Λ into $\{\Lambda_0, \Lambda_1\}$, where $p(|\phi\rangle\langle\phi||\lambda) = i$ for $\lambda \in \Lambda_i$. From the above we have that $\int_{-\infty}^{0} p(S_x|\lambda)dx \leq p(S|\lambda) \leq p_d$ on Λ_0. Hence, splitting the rhs of Eq. (7) into integrals over Λ_0 and Λ_1 and integrating over $x < 0$ gives

$$\int_{-\infty}^{0} |\langle\phi|M_x|\psi\rangle|^2 dx \leq \int_{-\infty}^{0} \int_{\Lambda_1} p(S_x|\lambda)p(\lambda)d\lambda dx + p_d.$$

Applying Eq. (8) and recalling that Eq. (2) gives $\int_{\Lambda_1} p(\lambda)d\lambda = \int_{\Lambda} p(|\phi\rangle\langle\phi||\lambda)p(\lambda)d\lambda = |\langle\phi|\psi\rangle|^2 = p_\phi$, this gives

$$\frac{1}{p_\phi}\int_{-\infty}^{0} |\langle\phi|M_x|\psi\rangle|^2 dx \leq \frac{1}{2} + \frac{p_d}{p_\phi}, \tag{9}$$

in contradiction to Eq. (6).

As promised, I will now confirm that a projector Π with $\mathrm{Re}(\Pi_w) < 0$ implies the existence of a measurement $\{M_x\}$ with the properties assumed in Theorem 1.

Similarly to Ref. [1], the measurement begins by preparing a probe system in the Gaussian state $|\Psi\rangle = N\int_{-\infty}^{\infty} \exp(-x^2/2\sigma^2)|x\rangle dx$, with $N = (\pi\sigma^2)^{-1/4}$. This interacts with the system via the unitary (with $\hbar = 1$)

$$U = \exp(-i\Pi P) = \exp(-iP)\Pi + \tilde{\Pi}, \tag{10}$$

which defines our units of momentum and hence length, and then the probe is projectively measured in the $\{|x\rangle\langle x|\}$ basis. On the system this is a generalized measurement with $M_x = \langle x|U|\Psi\rangle$. Recalling that P generates translations we have

$$M_x = N\exp\left(-\frac{(x-1)^2}{2\sigma^2}\right)\Pi + N\exp\left(-\frac{x^2}{2\sigma^2}\right)\tilde{\Pi}. \tag{11}$$

This becomes a projective measurement in the limit $\sigma \to 0$, whereas it is known as a weak measurement for large σ. We can now calculate

$$E_x = M_x^\dagger M_x = p_n(x-1)\Pi + p_n(x)\tilde{\Pi}, \tag{12}$$

where $p_n(x) = N^2\exp(-x^2/\sigma^2)$ has median $x = 0$. Recalling that $p_n(x)$ is normalized and defining

$$\Delta := \int_{-\infty}^{\infty} \sqrt{p_n(x-1)p_n(x)}dx = \exp\left(-\frac{1}{4\sigma^2}\right), \tag{13}$$

we obtain

$$\begin{aligned} S &= \int_{-\infty}^{\infty} M_x^\dagger|\phi\rangle\langle\phi|M_x dx \\ &= \Pi|\phi\rangle\langle\phi|\Pi + \tilde{\Pi}|\phi\rangle\langle\phi|\tilde{\Pi} + \Delta(\Pi|\phi\rangle\langle\phi|\tilde{\Pi} + \tilde{\Pi}|\phi\rangle\langle\phi|\Pi) \\ &= \frac{1+\Delta}{2}|\phi\rangle\langle\phi| + \frac{1-\Delta}{2}(\Pi - \tilde{\Pi})|\phi\rangle\langle\phi|(\Pi - \tilde{\Pi}). \end{aligned} \tag{14}$$

Setting $p_d = (1-\Delta)/2 < 1$ and $E_d = (\Pi - \tilde{\Pi})|\phi\rangle\langle\phi|(\Pi - \tilde{\Pi})$ (which is a projector) we have Eq. (5).

Finally we need to calculate

$$\begin{aligned} p_- &= \frac{1}{p_\phi}\int_{-\infty}^{0} |\langle\phi|M_x|\psi\rangle|^2 dx \\ &= A|\Pi_w|^2 + B|\tilde{\Pi}_w|^2 + 2C\mathrm{Re}(\Pi_w\tilde{\Pi}_w^*), \end{aligned} \tag{15}$$

where we have recalled Eq. (1) and defined the integrals

$$A = \int_{-\infty}^{0} p_n(x-1)dx = \frac{1}{2}\left[1 - \mathrm{erf}\left(\frac{1}{\sigma}\right)\right], \tag{16}$$

$$B = \int_{-\infty}^{0} p_n(x)dx = \frac{1}{2}, \tag{17}$$

$$\begin{aligned} C &= \int_{-\infty}^{0} \sqrt{p_n(x-1)p_n(x)}dx \\ &= \frac{1}{2}\exp\left(-\frac{1}{4\sigma^2}\right)\left[1 - \mathrm{erf}\left(\frac{1}{2\sigma}\right)\right]. \end{aligned} \tag{18}$$

Expanding around $1/\sigma = 0$ we find $A \approx \frac{1}{2} - (1/(\sqrt{\pi}\sigma))$ and $C \approx \frac{1}{2} - (1/(2\sqrt{\pi}\sigma))$. Since $\Pi_w + \tilde{\Pi}_w = I_w = 1$ this gives

$$p_- \approx \frac{1}{2} - \frac{1}{\sqrt{\pi}\sigma}\mathrm{Re}(\Pi_w). \tag{19}$$

Meanwhile to leading order $p_d \approx 1/8\sigma^2$. Hence, provided $\mathrm{Re}(\Pi_w) < 0$, for sufficiently large σ we will satisfy Eq. (6). It is worth emphasizing that no approximations were made in the proof of Theorem 1, and in a concrete case one can simply plug values of σ into the exact formulas above to verify Eq. (6).

I will conclude by outlining three interconnected lessons from Theorem 1. The first is a classification of how anomalous weak values could arise in an ontological model. One possibility (perhaps the most common realist interpretation of anomalous weak values) is that some ontic states are predisposed to manifest such values, in violation of the first application of measurement noncontextuality using Eq. (4). Alternatively (along the lines of Ref. [18]) the weak measurement may disturb the system much more than the quantum formalism would suggest, in violation of the second application. The final possibility is that the postselection is not represented outcome deterministically (as in the interpretation where the ontic state is simply the quantum state) and so fails to identify a particular set of ontic states.

The second lesson is that a large number of aspects of the manifestation of anomalous weak values seem to be involved in preventing noncontextual explanation. The "anomaly" itself is only one ingredient. Some others may have been anticipated, such as the favorable information-disturbance tradeoff of weak measurements. But some seem somewhat surprising, for example the importance of the postselection being a projective measurement.

The final lesson is an indication of what it would take for an experiment involving anomalous weak values to exclude noncontextual theories that would provide a good classical explanation. Merely observing "anomalous pointer readings" under pre- and postselection is far from sufficient. Most fundamentally, the experiment must show that the probabilities in the statement of Theorem 1 really are the probabilities of discrete events, rather than mere (normalized) intensities. An experiment consistent with a classical field theory, so far the most common way to observe weak values, is therefore not sufficient [29]. One would also need to provide evidence for an operational version of Eqs. (4) and (5). Notice that these would be statements about how the weak measurement works on all preparations, not just the one corresponding to $|\psi\rangle$. Furthermore, one would need an operational counterpart to the inference from preparation noncontextuality to outcome determinism for the postselection measurement, perhaps by implementing preparations that make the postselection highly predictable (see Ref. [32] for how this can be done in more traditional proofs of contextuality). Turning these ideas into a concrete experimental proposal is an interesting avenue for future work.

Thanks are due to Aharon Brodutch, Joshua Combes, Chris Ferrie, Ravi Kunjwal, and Matt Leifer for useful discussions. I am particularly indebted to Matt Leifer for help in analyzing measurement disturbance, and to Aharon Brodutch for bringing the issue of intensities versus probabilities to my attention. Research at Perimeter Institute is supported in part by the Government of Canada through National Sciences and Engineering Research Council and by the Province of Ontario through Ministry of Research and Innovation.

*m@physics.org

[1] Y. Aharonov, D. Z. Albert, and L. Vaidman, Phys. Rev. Lett. **60**, 1351 (1988).

[2] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1996).

[3] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd, Rev. Mod. Phys. **86**, 307 (2014).

[4] Y. Aharonov and L. Vaidman, J. Phys. A **24**, 2315 (1991).

[5] K. J. Resch, J. S. Lundeen, and A. M. Steinberg, Phys. Lett. A **324**, 125 (2004).

[6] Y. Aharonov, A. Botero, S. Popescu, B. Reznik, and J. Tollaksen, Phys. Lett. A **301**, 130 (2002).

[7] J. S. Lundeen and A. M. Steinberg, Phys. Rev. Lett. **102**, 020404 (2009).

[8] Y. Aharonov, S. Popescu, D. Rohrlich, and P. Skrzypczyk, New J. Phys. **15**, 113015 (2013).

[9] T. Denkmayr, H. Geppert, S. Sponar, H. Lemmel, A. Matzkin, J. Tollaksen, and Y. Hasegawa, Nat. Commun. **5**, 4492 (2014).

[10] Y. Aharonov, S. Popescu, and J. Tollaksen, Phys. Today **63**, No. 11, 27 (2010).

[11] A. J. Leggett, Phys. Rev. Lett. **62**, 2325 (1989).

[12] Y. Aharonov and L. Vaidman, Phys. Rev. Lett. **62**, 2327 (1989).

[13] N. S. Williams and A. N. Jordan, Phys. Rev. Lett. **100**, 026804 (2008).

[14] M. E. Goggin, M. P. Almeida, M. Barbieri, B. P. Lanyon, J. L. OBrien, A. G. White, and G. J. Pryde, Proc. Natl. Acad. Sci. U.S.A. **108**, 1256 (2011).

[15] J. Dressel, C. J. Broadbent, J. C. Howell, and A. N. Jordan, Phys. Rev. Lett. **106**, 040402 (2011).

[16] J. Dressel and A. N. Jordan, Phys. Rev. A **85**, 022123 (2012).

[17] K. Y. Bliokh, A. Y. Bekshaev, A. G. Kofman, and F. Nori, New J. Phys. **15**, 073022 (2013).

[18] C. Ferrie and J. Combes, Phys. Rev. Lett. **113**, 120404 (2014).

[19] R. W. Spekkens, Phys. Rev. A **71**, 052108 (2005).

[20] R. W. Spekkens, Phys. Rev. Lett. **101**, 020401 (2008).

[21] J. Tollaksen, J. Phys. A **40**, 9033 (2007).

[22] R. Jozsa, Phys. Rev. A **76**, 044103 (2007).

[23] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Phys. Rev. A **86**, 012103 (2012).

[24] S. Kochen and E. Specker, Indiana University mathematics Journal **17**, 59 (1968).

[25] R. W. Spekkens, Found. Phys. **44**, 1125 (2014).

[26] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).

[27] Notice that although p_- is a combination of operationally defined quantities, it is not exactly the probability of getting a negative x under the pre-and postselection. To obtain this, instead of dividing by p_ϕ one would have to divide by $\langle\psi|S|\psi\rangle = (1 - p_d)p_\phi + p_d\langle\psi|E_d|\psi\rangle$, making the analysis slightly more complicated (but still tractable).

[28] J. L. Garretson, H. M. Wiseman, D. T. Pope, and D. T. Pegg, J. Opt. B: Quantum Semiclass. Opt. **6**, S506 (2004).

[29] This is because the analysis presented here, like any proof of contextuality, is for an ontological model that produces individual measurement results with the correct probabilities. This requirement immediately rules out a straightforward field ontology that, while perhaps offering interesting explanations of the weak value [17,30,31], only produces intensities. To exclude noncontextual explanation, an experiment based on fields would have to justify this

requirement by working at the level of single field quanta. Compare with the classic "double-slit experiment": while an interference pattern in intensities has a simple classical explanation in terms of fields, the same interference pattern in what are unambiguously probabilities defies classical intuitions.

[30] M. V. Berry and S. Popescu, J. Phys. A **39**, 6965 (2006).
[31] J. Dressel, K. Y. Bliokh, and F. Nori, Phys. Rev. Lett. **112**, 110407 (2014).
[32] R. W. Spekkens and R. Kunjwal, Perimeter Institute Seminar (2014), http://pirsa.org/14010102/.