

Встроенный наблюдатель и границы самопознания: квантовая теорема и трансцендентальная эпистемология

(с)2026 С.Ю. Матеров

Введение

В 2026 году нами была сформулирована и строго математически доказана Embedded-Observer No-Go Theorem - результат из области квантовой информации, утверждающий, что наблюдатель, реализованный как собственная подсистема замкнутой унитарной квантовой системы, не может одновременно (а) восстановить полное глобальное состояние системы из доступных ему данных и (b) эволюционировать по фиксированному автономному унитарному закону, если он действительно, нетривиально связан с остальной частью системы.

Формально это скромный результат: он не претендует на пересмотр интерпретаций квантовой механики и явно позиционирует себя как техническое уточнение уже известной линии рассуждений: от аргумента Вигнера о друге - до механизма Пейджа–Вуттерс и теоремы Фраухигер–Реннер. Новизна работы не только в объединении этой линии рассуждений, но и в двух узких, но точных технических результатах: точной световой границе для дискретных локальных квантовых схем и явном логическом разделении двух разных типов неудачи самопознания наблюдателя - «запаса» (stock, накопленная история, формализуемая энтропией фон Неймана) и «потока» (flow, свойство мгновенного взаимодействия, формализуемое унитарностью канала).

При этом сама архитектура результата почти неизбежно вызывает философскую ассоциацию: она воспроизводит структуру, до мельчайших деталей знакомую всякому читателю “Критики чистого разума”: различение вещи-в-себе и явления, тезис об априорных формах чувственности как условии всякого возможного опыта, идею конституирования предметности познающим субъектом, а не пассивного её считывания. Но здесь необходима методологическая дисциплина. Задача этого текста - не в том, чтобы объявить физическую теорему “математическим доказательством Канта” (что было бы категориальной путаницей), а в том, чтобы честно взвесить эпистемический статус самой параллели: что именно формальный, физико-математический результат может, а чего принципиально не может сказать философии науки об условиях познания, если субъект познания сам оказывается частью познаваемого мира.

1. Теорема без формул

Огрубляя до предела: пусть замкнутая физическая система описывается единой унитарной, обратимой, сохраняющей информацию эволюцией. Пусть наблюдатель - не что-то внешнее этой системе, не архимедова точка обзора, а её собственная часть: одна из двух подсистем в разбиении целого на “наблюдателя” (O) и “остальное” (R). Тогда:

- 1. Редукция принудительна.** Всё, что наблюдатель в принципе может знать, исчерпывается редуцированной матрицей плотности ρ_O - результатом операции частичного следа, «стирающей» степени свободы R. Это не эпистемическая небрежность и не приближение, а единственный логически возможный способ описать то, что доступно измерениям, ограниченным подсистемой O. Условие возможности знания здесь буквально совпадает с условием его ограниченности.
- 2. Редукция необратима эпистемически.** Если ρ_O смешанная, то бесконечно много

различных глобальных состояний целого дают ровно одну и ту же ρ_O (Hughston–Jozsa–Wootters, 1993). Наблюдатель, ограниченный O , принципиально не может по своим данным отличить эти состояния - глобальное состояние недоопределено его локальными данными не по случайности незнания, а по структуре самой ситуации.

3. **Локальная динамика не автономна.** Если взаимодействие O и R реально запутывающее, не существует фиксированного унитарного закона на одном лишь O , который описывал бы эволюцию ρ_O - она подчинена тому, что происходит в R , к чему у наблюдателя нет и не может быть доступа.
4. **Причинность геометрична.** Скорость, с которой вообще могут возникать корреляции между удалёнными частями системы, ограничена структурой локальных взаимодействий - приближённо для непрерывных гамильтоновых систем (граница Либа–Робинсона) и точно для дискретных локальных схем.

Философски существенная черта результата: он безразличен к любым специфическим физическим деталям - ни к массе, ни к температуре, ни к конкретной архитектуре субстрата. Он следует чисто из формы отношения между частью и целым, из самого факта, что наблюдатель есть часть системы, а не внеположная ей точка обзора. В этом смысле теорема сформулирована не о содержании опыта, а о его форме - что и делает философскую параллель небезосновательной.

2. Структурный изоморфизм с кантовской эпистемологией

Кантовский трансцендентализм - это учение об условиях возможности всякого опыта как такового: пространство и время как априорные формы чувственности, категории рассудка как условия синтеза многообразного восприятия в единство объективного опыта. Это не эмпирическая гипотеза о физическом устройстве мира, которую можно подтвердить или опровергнуть теоремой или экспериментом; Кант сознательно строил трансцендентальный аргумент так, чтобы он оставался независимым от содержания любой конкретной физической теории, обосновывая условия возможности познания “до” и “независимо” от эмпирики, à la “коперниканский переворот”: не познание соотнобразуется с предметами, а предметы с условиями нашего познания. Именно поэтому тезис “теорема доказывает философию Канта” есть категориальная ошибка в строгом райловском смысле: физическая теорема устанавливает свойства систем, удовлетворяющих её гипотезам, трансцендентальный же аргумент устанавливает условия возможности опыта в принципе, до всякой конкретной физики.

То, что теорема действительно даёт - вещь более скромная и, на наш взгляд, ничуть не менее интересная в качестве предмета философской рефлексии: **структурный изоморфизм**, воспроизводящий кантовскую архитектуру познания пункт за пунктом в достаточно общем классе физических систем.

(а) Вещь-в-себе и глобальное состояние. У Канта: существует объективная реальность (“вещь в себе”, Ding an sich), принципиально недоступная познанию, - нам даны лишь её явления. В теореме: существует чистое, симметричное, унитарное глобальное состояние $|\Psi\rangle$, которое наблюдатель, будучи сам частью системы, не может ни вычислить, ни однозначно восстановить локальным измерением (Lemma A). Оговорка обязательна: у Канта непознаваемость вещи-в-себе - трансцендентальный тезис, выводимый из анализа структуры познания как таковой; в теореме недоступность глобального состояния - это следствие конкретной математической структуры (энтропии смешанного состояния, не единственности его очищения). Изоморфизм

структурный, не логический - из одного не следует другое, но оба воспроизводят одну и ту же архитектонику принципиального разрыва между целым и доступным.

(b) Априорные формы чувственности и световой конус. Пожалуй, наиболее точное звено параллели. У Канта пространство и время - не свойства мира самого по себе, а формы, которые познающий субъект с необходимостью привносит в самый акт восприятия; именно поэтому геометрия синтетически применима к опыту *a priori*, не будучи выведена из него апостериорно. В теореме световой конус играет структурно ту же роль: не эмпирический факт о частной системе, а геометрическое ограничение, встроенное в само устройство взаимодействий, определяющее заранее, какие корреляции вообще могут возникнуть и когда. Наблюдатель не выбирает воспринимать мир причинно-упорядоченным - это принудительное, трансцендентально-подобное следствие самого факта локального взаимодействия с локально устроенной структурой.

Отдельно стоит развести это звено с другим элементом теоремы – “запасом” (Stock, накопленная энтропия истории взаимодействия). Категории и формы чувственности у Канта фиксированы и не накапливаются во времени - они даны заведомо, одинаково для всякого момента возможного опыта. Stock же есть именно накопленная, в общем случае монотонно растущая история; более уместная параллель для него - не кантовская трансцендентальная эстетика, а скорее вопрос о следе прошлого, впечатанном необратимо в наличное состояние системы, - отчётливо не тематизированная у самого Канта в этом виде и ближе к феноменологии внутреннего сознания времени (Гуссерль) или к философии длительности, чем к “Критике чистого разума”. Мы сознательно не форсируем эту вторую параллель.

(c) Феномен и частичный след. У Канта то, что дано восприятию - столы, стулья, звёзды, - суть явления (Erscheinungen): продукт обработки непознаваемой вещи-в-себе синтезирующей деятельностью рассудка. В теореме классически определённая картина мира, которую видит встроенный наблюдатель, - продукт операции частичного следа, принудительно стирающей информацию о R. И там, и там определённая воспринимаемого мира не дана непосредственно, а конституирована структурным ограничением способа доступа к целому: представление предстаёт не как зеркало реальности, а как её функция от точки зрения.

3. Что теорема не показывает

Дисциплина требует очертить границы аналогии не менее тщательно, чем саму аналогию, иначе изоморфизм рискует незаметно соскользнуть в метафизическую спекуляцию.

Теорема не касается онтологии. Она не утверждает и не отрицает существование объективной реальности как таковой; она говорит исключительно об ограничениях эпистемического доступа наблюдателя, встроенного в конкретный формальный класс систем. Её ближайший философский родственник - не радикальный идеализм и не наивный реализм, а нечто вроде структурного реализма (Уорролл, 1989; Ладимен и Росс, 2007): позиции, согласно которой познанию доступна именно и только *структура* отношений, тогда как вопрос о внутренней природе релятов остаётся открытым или вовсе снимается как некорректно поставленный.

Теорема не имеет отношения к вопросу о существовании трансцендентного начала. Формальная область действия теоремы ограничена наблюдателем как собственной подсистемой замкнутой системы. Всё, что не подпадает под это условие, то есть не является частью описываемой физической системы, - просто вне сферы её действия. Классические теистические концепции, настаивающие на трансцендентности божественного знания по отношению к сотворённому порядку, формально не противоречат теореме ни в одном пункте; если что, они

структурно ближе к её логике, чем позиция, пытающаяся поместить какое бы то ни было всеведущее начало внутрь системы как ещё одну подсистему, - именно это теорема и исключает. Вопрос о существовании такого трансцендентного начала лежит в иной, метафизической и теологической плоскости, куда физическая теорема о свойствах embedded-наблюдателей не проникает ни в одну, ни в другую сторону.

Теорема - не дедуктивное доказательство Канта. Структурный изоморфизм не эквивалентен логическому выводу одного из другого. Кантовский тезис об априорности пространства и времени обоснован трансцендентальной дедукцией - анализом условий возможности синтетических суждений a priori, - а не физическим экспериментом или теоремой; он логически автономен по отношению к тому, воспроизводится ли аналогичная архитектура в некотором классе формальных физических моделей. Показательно в этом отношении программа Мишеля Битболя (1998), явно поставившего вопрос: можно ли дать «трансцендентальную дедукцию» самой квантовой механики, обобщив кантовский метод так, чтобы вывести часть её формальной структуры из условий возможности эмпирического знания, минуя как реализм, так и чистый эмпиризм. Это дисциплинированный, методологически осторожный проект, в отличие от него настоящая теорема не претендует на трансцендентальную дедукцию квантовой механики как таковой; она значительно скромнее: она показывает, что в рамках уже принятой квантовой механики схема, изоморфная кантовской, принудительно возникает для любого встроенного наблюдателя. Это интересный факт независимо от того, прав ли был Кант в собственном, трансцендентальном смысле.

4. Место в философии науки

Если снять кантовскую рамку и взглянуть на результат в более широком контексте, он попадает в давно существующую линию рассуждений об относительности наблюдателя - физического описания, рядом с реляционной квантовой механикой Ровелли (1996), постулирующей, что состояние системы всегда относительно другой системы, а не абсолютно; рядом с QBism (Fuchs, Mermin & Schack, 2014), трактующим квантовое состояние как выражение убеждений конкретного агента, а не свойство мира самого по себе; рядом с дискуссиями о квантовых системах отсчёта (Джакомини, Кастро-Руис, Брукнер, 2019). Специфика именно этой теоремы в предельной экономности формулировки, минимальном наборе гипотез (замкнутость, встроенность, локальность, невырожденность взаимодействия) и в отказе апеллировать к какой-либо конкретной интерпретации квантовой механики или строить многоагентный мысленный эксперимент по образцу Фраухигер–Реннера.

Разрыв между тем, что доступно точке зрения изнутри, и тем, что было бы дано “взгляду ниоткуда” (view from nowhere, если воспользоваться формулой Нагеля) - вообще одна из сквозных тем философии XX–XXI века, далеко выходящая за пределы философии физики. Кентен Мейясу (2008) диагностирует всю послекантовскую философию как плененную “корреляционистским кругом” - невозможностью помыслить бытие иначе, чем как коррелят акта мышления о нём, и именно против этой пленённости строит свой спекулятивный реализм. Настоящая теорема не разрешает этот спор и не встаёт ни на одну из сторон, но интересно освещает его иным светом: она показывает, что нечто похожее на “корреляционистский круг” - принципиальную недоступность немедианного целого можно получить не как трансцендентальную предпосылку и не как метафизический тезис, а как строгое следствие того, что познающая система физически вложена в познаваемую. Это не довод в пользу корреляционизма и не довод против спекулятивного реализма - это независимый, физически мотивированный источник той же структуры, о которой спорят по существу иными средствами.

Родственные интуиции обнаруживаются и в теории вычислимости и кибернетике, хотя и по касательной, без прямого формального родства с настоящим результатом: теоремы Гёделя (1931) о неполноте формальных систем - отдалённый, но напрашивающийся структурный аналог невозможности полного самоописания изнутри; закон необходимого разнообразия Эшби (1956) в кибернетике - тезис о том, что регулирующая система должна обладать не меньшим разнообразием состояний, чем регулируемая, и потому не может её полностью объять, оставаясь её частью. Всё это вариации на общую тему: часть, стремящаяся описать целое, частью которого она является, наталкивается на структурный, не устранимый улучшением метода предел.

Заключение

Embedded-Observer No-Go Theorem - не физическое доказательство трансцендентальной философии и не аргумент за или против существования Бога. Это узкий, аккуратно доказанный технический результат о конкретном классе физических систем. Его философская ценность в другом: он даёт редкий случай, когда эпистемическая архитектура, сформулированная философией на языке трансцендентальной аргументации два с половиной века назад, обнаруживает точный структурный аналог в формальном аппарате современной квантовой информации. Это не подтверждает Канта и не заменяет философский анализ условий познания - но даёт философии науки повод отнестись к вопросу серьёзно: возможно, разрыв между целым и локально доступным это не случайность конкретной физической теории и не артефакт конкретной метафизической позиции, а нечто, принудительно возникающее всякий раз, когда познающее оказывается частью познаваемого.

Библиография

- Ashby, W. R. (1956). *An Introduction to Cybernetics*. London: Chapman & Hall.
- Bitbol, M. (1998). Some Steps Towards a Transcendental Deduction of Quantum Mechanics. *Philosophia Naturalis*, 35, 253–280.
- Bitbol, M. (2010). *Reflective Metaphysics: Understanding Quantum Mechanics from a Kantian Standpoint*. Paris: Éditions Rue d'Ulm/Presses de l'École normale supérieure.
- Bong, K.-W., Utreras-Alarcón, A., Ghafari, F., Liang, Y.-C., Tischler, N., Cavalcanti, E. G., Pryde, G. J., & Wiseman, H. M. (2020). A Strong No-Go Theorem on the Wigner's Friend Paradox. *Nature Physics*, 16, 1199–1205.
- Brukner, Č. (2018). A No-Go Theorem for Observer-Independent Facts. *Entropy*, 20(5), 350.
- Frauchiger, D., & Renner, R. (2018). Quantum Theory Cannot Consistently Describe the Use of Itself. *Nature Communications*, 9, 3711.
- Fuchs, C. A., Mermin, N. D., & Schack, R. (2014). An Introduction to QBism with an Application to the Locality of Quantum Mechanics. *American Journal of Physics*, 82, 749–754.
- Giacomini, F., Castro-Ruiz, E., & Brukner, Č. (2019). Quantum Mechanics and the Covariance of Physical Laws in Quantum Reference Frames. *Nature Communications*, 10, 494.
- Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 38, 173–198.

- Husserl, E. (1928). *Vorlesungen zur Phänomenologie des inneren Zeitbewusstseins*. Halle: Max Niemeyer.
- Kant, I. (1781/1787). *Kritik der reinen Vernunft* [Критика чистого разума]. Riga: Johann Friedrich Hartknoch.
- Ladyman, J., & Ross, D. (2007). *Every Thing Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press.
- Meillassoux, Q. (2008). *After Finitude: An Essay on the Necessity of Contingency* (R. Brassier, Trans.). London: Continuum. (Original work published 2006 as *Après la finitude*)
- Nagel, T. (1986). *The View from Nowhere*. Oxford: Oxford University Press.
- Page, D. N., & Wootters, W. K. (1983). Evolution Without Evolution: Dynamics Described by Stationary Observables. *Physical Review D*, 27, 2885.
- Rovelli, C. (1996). Relational Quantum Mechanics. *International Journal of Theoretical Physics*, 35(8), 1637–1678.
- Wigner, E. P. (1961). Remarks on the Mind-Body Question. In I. J. Good (Ed.), *The Scientist Speculates*. London: Heinemann.
- Worrall, J. (1989). Structural Realism: The Best of Both Worlds? *Dialectica*, 43(1–2), 99–124.
- Materov, S. (2026). *The Embedded-Observer No-Go Theorem: Causal Order, Geometrically Induced Reduction, and the Impossibility of Complete Self-Description in Any Closed Unitary Quantum System*. DOI <https://doi.org/10.5281/zenodo.21864707>.