Perspective Duality as a Physical Requirement

Wayne C. Myrvold[a]

[a]Department of Philosophy , The University of Western Ontario , wmyrvold@uwo.ca

Abstract

Philosophers of physics, when engaged in matters they regard as fundamental, tend to focus their analyses on fictitious systems that are wholly isolated from their environments. When pressed, they retreat to the fiction of treating the universe as a whole as an object of scientific study. This is nothing at all like the way science is practiced. Even if a system can be insulated in such a way that its interactions with its surroundings are negligible, one only ever explicitly models a minute fraction of the degrees of freedom of the system, and, as the degrees of freedom explicitly considered interact with those that are not, those degrees of freedom that are treated in the model constitute what is, in effect, an open system.

The thought behind this seems to be that this situation is a mere expedient, and that answers to all questions of import are to be found in the investigation of isolated systems. In this paper, I will argue that this is incorrect. Although J. S. Bell was correct to say that a fundamental physical theory should be formulable without recourse to terms such as *system*, *apparatus*, and *environment*, I will argue that any physical theory worth its salt should also be capable of *not* representing the entire universe, and should admit of a formulation with exogenous parameters. This leads to a stance I call *perspective duality.*

1 Introduction

There is a tension lurking in the literature on philosophy of physics, between two strands of thought. One strand, exemplified in J. S. Bell's "Against 'Measurement'", (Bell, 1990) contends that a fundamental theory ought to be capable of being formulated without a split between *system* and *environment.* Any such split, the thought goes, is a somewhat arbitrary distinction imposed by us, for convenience in our modelling, and corresponds to nothing in physical reality—hence, when speaking of the really deep issues, the questions about what the world is like, at the fundamental level, it must be dispensed with. This was, in part, a reaction against the insistence, from those infected with the Copenhagen virus, that quantum mechanics requires for its formulation a split between system and apparatus, with the apparatus treated classically, or worse, a split between system and *observer.* Against this, it has been emphasized that there are formulations of quantum theory, such as dynamical collapse theories and the de Broglie-Bohm pilot wave theory, that require no such split, and can be taken to apply to both macroscopic and microscopic objects, to both system experimented on and experimental apparatus, modelling the interactions between the two *within* the theory.

The other strand of thought is also found in Bell.

> A respectable class of theories, including contemporary quantum theory as it is practiced, have 'free' 'external' variables in addition to those internal to and conditioned by the theory. These variables are typically external fields and sources. They are invoked to represent experimental conditions. (Bell 1977, 81; 2004, 101)

To treat a variable as *external*, or *exogenous*, is to treat its value as simply given, not to be subjected to the dynamical laws of the theory being used to model the dynamics of the system under investigation. Whether or not a given parameter is exogenous is thus not a property of the physical situation, but how it is being treated in a given investigation. One and the same quantity may be treated exogenously in one investigation and endogenously in another. As an illustrative example, consider the familiar progression of textbook treatments of planetary motion. As a first pass, one considers the motion of a planet subjected to the Sun's gravitation, taken as a fixed gravitational potential unaffected by the planet. On such a treatment, the Sun's position is not treated as a dynamical variable, subject to the laws of motion, but is treated as fixed. As a next step one includes the Sun's position among the dynamical variables considered, and deals with the system as a two-body problem.

Exogenous variables—or, we should say, variables treated exogenously, to emphasize that the distinction is in our treatment, not in the nature of things—play two crucial, and related, roles in science. In statistical hypothesis testing, the choice of experiment, or choice of intervention on the system, is treated as an exogenous, or, as Bell would say, a free variable, and this is absolutely essential to statistical inference. In the causal modelling literature, it is common to take an interventionist approach, on which the effects of a change of a variable are studied by regarding the value of that variable as set exogenously.

This poses a problem. If the theory under investigation is a comprehensive theory of fundamental physics, applicable in principle to everything, what sense can we make of treating certain variables as exogenous? If the notion of *intervention* is essential for the application of causal concepts, does this mean that, as Pearl (2000, 350) put it, when the subject matter is the whole of reality, "causality disappears because interventions disappear—the manipulator and the manipulated lose their distinction"? And, if our experimental apparatus, including any devices (such as randomizers) used to determine the choice of treatment, are subject to the dynamical laws of our theory, what sense does it make to treat the outputs of such devices as 'free' variables? Some will say that this, too, makes no sense when fundamental physics is under investigation.

Things are not so dire, I will argue. The apparent tension can be resolved, and in their resolution, we gain insight into physical theorizing. In what follows, I will argue for what I call *perspective duality.* Bell is right that any candidate for a fundamental physical theory worth its salt should admit a formulation without a system/environment split. That is, it should be possible to adopt a *totalizing* perspective, an ideal point of view that would leave nothing out. But also—and this is something that has not, as far as I know, been articulated in the existing literature—to avoid the threat of making nonsense of statistical hypothesis testing, and of causal reasoning, a physical theory worth its salt should *also* admit of a formulation with exogenous variables. We should be able to adopt a non-totalizing, *limited* perspective. As we will see in section 7, the possibility of switching between these two perspectives, consistently, places a constraint on what a reasonable physical theory can be like.

2 Comprehensive theories with deterministic, invertible dynamics

The theories of the sort we'll be interested in are those with certain features that are usually assumed by those who imagine that they know what a fundamental physical theory would look like, if we had one.

The first is that there be no built-in limitation in the scope of a theory. Unlike theories of the special sciences, which deal with limited domains of reality, the theories we will be concerned with will be capable, at least in principle, of treating of the behavior of everything that there is. They will also (unlike, say, a theory of gravity or of electromagnetism) deal with all sorts of interactions. Call this feature *comprehensiveness*.

The dynamical laws of the theory will be assumed to be *deterministic*, and *invertible*. *Determinism* is the condition that a complete specification of the state of the world up to a given time, together with the theory's dynamical laws, uniquely determines its state at all future times. That is, if two dynamically possible complete histories of the world agree at given time t, they agree at all times to the future of t. Another way of saying this is that, if two complete histories of the world differ at some time t, they differ at all earlier times. In a relativistic spacetime, talk of the state at a time is to be replaced by talk of the state on a spacelike slice of spacetime that is a *Cauchy surface*, meaning that every past- and future-inextendible timelike curve intersects it exactly once. (We assume that we are dealing with a globally hyperbolic spacetime, that is, one that admits of Cauchy surfaces; otherwise, talk of determinism fails to make sense.)

Invertibility is the temporal inverse of determinism; it is the condition that a complete specification of the state of the world at a time, together with the theory's dynamical laws, uniquely determines its state at all past times. That is, if two dynamically possible complete histories differ at some time t, they differ at all times to the future of t; there is no forgetting of the past.

Determinism is often taken as the conjunction of what we are calling "determinism" and "invertibility." For our purposes it is useful to distinguish them, as the two conditions play distinct roles.

3 Physics and causal structure

In some quarters, at least, the causal structure of a physical theory is regarded as an important feature of the theory. In particular, having a causal structure compatible with relativistic spacetime structure is often taken to be a *desideratum* for a proposed physical theory, perhaps even a necessary condition for the theory to be taken into serious consideration. Relativistic structure (whether special or general) has the following features. The past of any event e is the set of event on or within its past light-cone, that is, the set of events from which a signal or other influence propagating at the speed of light or slower can reach the event. The future of event e is the set of events on or within its future light-cone, that is, the set of events that can be reached from e by a signal propagating at light-speed or slower. All other events are *spacelike separated* from e, neither to the past of e nor to the future of it. If, as is usual, we take the cause-effect relation to be a temporally asymmetric one, with causes temporally preceding their effects, compatibility with relativistic spacetime structure requires there to be no cause-effect relations between spacelike separated events. (Absent this assumption about temporal asymmetry, there is no reason to hold that relativity prohibits superluminal signalling or other cause-effect relations between spacelike separated events.) The constraint of relativistic causality is distinct from, and logically independent of, the requirement that the theory's dynamically laws be invariant under the symmetries of Minkowski spacetime, a requirement often called *Lorentz covariance*

In the case of a classical field theory, the requirement of relativistic causality is taken to amount to the requirement that the values of a field in a given spacetime region be determined by the values it takes on in its past light-cone, and that field values in two spacelike separated regions can be specified independently of each other. In the case of a quantum field theory, the requirement is that the theory satisfy a condition known as *microcausality.*

The physical content of the microcausality condition is the demand that a choice of what, if any, experiment to perform not affect the probability of outcomes of an experiment at spacelike separation. It is cashed out mathematically as follows.

As in nonrelativistic quantum mechanics, in a quantum field theory one associates operators with the dynamical variables ("observables") of a system. These operators can be added to each other, and one can multiply operators by scalars and by each other. Subject to appropriate conditions on these mathematical operations, the set of operators associated with a physical system form an *algebra.* In quantum theories the multiplication operation is, in general, noncommutative, meaning that the product of two operators may differ depending on the order. That is, for two operators $\hat{A}$, $\hat{B}$, $\hat{A}\hat{B}$ may differ from $\hat{B}\hat{A}$. This is what distinguishes quantum theories from classical ones. Of course, for some operators, the product could be independent of order. When $\hat{A}\hat{B}$ is equal to $\hat{B}\hat{A}$, the operators $\hat{A}$ and $\hat{B}$ are said to *commute.*

In a relativistic quantum field theory, for any region of spacetime with nonempty interior, there is a set of observables corresponding to outcomes of experiments that can be performed in the causal closure of that region. We will call the algebra of operators containing the operators corresponding to observables associated with a spacetime region $\mathcal{O}$, $\mathfrak{A}(\mathcal{O})$.

One typically models experiments as follows. It is assumed that, associated with the experiment to be performed, there is a set of operators $\{\hat{M}_i\}$, satisfying the condition that $\sum_i \hat{M}_i^\dagger \hat{M}_i$ is the identify operator, such that, when the experiment is performed on a quantum state represented by a vector $|\Psi\rangle$, the resulting is to produce, for some i, the state $\hat{M}_i|\Psi\rangle$, with probabilities proportional to $\|\hat{M}_i|\Psi\rangle\|^2$. This is the condition that the state-transformation correspond to what is called a *nonselective completely positive operation.*

The microcausality condition requires that these goings-on be irrelevant to probabilities of an experiment regarding some observable B, performed at spacelike separation from the A-experiment. And that means that, when computing probabilities for outcomes of a B-experiment, it shouldn't matter whether we use the original state vector $|\Psi\rangle$, or a weighted average of the probabilities computed using state vectors $\{\hat{M}_i|\Psi\rangle\}$, weighted by the probabilities of the corresponding outcomes. It is not difficult to show that the operation always leaves the probabilities of a measurement of an observable B unchanged if and only if the operator $\hat{B}$ commutes with every one of the operators $\{\hat{M}_i\}$ (see appendix of Myrvold 2021 for a simple proof).

The constraint of relativistic causality also has implications for proposed modifications of the usual quantum dynamics. One approach to the so-called "measurement problem" of quantum mechanics involves a modification of the dynamics to suppress superpositions of macroscopically distinct states. It can be shown that the no-signalling condition places non-trivial constraints on such modifications (see Gisin 1989; Simon, Bužek, & Gisin 2001).

4 The problem

Randomization posits employed in statistical hypothesis testing involve a division between the system under investigation and the randomizing device used. In the literature on causal inference, the notion of *intervention* plays a role, which involves treating the value of the variable intervened upon as exogenously given.

Some theories are inherently of limited scope, and may be such that we would not and could not expect them to have anything to say about the behaviour of the randomizing device and other experimental apparatus. A comprehensive physical theory has no such excuse.

Moreover, for a system that is not acted upon from the outside, a solution of the equations of motion specifies what happens at *all* times, past, present, and future. It is not clear what it would mean for one part of the system to *change* the state of another, or to talk of cause-effect relations, or sender-receiver signalling relations.

Take, for example, the case of signalling. We usually think of Bob, the sender, as having a choice of signals to send, and of Alice, the receiver, as learning, upon receipt of the signal, about the choice that Bob has made. But if Bob and Alice and their apparatus are included in our theoretical description, and if the dynamics are invertible, then any two solutions in which Bob sends different signals are different at all times to the past of the sending event, and this will include times to the past of Alice's reception. Because of determinism, a state to the past of Alice's reception determines what Alice will receive. That is, a state to the past of both events determines both Bob's choice of signal and what Alice receives. How, then, are we to make sense of talk of Bob's choice of signal having an effect on what Alice receives?

This worry has, in fact, been expressed, in connection with the usual sorts of rationale for the microcausality constraint on quantum field theories. Laura Ruetsche, after introducing a proposed motivation for the constraint, along the lines of what we discussed in section 3, writes,

> I take this motivation to fail because I take one of its presuppositions to be incompatible with *other* axioms of Axiomatic QFT. The presupposition is that the way to model an operation performed in the spacetime region $\mathcal{O}$ is by the change of state wrought by acting on the global state $|\Psi\rangle$ with some element A of $\mathfrak{A}(\mathcal{O})$. My tendentious contention is that, whatever one thinks of measurement in non-relativistic QM, there's no way to square this model with the idea…that the quasilocal algebra $\mathfrak{A}(\mathcal{M})$ admits a state.[1] If $|\Psi\rangle$ is that global state, it's a state for all space and all time, and the notion of changing it by an operation executed within spacetime is nonsense. If $|\Psi\rangle$ is not that global state, we need an account of what state $|\Psi\rangle$ is, an account that lends coherence to the idea that acting on $|\Psi\rangle$ with an element of the local

---

[1] Recall, from §3 above, that $\mathfrak{A}(\mathcal{O})$ is the algebra of operators associated with spacetime region $\mathcal{O}$, which contains operators pertaining to observables measurable by operations in the causal closure of $\mathcal{O}$. The quasilocal algebra $\mathfrak{A}(\mathcal{M})$ is the algebra associated with the whole of Minkowski space; it's the smallest algebra that contains all the algebras associated with bounded regions of spacetime.

> algebra $\mathfrak{A}(\mathcal{O})$ corresponds to carrying out a local operation performable in $\mathcal{O}$ (Ruetsche 2011, 110; see also Ruetsche 2021, 319).

Ruetsche goes on to express further worries about arguments along the lines we have considered, having to do with prohibition of action-at-a-distance. Since such arguments involve modelling a change of the quantum state brought about by an operation local to a region $\mathcal{O}$ via operation on the state vector with an operator in the local algebra $\mathfrak{A}(\mathcal{O})$, the worry expressed in the quoted passage applies also to them.

Similar concerns have been expressed by others, about how to make sense of intervention when the system being modelled is the entire universe; see Pearl (2000, 350), Hitchcock (2007, 53–54), Woodward (2007, 93).

5 Wrong answers

There are a number of ways to respond to this worry. In this section I will present some that strike me as steps in the wrong direction.

The first might be called *exceptionalism*. This is the idea that there are certain kinds of things that are exempt from being treated as physical objects, subject to physical laws. The most common sort is that conscious beings are endowed with special powers that transcend the laws of physics.

This would be a radical move, unwarranted by anything we know about the behaviour of the bodies (including the brains) of conscious beings, which suggests, on the contrary, that physical principles *can* fruitfully be applied to them. It would also fail to solve the problems raised, as they arise also when considering the effects of one physical system on another, with no conscious beings involved.

Another proposal that might be made is that, though a comprehensive physical theory ought to be capable of embracing *anything*, it makes no sense to talk of simultaneously modelling *everything*; scientific investigation always involve a split between system under investigation, and system doing the investigation. This position is less obviously problematic than the position that there are some things that are forever exempt. Nonetheless, as I will argue below, it is more extreme than the logic of the situation dictates, and there is a more modest escape from the difficulties.

Another move is what might be called *eliminativism*. Bertrand Russell famously remarked, more than a century ago, "The law of causality, I believe, like much that passes muster among philosophers, is a relic of a bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm" (Russell, 1913, 1). In a similar vein, John Norton, more recently, has written,

> Mature sciences, I maintain, are adequate to account for their realms without need of supplement by causal notions and principles. The latter belong to earlier efforts to understand our natural world or to simplified reformulations of our mature theories, intended to trade precision for intelligibility. (Norton, 2007, 12)

And Ruetsche:

> Soberly viewed all the special theory of relativity (STR) can be taken to demand of a theory set in Minkowski spacetime is that it exhibit Lorentz covariance. And this demand is met by QFTs in its scope, provided they

> satisfy the Covariance axiom—*even if they violate the Microcausality axiom!* (Ruetsche, 2011, 109)

On this view, causal notions, though they may be useful in everyday life and in the special sciences, are to be dispensed with in fundamental physics.

Is this right? When we were children, we spoke as children, understood as children, thought as children. Now that we have matured, is causality one of those childish things to be put away?

This move strikes me as throwing the baby out with the bathwater. For one thing, as we have seen, causal constraints are fruitfully employed in construction and evaluation of physical theories. The requirement of relativistic causality imposes constraints not entailed by Lorentz covariance, and, in the context of the closest things we have to fundamental theories, quantum field theories, it has teeth; it requires that the microcausality axiom be satisfied, which, as Ruetsche notes, is a requirement independent of covariance.

It has also been proposed that in the context of fundamental physical theories, independence posits of the sort prevalent in statistical hypothesis testing be abandoned. This proposal has arisen in the context of experimental tests of Bell inequalities, the results of which, on the usual interpretation, are taken to indicate the reality of some kind of nonlocality in the physical world, albeit a sort that is arguably not at odds with the requirements of relativistic causality.[2]

It was noted by Shimony, Horne, & Clauser (1976) that there was an assumption required for derivation of Bell inequalities that up to that point had been implicit but not explicit in Bell's derivations. This is the assumption that the choice of experiment to be performed can be made in such a way that it is statistically independent of the state of the system to be experimented on, an assumption that has been called the assumption of *measurement independence*, or, somewhat misleadingly, the *free will assumption.* In recent literature it is often referred to as the assumption of *statistical independence*, terminology that is less than ideal, as it fails to specific what variables are to be taken as statistically independent. The rejection of the measurement independence assumption has come to be known as *superdeterminism.*

Shimony, Horne, and Clauser (SHC) illustrate the point with a tale about a conspiracy involving manufacturers of experimental apparatus and experimenters' assistants. The director of the conspiracy concocts a set of experimental data, consisting of a sequence of experimental settings and results obtained. The experimenters' assistants arrange things so that the experiments performed are the pre-arranged ones, and the experimental apparatus is pre-programmed to yield the stipulated results. Obviously, any set of results can be used in such a scheme, and one can obtain violations of the Bell inequalities without making use of any physical nonlocality.

SHC point out that the assumption of measurement independence is not required by any sort of relativistic causality condition.

> After all, the backward light cones of those two acts do eventually overlap, and one can imagine one region which controls the decision of the two experimenters who chose a and b. We cannot deny such a possibility. But we feel that it is wrong on methodological grounds to worry seriously about it if no specific causal linkage is proposed. In any scientific experiment in which two or more variables are supposed to be randomly selected, one can always conjecture

---

[2] See Myrvold (2016) for an argument of this sort.

> that some factor in the overlap of the backwards light cones has controlled the presumably random choices. But, we maintain, skepticism of this sort will essentially dismiss all results of scientific experimentation. Unless we proceed under the assumption that hidden conspiracies of this sort do not occur, we have abandoned in advance the whole enterprise of discovering the laws of nature by experimentation. (Shimony et al. 1976, 6; Shimony 1993, 168)

The conspiracy need not involve the plotting of agents; one can (perhaps) imagine theories on which all interactions are local but initial conditions are constrained to produce the desired experimental settings and results. In order for a theory of this sort to be empirically equivalent to standard quantum mechanics, this must hold *no matter what* randomization method is employed. For example, in the experiment of Shalm *et al.* (2015), the measurement decisions were determined by applying an XOR (exclusive or) operation to three bits from three independent sources, one of which was a pseudorandom sequence constructed from digits of π and binary strings derived from various popular culture sources, including *Monty Python and the Holy Grail*, the *Back to the Future* trilogy, and episodes of *Doctor Who*, *Saved by the Bell*, and *Star Trek*. A hypothetical cause that achieved the observed statistics via correlation between states of the photons studied and the choice of measurements would have to orchestrate the creative processes leading up to the digitized versions of these cultural artifacts in such a way that, when processed in conjunction with the outputs of the random number generators, produced just the right sequence of experimental choices.

SHC are correct that measurement independence assumptions underly virtually all experimental science. Any conclusion whatsoever drawn from statistical hypothesis testing could be undermined by a rejection of measurement independence. For that reason, SHC argue, a methodology that countenanced rejection of a measurement independence assumption whenever its employment would lead to an unwanted conclusion amounts to an abandonment of the enterprise of investigating the world by experimentation. Nonetheless, superdeterminism has found some advocates; see Myrvold, Genovese, & Shimony (2024) for an overview and references.

I think that SHC are right that we should *not* adopt a methodology according to which it is acceptable to conclude a failure of our randomizers to do their job whenever the ordinary canons of statistical hypothesis testing would lead to an unpalatable conclusion, in the absence of independent reason to suspect a failure of randomization. Such a methodology would be a blunt instrument that could be used to undermine any conclusion one wishes to reject, and would amount to a violation of Peirce's maxim, "Do not block the way of inquiry" (Peirce, 1931, ¶135).

6 The solution

To understand how to deal with these worries, let us consider again the concern raised by Ruetsche in connection with quantum field theories. A global state yields expectation values for all observables, distributed over all of spacetime. The idea of changing this state by an operation executed within spacetime, it is said, is nonsense.

The first thing to note is that there is nothing at all distinctively *quantum* about this argument. If cogent, it applies equally well to, say, classical field theories. A solution of the field equations of a classical field theory is a global solution, yielding field values distributed over all of space and time. We could replace talk of a global state, in quantum field theory, with talk of a global solution of the field equations, in a classical theory, and the worry would retain its full force.

And this gives us a key to the solution of the puzzle. We do, indeed, apply causal notions to classical field theories; part of the content of a course on classical electromagnetism consists of learning how to generate and manipulate electromagnetic waves! This is possible because, though the theory is capable of treating of the behaviour of charges and currents within it, one can also examine solutions of the field equations with exogenously given charges and currents. So, for example, to examine the effect of a given oscillating current as a source of electromagnetic waves, one can compare two solutions: one in which the current is absent, for all time, and one in which it is endowed with a certain history, which might involve being zero up until some time at which it is switched on. One can then consider solutions to the field equations with the current as a source term, selecting retarded solutions, that is, solutions in which the field at a given spacetime point depends on charge-current distribution within the past light-cone of that point. There are, of course, also advanced solutions, in which the field depends on the charge-current distributions within the *future* light-cone, but these are commonly rejected as unphysical, and textbooks often explicitly invoke considerations of causality in taking the physical solutions to be the retarded solutions.

One can compare solutions of the field equations with and without the source charges and currents. The difference between these solutions is the effect of the sources.

What has been said about classical field theories holds equally well for quantum field theories. Just as in the classical case, quantum field theories admit of formulations in which source fields or charges are treated exogenously, as classical (“c-number”) terms added to the Lagrangian of the theory. One can, therefore, compare a global solution that lacks the source terms with one in which the source is zero up until some point at which it is turned on, and, just as in the classical case, take the effect of the source term on the quantum field as the difference between a global state that lacks the sources and a global state of the field with source term included. Here, again, one has a distinction analogous to retarded and advanced solutions, in the form of retarded and advanced propagators.

Thus, there is a respectable class of theories, which includes both classical and quantum field theories, that admit of formulations with exogenous source terms. Moreover, it is *essential* to certain sorts of applications that a theory admit of a formulation in which certain parameters are treated exogenously. This, as we have seen, is essential for statistical hypothesis testing, and for the employment of causal considerations.

I propose the following adequacy condition for a physical theory:

> Even if one accepts that a physical theory with aspirations to comprehensiveness ought to be capable in principle of delivering a global representation of the entire universe, leaving nothing out, any physical theory worth its salt should also be capable of *not* representing the entire universe, and should admit of a formulation with exogenous parameters.

The formulation with exogenous parameters might be limited to certain regimes of the theory’s space of possibilities. It is a requirement on a comprehensive physical theory that it admit the possibility of systems that act appropriately like our experimental apparatus. It is *not*, of course, a requirement that every solution of the theory contain such things. One might, for example, have a theory of quantum gravity according to which a spacetime description is possible only if certain conditions are met, and then at a suitably course-grained level. The proposed condition is that the theory be capable of admitting exogenous parameters in the regimes in which we need to employ them.[3]

I am tempted to elevate this into an *a priori* condition for the possibility of physical theorizing, but, in light of the chequered history of such proposals, confine myself to a more cautious formulation, echoing what Einstein said of the principle of separability: without it, "physical thought in the sense familiar to us would not be possible."[4] This leaves open the possibility of some hitherto unthought-of, unfamiliar modes of physical theorizing.[5]

So, for example, were someone to propose a viable superdeterministic theory that accounted for quantum phenomena via local means, *and* if an account were given of why one should take that theory seriously that was not vulnerable to SHC's charge that one is simply rejecting the usual canons of scientific inquiry because they lead to undesirable conclusions, then we would have motivation for reconsidering the proposed adequacy condition. The attitude expressed by Bell seems to be a judicious one.

> Of course it might be that these reasonable ideas about physical randomizers are just wrong — for the purpose at hand. A theory may appear in which such conspiracies inevitably occur, and these conspiracies may then seem more digestible than the non-localities of other theories. When that theory is announced I will not refuse to listen, either on methodological or other grounds. But I will not myself try to make such a theory. (Bell 1977, 83; 2004, 103)

I am skeptical that it is possible to produce a theory like that *and also* answer SHC's charge. But, like Bell, if someone were to make such a proposal, I would not refuse to consider it.

7 Perspective duality, and the nature of randomizers

Suppose we have some physical system S that is to be manipulated and experimented upon via some piece of apparatus A. On our proposal, a physical theory ought to be capable of treating the combined system $S + A$ as a physical system in its own right, subject to the dynamics of the theory, and of *also* treating just S, with the influence of the apparatus treated as exogenous. Call these two sorts of treatment the *totalizing* and *limited* perspectives, respectively.

Suppose we are interested in the effect on the system of choosing between two parameter settings, a and b. Then we can treat the parameters as exogenous variables, and contrast solutions to the equations of motion for the system S with setting a, and with setting b. These will be global solutions that agree up to the time at which the parameters are set, and differ only after this.

We should be able to switch to a totalizing perspective, and include the apparatus in our physical description. The two modes of description should mesh.

---

[3] I am grateful to an anonymous referee for raising this point.

[4] "Ohne die Annahme einer solchen Unabhängigkeit der Existenz (des «So-Seins ») der räumlich distanten Dinge voneinander, die zunächst dem Alltags-Denken entstammt, wäre physikalisches Denken in dem uns geläufigen Sinne nicht möglich" (Einstein, 1948, 321). I quote from the translation of Don Howard (1985, 187).

[5] Some have attributed a less cautious view to Einstein, but that's not his fault.

Exactly how this can work in the context of quantum field theory will be the subject of a sequel, but the gist of such an account will be as follows. We should be able to model the apparatus, at least schematically, within the theory. A limited perspective is obtained by tracing out the degrees of freedom of the apparatus. The presupposition motivating the microcausality axiom is that, if experimental apparatus couples to the fields under study only in a bounded spacetime region $\mathcal{O}$, its effect on the state of the fields can be modelled by a set of operators $\{\hat{M}_i\}$ belonging to $\mathfrak{A}(\mathcal{O})$. This should be a consequence of modelling the apparatus within the theory as something that couples locally to the field, and then tracing out the degrees of freedom of the apparatus.

We can compare two solutions pertaining to differing values of experimental parameter settings; the differences between the two solutions are to be counted as the effects of varying the parameters. An approach of this sort lends insight. Rafael Sorkin (1993) has shown that simply importing a measurement postulate into quantum field theory borrowed from non-relativistic quantum mechanics, to the effect that for any observable corresponding to a bounded spacetime region there is a corresponding experiment whose effect is projection onto an eigenstate of the observable, can lead to superluminal signalling. If one holds that superluminal signalling is impossible, measurements of that sort are impossible, but some principled reason is needed for *why* one should regard them as impossible. The answer is that modelling the apparatus within the theory, as coupling locally to the fields, eliminates such impossible measurements (see Papageorgiou & Fraser 2024 and references therein).

There's a potential obstacle to the meshing of the two modes of description. If the dynamics are deterministic, there are no solutions that differ only to the future of the setting event; if two solutions differ in the parameter settings, they differ at all times.

Here we can avail ourselves of another principle of physical theorizing. It should be capable of making empirically testable predictions (of either a deterministic or probabilistic sort). As it is never possible to obtain *exact* values of all the relevant parameters, any such predictions will have to be, at least to a certain extent, robust under approximation. And, though invertible dynamics mean that there can be no pair of exact solutions that agree up to a certain time and subsequently diverge, there *can* be solutions that *approximately* agree up to a certain time and subsequently diverge to a significant extent. This means that even a slight degree of coarse-graining suffices for the possibility of two coarse-grained solutions that agree up to a certain point and thence diverge.

This is the phenomenon of sensitive dependence on initial conditions. It is a familiar fact that certain sorts of systems with deterministic dynamics are *effectively* unpredictable. That is, though an exact specification of the state at a given time uniquely determines future states, an approximation to the state at a given time may be compatible with very different states at future times. If the device responsible for the instrument settings exhibits this sort of sensitive dependence, it is possible for two solutions to agree to a high degree of approximation up to a certain point, and nevertheless contain different parameter settings, and to differ in all the consequences of these different settings.

This is, of course, true of the devices that we think of as paradigmatic randomizers. A properly flipped coin, or a die thrown, exhibits exquisite dependence on initial conditions, in such a way that information about the state of the world that is empirically accessible prior to the coin flip or dice toss is useless for prediction of the outcome.

This, then, is one way to reconcile a limited perspective, in which two solutions are compared that agree up to a point and then differ as a result of differing values of some exogenous variables, and a totalizing perspective that includes the experimental apparatus in the description. If we take the totalizing description to be one involving, not an exact specification of state, but some degree of coarse-graining, we can have, despite determinism at the level of fine-grained dynamics, two histories that agree up to a certain point and thence diverge.

Not every device that is used as a randomizer is of the effectively unpredictable sort. We also use pseudo-random number generators, whose output is predictable, but expected to be uncorrelated with the state of the system being experimented on. This feature is more important than unpredictability of the output of the randomizer.

Whence this lack of correlation? We might invoke a common-cause principle, to the effect that there be no correlations between the states of things that have not interacted, either directly or via some third mediating system. But this, by itself, would be empty, unless we imagine a randomizer that has somehow been kept for all eternity in isolation not only from the target system but also from anything that will subsequently be linked via a causal chain to the target system.

Isolation of this sort is an unsatisfiable, unreasonable, and unnecessary demand. Our conviction that we can produce outputs of a randomizer that are effectively uncorrelated with the state of the system being experimented upon is based, not on isolation of the randomizing devices, but on the opposite, on the idea that the systems are subjected to influences that effectively efface any correlations between the device and the target system. The idea is that the dynamics of the randomizing device render events in the common past of the device and the target system irrelevant to its output. They are, as J.S. Bell has put it, 'forgetting machines.'

> In such a device the complete final state perfectly determines the complete initial state — nothing is forgotten. And yet for many purposes, such a device is precisely a 'forgetting machine'. A particular output is the result of combining so many factors, of such a lengthy and complicated dynamical chain, that it is quite extraordinarily sensitive to minute variations of any one of many initial conditions. It is the familiar paradox of classical statistical mechanics that such an exquisite sensitivity to initial conditions is practically equivalent to complete forgetfulness of them. To illustrate the point, suppose that the choice between two outputs, a and a', depended on the oddness or evenness of the digit in the millionth decimal place of some input variable. Then fixing a or a' indeed fixes something about the input — i.e. whether the millionth digit is odd or even. But this peculiar piece of information is unlikely to be the vital piece for any distinctively different purpose, i.e., it is otherwise rather useless. With a physical shuffling machine, we are unable to perform the analysis to the point of saying just what peculiar feature of the input is remembered in the output. But we can quite reasonably assume that it is not relevant for other purposes. In this sense the output of such a device is indeed a sufficiently free variable for the purpose at hand. (Bell 1977, 82–83; Bell 2004, 102–102)

As Bell notes, this phenomenon is related to the process of relaxation to equilibrium in statistical mechanics. This is a process in which there is a failure of invertibility at the macroscopic scale. For many systems there is, for given values of total energy and the external parameters that define the thermodynamic state of a system, a unique equilibrium state that the system relaxes to, independent of the initial state. This is to be contrasted with invertibility at the level of microphysical dynamics, which preserves distinguishability of states. The project of explaining the tendency of systems to equilibrate can be thought of as the task of reconciling preservation of memory of the initial state at the level of microdynamics with macroscopic forgetfulness (see Myrvold 2020 for further discussion of this point).

8 Conclusion

The worries presented can be addressed. Moreover, it seems to me, seeing how they are to be addressed lends insight into the nature of scientific theorizing. Those who have stressed that a fundamental physical theory should be capable of encompassing everything in its embrace, and should be formulable without reference to as split between system and observer, or system and apparatus, are right. But, in order for it to be applicable in anything like the way we do, in fact, apply physical theories, a theory should *also* be capable of *not* embracing everything at once in its description.

9 Acknowledgments

Preliminary versions of this work were presented at a symposium, *Randomness in Scientific Practice*, at the 27th Biennial Meeting of the Philosophy of Science Association. November 2021, and at the *Causalworlds* conference, at the Perimeter Institute for Theoretical Physics, September 2024. I am grateful to the discussants at both events for helpful feedback. I am particularly grateful to Marshall Abrams and Doreen Fraser for helpful discussions.

References

Bell, J. S. (1977). Free variables and local causality. *Epistemological Letters*, *15*, 79–84. Reprinted in *Dialectica* **39** (1985), 103–16, and in (Bell, 2004, 100-104).

Bell, J. S. (1990). Against 'measurement'. *Physics World*, *3*, 33–40. Reprinted in Bell (2004, 213–231).

Bell, J. S. (2004). *Speakable and Unspeakable in Quantum Mechanics*. Cambridge University Press, 2nd ed.

Einstein, A. (1948). Quanten-Mechanik und Wirklichkeit. *Dialectica*, *2*, 320–324.

Gisin, N. (1989). Stochastic quantum dynamics and relativity. *Helvetica Physica*, *62*, 363–371.

Hitchcock, C. (2007). What Russell got right. In Price & Corry (2007), (pp. 45–65).

Howard, D. (1985). Einstein on locality and separability. *Studies in History and Philosophy of Science*, *16*, 171–201.

Knox, E., & Wilson, A. (Eds.) (2021). *The Routledge Companion to Philosophy of Physics*. Routledge.

Myrvold, W., Genovese, M., & Shimony, A. (2024). Bell's Theorem. In E. N. Zalta, & U. Nodelman (Eds.) *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2024 ed.

Myrvold, W. C. (2016). Lessons of Bell's theorem: Nonlocality, yes; action at a distance, not necessarily. In M. Bell, & S. Gao (Eds.) *Quantum Nonlocality and Reality: 50 Years of Bell's Theorem*, (pp. 238–260). Cambridge University Press.

Myrvold, W. C. (2020). Explaining thermodynamics: What remains to be done? In V. Allori (Ed.) *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*, (pp. 113–143). World Scientific.

Myrvold, W. C. (2021). Relativistic constraints on interpretations of quantum mechanics. In Knox & Wilson (2021), (pp. 99–121).

Norton, J. D. (2007). Causation as folk science. In Price & Corry (2007), (pp. 11–44).

Papageorgiou, M., & Fraser, D. (2024). Eliminating the 'impossible': Recent progress on local measurement theory for quantum field theory. *Foundations of Physics*, *54*, 26.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Peirce, C. S. (1931). *The Collected Papers of Charles Sanders Peirce, Volume 1: Principles of Philosophy*. Cambridge: Harvard University Press.

Price, H., & Corry, R. (Eds.) (2007). *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*. Oxford: Oxford University Press.

Ruetsche, L. (2011). *Interpreting Quantum Theories*. Oxford University Press.

Ruetsche, L. (2021). Locality in (axiomatic) quantum field theory: A minority report. In Knox & Wilson (2021), (pp. 311–322).

Russell, B. (1913). On the notion of cause. *Proceedings of the Aristotelian Society, New Series*, *13*, 1–26.

Shalm, L. K., *et al.* (2015). Strong loophole-free test of local realism. *Physical Review Letters*, *115*, 250402.

Shimony, A. (1993). *Search for a Naturalistic Worldview, Volume II: Natural Science and Metaphysics*. Cambridge University Press.

Shimony, A., Horne, M. A., & Clauser, J. F. (1976). Comment on "The Theory of Local Beables". *Epistemological Letters*, *13*, 1–8. Reprinted in *Dialectica* **39** (1985), 97–102, and in Shimony (1993, 163–167).

Simon, C., Bužek, V., & Gisin, N. (2001). No-signalling condition and quantum dynamics. *Physical Review Letters*, *87*, 170405.

Sorkin, R. D. (1993). Impossible measurements on quantum fields. In B. L. Hu, & T. A. Jacobson (Eds.) *Directions in General Relativity: Proceedings of the 1993 International Symposium, Maryland, Volume 2: Papers in Honor of Dieter Brill*, (pp. 293–305). Cambridge University Press.

Woodward, J. (2007). Causation with a human face. In Price & Corry (2007), (pp. 66–105).