Published in *Symmetries in Physics: Philosophical Reflections*, K. Brading and E. Castellani, eds., Cambridge University Press, 2003

Physics and Leibniz's Principles

Simon Saunders

Leibniz's principles made for an elegant and coherent philosophy. In part metaphysical, in part methodological, they addressed fundamental questions - in the treatment of symmetry, in the relationship of physics to mathematics, in logic - that are if anything even more pressing today than they were in Leibniz's time. As I shall read them, they also expressed a distinctive and uncompromising form of realism, a commitment to the adequacy of purely descriptive concepts. This doctrine has been called 'semantic universalism' by van Fraassen (1991), and the 'generalist picture' by O'Leary-Hawthorne and Cover (1996): it will become clearer in due course just what it entails.

The principles that I shall consider are the Principle of Sufficient Reason (PSR) and the Principle of Identity of Indiscernibles (PII). In the first instance I shall take them both to be methodological principles. The former I shall read as requiring that the concepts of physics be entirely transparent. Analysis and explanation are to proceed without any limits. The perspective is impersonal: any epistemological limitation, to do with our human situation or perceptual apparatus, is to be viewed as a purely practical matter, reflecting no fundamental constraint. This puts in place a part of the generalist picture.

The PSR clearly promotes the use of mathematical concepts in physics. The PII, in contrast, depends on a sharp distinction between *purely* mathematical concepts, and physical ones. Leibniz too made use of this distinction (between 'real' and 'notional' or 'ideal' concepts), but in his hands the principle depended heavily, though often tacitly, on his metaphysical theory of substance (and, with qualifications, on his philosophical logic). He was led to a restrictive formulation of it in consequence:

> There are never in nature two beings which are perfectly alike and in which it would not be possible to find a difference that is internal or founded upon an intrinsic denomination. (Leibniz, 1714, §9.)

In due course we shall come across two possible candidates for what Leibniz called 'intrinsic' denominations. But there have been major changes in logic since Leibniz's time, and a purely metaphysical theory of substance is unlikely to command much assent today: we shall find little reason to plump for either of them.

There is a widespread view that, apart from the trivialization of the PII whereupon identity is made out in terms of predicates that themselves involve identity, there are straightforward exceptions to every version of this principle.

It is a mistaken view; as we shall see, there is a natural analysis of identity available for any formal language that is immune to the usual counter-examples; the principle is not, I hold, in any difficulties from this quarter. The problem, rather, concerns the *justification* for the PII - why embrace such a principle? What is wrong with identity taken as primitive?

In the most general context, I see nothing wrong with identity. But in physics - specifically identity as it figures in physical *theory* - there are special reasons to view it as derivative. I take we are concerned with physical objects in the logical sense, as objects of predication. I suggest it is through *talk* of objects, in the light of mathematical theories and experiments, that we achieve a clear interpretation of these theories and experiments in terms of physical objects - our understanding of what objects there are, I am suggesting, is clearest in our use of simple declarative sentences. And it is here that purely formal, logical considerations come into play; Quine and the logical empiricists had something important to say in this respect. But my suggestion is not that physical theories should be *reconstructed* in a formalized language (they should not be *rewritten,* as a construction in set theory). What I have in mind is *description*, as informed by theory, in predicative terms. (This puts in place the other part of the generalist picture.)

Taking this route, our first concern is with syntax. Here, I suggest, the non-logical symbols of the language - for simplicity I shall consider only finite, first-order languages - can be derived more or less directly from the physical theory. They are to be interpreted in terms of the real physical functions, properties and relations. Our guide here, as for Leibniz, lies in the measurable quantities. Not so identity, as the relation that every physical object has to itself and to no other. It would be hard to imagine a quantity whose measurement could tell us about this directly. Nor is the identity relation itself under investigation in physical theorizing, unlike measurable properties and relations (in this sense it is not treated as a *physical* relation at all). From a formal point of view, the mathematics used in physics is far away from set theory, and still further from formal logic: the identity sign, as it figures in extant physical theories, signifies only the equality or identity of mathematical expressions, not of physical objects.

In summary, we may read off the predicates of an interpretation from the mathematics of a theory, and, because theories are born interpreted, we have a rough and ready idea of the objects that they are predicates of. But there is nothing systematic to learn from the formalism to sharpen this idea of object. It is plausible, in this situation, that we should look to a purely logical aid.

It is the fact that there is an essentially unique prescription for how to use the identity sign, available for any formal language whose predicates do not involve identity, that now is really telling: this is the chief selling-point for the PII as I shall understand it. Indeed, given a finite lexicon, this prescription even generates an explicit *definition* of identity.

I shall first sketch the details, and then show how the principle fares in the face of the usual scenarios offered as counter-examples to Leibniz's principles. The generalist picture is also supposed to be in trouble in these contexts; that too will need some defending. To proceed from that, we shall need the rudiments of Leibniz's theory of possible worlds; as we shall see it can be taken over for our purposes with little change. The most important question, from that point on, is how we are to distinguish the real physical quantities of a theory from the purely mathematical ones. Leibniz too needed this distinction, but he fell back on a fairly crude form of verificationism. Measurable quantities will be our guide as for Leibniz, but they are only the starting point of our analysis: symmetries are the essential tool for going beyond them. At the end I will return to the PSR, and its relationship to the PII.

1 Identity

How, in the interpretation of a physical theory, is identity to be analyzed in terms of other properties and relations, which do not themselves involve identity? There is a canonical answer to this question. Given the simplest case of a language with only finitely-many predicates, for each $n \leq N$, let there be K_n n-ary predicate symbols $P_1^n, P_2^n, ..., P_{K_n}^n$. Now let '$s$' and '$t$' be terms (variables, names, or functions of such). The familiar axiom scheme for identity is:

$$s = s$$

$$s = t \rightarrow (Fs \rightarrow Ft) \tag{1}$$

where F is any predicate of the language - expressing, essentially, the substitutivity of identicals. As Gödel showed, a complete proof procedure for the predicate calculus without identity, supplemented by this scheme, yields a complete proof procedure for the predicate calculus *with* identity (Gödel 1930 Th.VII). It is therefore enough, from the point of view of completeness, to take the conjunction of every instance of (1) as *sufficient* for identity. In the case of 1-place predicates, we obtain formulae of the form

$$P_i^1 s \leftrightarrow P_i^1 t \tag{2}$$

for $1 \leq i \leq K_1$. If it is right to read '1-place predicate' for 'intrinsic denomination', this would clearly do as a formalization of the principle of identity as Leibniz stated it (Eq.(2) is often called the *strong* version of his principle). In the case of 2-place predicates, generalizing on the free variable that remains in instances of (1) we obtain

$$\forall z_1((P_i^2 s z_1 \leftrightarrow P_i^2 t z_1) \wedge (P_i^2 z_1 s \leftrightarrow P_i^2 z_1 t)) \tag{3}$$

for $1 \leq i \leq K_2$. Likewise, for 3-place predicates:

$$\forall z_1 \forall z_2 (P_i^3 s z_1 z_2 \leftrightarrow P_i^3 t z_1 z_2) \text{ and permutations}$$

for $1 \leq i \leq K_3$. And so on, up to predicates in N variables. Call such formulae *identity conditions* for s and t. The conjunction of all these identity conditions is to serve as our definition of identity, where the question of which conditions in fact hold, for given terms, is to be settled by appeal to the physical theory from which the non-logical vocabulary is derived.

This principle is the *only* analysis of identity that is really workable from a modern logical point of view, embracing as it does every deductive consequence of the axioms of identity. It was first proposed as such by Hilbert and Bernays in the *Grundlagen der Mathematik* in 1934. It was subsequently defended by Quine in the above, definitional sense, for any first-order language with a finite lexicon (in *Set Theory and Its Logic*, in *Word and Object*, and in *Philosophy of Logic*). Quine's interest in the principle was that it allowed him to extend his view of logical truth, as truth by virtue of grammatical form alone, to truths involving identity. As such he considered it an account of identity sufficient for mathematics as well.

There are plenty of reasons to be skeptical of Quine's program, understood to have the generality that he intended for it, but there is no need to consider them in any detail; the proposal I am making is quite different. I do not suppose that there is anything wrong with identity, taken in an irreducible sense; whatever objects there are, we know what the identity relation is among them; given objects, identity can look after itself. Neither are we concerned here with the ground for logical truth. The proposal, rather, is that in a situation in which we *do not know* what physical objects there are, but only, in the first instance, predicates and terms, and connections between them, then we should tailor our ontology to fit; we should admit no more as entities than are required by the distinctions that can be made out by their means. The most common objection to Quine's proposal is that a language-relative notion of identity cannot possibly do - that we end up with is not *really* identity.[1] But in the present context that is either to call in question the correctness of the underlying physical theory (we may not have the right vocabulary or identity conditions), or the method of interpreting it in terms of objects (certain identities may be negated in an unanalyzable sense). Certainly theory or method may be wrong; both of course are defeasible; but just for that reason, neither can be rejected *a priori*.

There is certainly plenty of evidence in favour of the principle when it comes to ordinary physical objects.[2] Following Quine (1960 p.230), call two objects *absolutely discernible* if there is a formula with one free variable true of one of them but not of the other. With the obvious extension of this terminology to sets of objects, it is clear that ordinary solid objects are all absolutely discernible: no two solid objects can occupy the same spatiotemporal position, and given an asymmetric distribution of such objects each will satisfy different spatiotemporal

[1] See e.g. Wiggins (2001, Ch.6). Having considered Quine's original proposal, Wiggins goes on to consider a supervenience thesis of identity (*ibid* p.187-8) that is closer to what I am proposing, but at this point he puts the lessons that should have been earlier learned to one side (the counter-examples he adduces against the supervenience thesis are examples of relative or weak discernibles - see below).

[2] Quine has offered an extensive account of how concepts of objects are first acquired, according to which predications (and connections among predicates) well precede the full-blown notion of object involved in the use of the identity sign (Quine 1974). Whether or not he is right on this, it can hardly be denied that the simplest ways in which we discriminate among ordinary objects is by qualitative, sensible differences.

relations with every other (referred to by bound variables). Given any countable set of absolute discernibles, for each there will exist a finite formula in one free variable that applies to it uniquely; call it an *individuating predicate* for that object. It follows that every solid object in an asymmetric universe has an individuating predicate.

If the PII identified any two objects not absolutely discernible, we would have what is usually called the *weak* version of Leibniz's principle.[3] But this is only one category of discernibility, according to the PII. Call two objects *relatively discernible* if they are not absolutely discernible and there is a formula with two free variables that applies to them in only one order. For an example, consider the instants in time in an empty Newtonian spacetime; they are all relatively discernible (of any two, one will be earlier than the other, but not *vice versa*).

There is a third and final category. An identity condition may fail even when objects have exactly the same properties and exactly the same relations to all other objects *and* exactly the same relations to each other; (3) will be false if x and y only satisfy an *irreflexive* relation A (for then $\exists z_1 \sim (Axz_1 \longleftrightarrow Ayz_1)$, namely when $z_1 = x$ or when $z_1 = y$). Call objects not absolutely or relatively discernible, that satisfy an irreflexive relation, *weakly discernible;*[4] if none of absolutely, relatively, or weakly discernible, *indiscernible.* Using these definitions the Hilbert-Bernays principle is precisely the principle of the identity of indiscernibles. (This is what I shall mean by the PII from this point on.)

For an example of weakly discernible objects, consider Black's two iron spheres, one mile apart, in an otherwise empty space (this volume, p.xxx). The irreflexive relation A is '...one mile apart from....'. It is *because* this relationship holds that we may say that there are two - that it is intuitively evident that there are two. The example was intended as a counter-example to Leibniz's original principle, in either the strong or weak form, and so it is; what went unnoticed is that it is the PII that sanctions the example, and shows us its logical form - and with which, of course, it is not in contradiction.

There are plenty of realistic examples of weak discernibles. Consider the spherically-symmetric singlet state of two indistinguishable fermions. Each has exactly the same mass, charge, and other intrinsic properties, and exactly the same reduced density matrix. Since the spatial part of the state has perfect spherical symmetry, each has exactly the same spatiotemporal properties and relations as well, both in themselves and with respect to everything else. But an irreflexive relation holds between them, so they cannot be identified (namely '....has opposite direction of each component of spin to ....'). Since symmetric, they are weakly, not relatively discernible. Indeed, indistinguishable fermions

[3] As a reading of Leibniz, this is to include relations involving bound variables as among the intrinsic denominations of a substance. (For a defence of the view that Leibniz was prepared to countenance relations not reducible to monadic predicates, see Ishiguro 1972.)

[4] This category went unnoticed in Quine (1960), where the terms "absolute" and "relative" indiscernibles were introduced. Quine remarked on it later (Quine 1976), but there he introduced a different terminology (in terms of 'discriminability'). It is true that weakly discernible objects are indiscernible under the strong or weak versions of Leibniz's principle - the traditional ones in the philosophy literature - but then so, usually, are relative discernibles; I see no reason to follow Quine in this shift in terminology.

are *always* at least weakly discernible; an irreflexive relation exists between any pair of fermions, whatever their state.[5] There has been plenty of discussion of the bearing of Pauli's exclusion principle on principles of identity; the prevelant view is that none can be secured by it under the standard, minimial interpretation of quantum mechancis.[6] But the relations that I have made use of follow from the eigenvector-eigenvalue link, and are not in any doubt.

One might conclude that the PII is so weak that it can never be compromised, but that view too is mistaken. Indistinguishable elementary bosons may all exist in exactly the same state, and satisfy no irreflexive physical relation. It was argued by Cortes (1976) that photons are a counterexample to Leibniz's principle; free photons are certainly a counterexample, even to the PII. Does it follow that the principle should be abandoned? But the argument can be turned on its head. The stable constituents of ordinary matter are all fermions.[7] Apart from the Higgs particle - not so far observed - all elementary bosons are gauge quanta; they all mediate forces between fermions. The number of elementary bosons all in exactly the same state may better be thought of as the excitation number of a certain mode of a quantum field. It is the discrete measure of the strength of dynamical couplings, dependent on the mode, between the genuine physical objects of the theory, whether fermions or other modes of quantum fields.

Schrödinger argued very early on for such a view (Schrödinger 1926). For a recent proposal of this sort, but with a somewhat different motivation, see Redhead and Teller (1992). Our conclusion, however, is that only boson numbers should be viewed as properties of things: the PII treats fermions quite differently. Given the contrast between the two, as gauge fields and sources respectively, it is a merit of the principle that it does.

---

[5] The most general antisymmetrized 2-particle state is of the form $\Psi = \frac{1}{\sqrt{2}}(\phi \otimes \psi - \psi \otimes \phi)$, where ϕ and ψ are orthogonal. Analogues of operators for components of spin can be defined as $S = P_\phi - P_\psi$, where P_ϕ, P_ψ are projections on the states ϕ, ψ. Each of the two particles in the sate Ψ has opposite value of S, but no particle can have opposite value of S to itself. (For a general theory of the state in terms of systems of relations, see Mermin 1998.)

[6] As first argued by Margenau (1944), Leibniz's principle must be rejected as the reduced density matrix for each fermion in any antisymmetrized state is exactly the same. A more general argument was given by French and Redhead (1988) and by Butterfield (1993); the theorems proved there, like those proved by Huggett (this volume), apply only to either the strong or the weak version of Leibniz's principle, not to the PII. (For a criticism of such methods from a rather different perspective, see Massima 2001; for a commentary on interpretation-dependent treatments, see Castellani and Mittelstaedt 2000.)

[7] It is unclear to French and Rickles (p. 20, fn.19, this volume) 'why metaphysics should follow the physics in this particular way, or at all'. Perhaps it need not. But if I am concerned with metaphysics at all, it is *descriptive* metaphysics, in Strawson's sense, as an aid to the interpretation of physics, and to that end I aim to preserve a good part of established practise. Ordinary objects had better turn out to be objects, on any account, and so they do on mine; it is as an extension from this that their stable constituents had better turn out to be objects as well. With the rest there is more latitude.

2 The Generalist Picture

The immediate difficulty is not that the PII identifies indiscernibles that are not even weakly discernible; it is that it does not identify those that are. On familiar, Strawsonian lines, such highly symmetric situations call in question the adequacy of the generalist picture (of what he called 'descriptions-in-general-terms'), for given two weakly discernible objects, individual reference can be made to neither of them. Likewise in the case of relative discernibles. Yet confronted with two objects of this sort, there could be no obstacle to the use of indexicals; indexicals would do better here than any purely predicative description;[8] the generalist picture is therefore incomplete - and so the PSR is also in question.

There is an obvious flaw in this argument. The use of indexicals presupposes the existence of an observing agent, but introduce such an agent and the symmetry is broken. Each of two weakly discernible objects, once related to something as highly asymmetric as, say, a functioning human being, become absolutely discernible. And there is little point in envisaging a perfectly symmetric observer, so long as indexicals are tracking perception: attention to one rather than the other object, by whatever perceptual means, will surely break the symmetry. But for all that there is a difficulty, even concerning asymmetric observers; for it is easy to imagine a *large-scale* exact symmetry (for example, a spacetime containing a plane of exact mirror-symmetry), where the observer too has a symmetric duplicate. It need not be the perception of an observer of *two* weakly discernible objects that creates the problem, it is enough, given that she has an exact duplicate elsewhere, that she sees only one of them.

The scenario is fanciful; one can deny that it is a genuine physical possibility. We are not concerned with defending the generalist picture or the PII in the face of any conceivable physics. For example, a physical theory that is explicitly a first-order formal theory (with physical objects as values of variables) is at least conceivable, however remote from the theories that we have; in such a case one might have as a law a sentence involving identity in which the identity sign cannot be treated as a defined term (using the PII) without contradiction. But these are challenges we do not have to meet.

Nevertheless, I think we should grasp this nettle.[9] We may grant a certain limitation to the generalist picture. But it is not that by the use of indexicals one can provide something *more* than what is there available, it is that one can provide something less - one provides less than a complete account of what there

[8] It should be clear, here as elsewhere, that although the PII as stated permits the use of proper names ($0-$ary functions), it is contrary to the spirit of our program to invoke them; certainly no physical theory makes use of them explicitly. (They may do so tacitly; indexicals are obviously unavoidable in practise; the question, from the point of view of the PSR, is whether they are avoidable in principle.)

[9] Here I follow Pooley (2002 Ch.2). Yet another alternative is to embrace Hacking's strategy (Hacking 1975), as recently endorsed by Belot (2001); I do not believe this strategy can be implemented with the generality Hacking claimed for it (see French 1995), but it may be it can for the special case of large-scale spacetime symmetries, that gives rise to the present difficulty for the generalist picture.

is. For suppose - the example is due to Adams (1979) - that each of Black's identical globes is inhabited, in such a way that the symmetry between them is preserved. The inhabitant of each globe refers to his own uniquely. But *that* fact is perfectly well described in the generalist picture. Nothing is left out of it. That is the point of relative and weak discernibles in the generalist picture: one can only describe the part in terms of the whole.[10]

3 Possible Worlds

A similar challenge arises from a different quarter. What of possible worlds? Might there be possible worlds which are only weakly discernible[11]? In particular, might there be a possible world that is only weakly discernible from our own? We surely do refer uniquely to actual physical objects and to the actual world.

It should be evident that here the strategy just canvassed will hardly do. It is one thing if, in the generalist picture, in highly symmetric cases, we can only describe the entire world, but it is quite another if even in the physically realistic case we can do no better than describe a set of possible worlds - that we cannot describe the actual without describing all its possible simulacra as well.

In fact there is no such difficulty. We have been talking all along of *real* physical properties and relations. Objects can only be discernible, yet fail to be absolutely discernible, if they bear relations to each other - real, irreducible physical relations, relations that are not deductive consequences of their properties. If there are none such, the PII reduces to the identity of objects not absolutely discernible (and the latter in turn to the strong principle, Eq.(2)). What real physical relation can one possible world, a possible physical universe, bear to another? A world, in Leibniz's philosophy as in modern cosmology, is a system which is physically closed. For every real relation, from spatiotemporal and causal relations to quantum correlations, the relata always have to be included together to arrive at the closed physical system.[12]

[10] Another challenge to the generalist picture arises from Kant's argument from incongruent counterparts. Pooley (this volume, p.11) may be read as insisting on a principled sense in which ostensive definition is required, in the context of a relational account of handedness. According to him, what objects we call 'left' can only, in the final resort, be shown. Against this - assuming spatial inversion is a symmetry (the relevance of this will become clear shortly) - I would maintain that a description of the universe which depicts a handed object as congruent to the hand on the side of the heart of a typical human body *describes* that object as left-handed (in other words, that the causal processes to which Pooley refers can themselves be described in the generalist picture).

Pooley rightly remarks that such descriptions do not solve the Ozma problem, but then this problem is not a difficulty for the generalist picture *per se* (concerning, as it does, the question of whether orientation can be locally defined).

[11] The PII can hardly be applied to *all* possible worlds, since among them will be worlds governed by different physical laws; here I only consider possible worlds with the same physics as ours (and, naturally, the problem only arises for these).

[12] Teller's "liberalized relationism", therefore, is not an option for us (Teller 1991). Possible objects may, of course, be physically related, if they are described as such by a physical theory

There is of course a difficulty in saying just what are the real physical relations, as opposed to the purely nominal, mathematical ones - we shall come on to that in a moment. But we do not have to settle this question to justify the claim that possible worlds may bear no real relations to one another. I take this point to follow from the definition of physical closure: if we cannot make sense of the difference between real physical relations and mathematical ones, we will be equally hard put to say what physical closure really means. If the notion of "real relation" is too vague then so is the notion of "world". And I suggest there is no good example in which the idea of closure under physical relationships is really in doubt. Even admitting exotica like cosmic wormholes, or spacetimes with topological change, it is clear whether or not one is dealing with a closed physical system if only because one is considering a *single* solution to Einstein's field equations (or a single extension of a solution).[13] Of course this is not to rule out comparisons of solutions to Einstein's equations. It clearly makes sense to talk of the mean matter density of one space-time model, in relation to another. But these relations are reducible to properties; they are deductive consequences of the properties of their relata.

Given that possible worlds bear no physical relations to one another, it follows from the PII that numerically distinct worlds will be absolutely (and in fact strongly) discernible. A world is surely an individual substance, in Leibniz's original sense of the term, even if nothing else is. But it follows from this that any object which is absolutely discernible from every other in one possible world, will be absolutely discernible from any other in any other possible world - for we have only to take the conjunct of its individuating predicate in the one possible world with the individuating predicate of the possible world to which it belongs; that will absolutely discern it from any possible object in any other world.[14]

When it comes to possible worlds, we not only obtain a form of the PII that is recognizably Leibniz's, we recover a part of his original motivation for it too. For the alternative to an analysis of identity in terms of predicates - taking identity as unanalyzable - always amounted to a purely extensional account of it, in terms of whatever objects there are (as that relation that every object bears to itself and to no other). That is how Lewis, agnostic as to the nature of the full space of logical possibility, could declare himself agnostic on the principle: whether there exist indiscernible possible worlds depended, for him, on whatever possible worlds really exist. But for Leibniz, as for those of us who

(as in the Everett interpretation of quantum mechanics): in that case they do not belong to distinct possible worlds, in Leibniz's sense (and they may well be only weakly discernible).

[13] If there is a difficulty it seems likely to lie in quantum cosmology or, classically, in regions interior to Cauchy horizons (in the neighbourhoods of singularities). The two topics are connected. In these cases we do not, properly speaking, have a serviceable theory at all. They offer no real threat to the view that a real physical relation is a part of reality, to be solved for along with all the other physically meaningful quantities, rather than a link to something beyond.

[14] It is true that the individuating predicate we end up with in this way will no longer be finite in length (for there the number of possible worlds is surely uncountable), but we are familiar with this from Leibniz's philosophy: there, in comparisons across worlds, the individual concept of a substance must be infinitely complex (this played an important role in his theory of contingency).

do not believe that possible worlds exist independent of us, that account will hardly do.

4 Leibniz Equivalence

I come back to the distinction between real and mathematical properties and relations in its more general setting. As we have just seen, transworld relations are either reducible or purely mathematical; our concern is with properties and relations internal to worlds. Here, as remarked earlier, our chief guide is experiment: properties and relations defined in terms of directly measurable quantities are certainly real. But to restrict ourselves to these would be a crude form of verificationism.

Symmetries are the key to moving beyond them. As I shall understand it, any exact symmetry of a system of equations is a transformation that leaves its *form* unchanged (under which the equations are *covariant*). It is their mathematical *form*, I take it, that has real physical meaning. Such symmetries therefore leave all the physically real quantities unchanged - among them the measurable ones. In the first instance we look for symmetries of these quantities; in the second to a theory or theory-formulation which respects these symmetries - which has a corresponding invariant structure or form. Under such a symmetry, those elements of the formalism that *are* transformed will have no direct physical significance, and their associated properties and relations that are likewise modified will not be physically real. But if there is no such theory formulation, that is an indication that there are *further* properties and relations, which are modified by these transformations, that are physically real - whether or not they are measurable. (This, as it were, *explains* why the transformations in question are not after all symmetries.) We then move to a smaller group of transformations, with respect to which the structure of the theory *is* preserved: the invariant quantities under these are the ones we count as real.

An example will illustrate the procedure. Relative distances between particles in Newton's theory of gravity (NTG) are surely measurable; they are invariant under translations, rotations, and boosts to reference frames with constant velocities. Since the equations of motion are form-invariant under these transformations - they are indeed symmetries - it follows that (absolute) positions and (absolute) velocities, as properties, are not physically real. What now of (absolute) linear accelerations? They are not directly measurable; but here the equations of NTG prove to be uncooperative. It turns out that they do not preserve their form under boosts to linearly accelerating frames. So the latter are not symmetries, and the quantities modified by these transformations - absolute accelerations, that are invariant under the symmetries of NTG - should also be counted as physically real.

This procedure, relying as it does on the details of a theory formulation, has its risks: even if the theory is substantially correct, a different formulation of it may come to light leading to a different conclusion. So it was with NTG: it turns out that there is a reformulation of the theory that does count boosts to accelerating frames as symmetries (where the accelerations are arbitrary functions of the time, but are independent of position).[15] Absolute accelerations are no longer invariant, so they go the way of absolute velocities. Only relative accelerations and relative velocities, we learn from this theory, are invariants, along with relative distances, under the full symmetry group that this theory allows. Only they should be added to the list of real physical quantities.

But the story does not stop there. It turns out that there is an empiically adequate alternative to NTG according to which certain relative accelerations and relative velocities - those associated with non-zero total angular momentum - are necessarily zero. According to this theory, due to Barbour and Bertotti (1982), the affine structure of spacetime has a purely dynamical origin, deriving from a *geodesic* principle on the *relative* configuration space. It follows that masses and relative distances alone are fundamental; every other real physical quantity (including every time-dependent quantity and all termporal relations) can be defined in terms of these. The symmetry group of this theory is correspondingly greatly enlarged.

This example makes it clear that whilst we must start with measurable quantities, they may appear in a very different light at the end. Getting the right expression for measurable quantities in terms of theoretical ones has proved almost as difficult as moving beyond them. It is now a familiar story how this played out in the case of the diffeomorphism symmetries of the general theory of relativity (GTR). This symmetry group appeared to be a step too far, given the then standard interpretation of coordinates in a physical theory. It proved a considerable difficulty for Einstein to appreciate that observable quantities could in fact be coded into the theory in a diffeomorphic-invariant way.[16]

Physically real quantities are invariant under exact symmetries - this is the general lesson. It has been long in coming. Here there is a potential for confusion which it would be well to dispel. It is sometimes said that every symmetry transformation has a 'passive' and an 'active' interpretation. It is not the right distinction: what matters is the difference between transformations that are defined by their action on physical quantities that are not themselves modeled in the equations, as opposed to those that are not, whose physical meaning, if any, has to be expressed by those equations themselves. Call them *extrinsically* as opposed to *intrinsically* defined transformations. In the former case, obviously, one is concerned only with a subsystem of the universe: the transformation in question will alter real physical relations between that subsystem and the rest (the change in the distance to the shore, when Galileo's ship is set smoothly in motion, is perfectly real). It is by means of these real physical quantities that one goes on to interpret the equations. Extrinsically-defined symmetries, then,

[15] This was clear from Cartan's reformulation of NTG as a diffeomorphic-covariant system of equations, but nothing so elaborate is needed: the clue to it was already evident in the *Principia*. There Newton showed that his equations yielded the same results for the relative motions when referred to linearly accelerating frames (and needed to, to apply his principles to the Jupiter system); see Corollary VI, Book 1.

[16] Consensus on this now appears to have been reached: see Renn *et al*, (2000). (For further background, see Norton, this volume.)

can usually be given an operational meaning. It is quite otherwise when the system of equations - *the very same* equations - is used to provide a model of the entire universe, or of a part of it (the Solar System, say, in NTG) *without* reference to the rest. The correspondence between active and passive interpretations apply only to the former sort, to transformations that are extrinsically defined, not to intrinsic ones.[17]

Extrinsic transformations, that at least in some applications can be given a clear operational meaning, have also served as a guide to establishing the symmetries of a theory - the transformations under which its form should be preserved - and hence the theoretical, intrinsic symmetries too, but one misses some of the most important ones; one is not thereby led to local symmetry groups (symmetries which, viewed as Lie groups, are infinite dimensional). One is not even led to all the finite-dimensional ones (the symmetry of the Barbour-Bertotti theory under transformations to rotating frames of reference cannot be realized operationally, applying as it does only to the universe as a whole). What similarities there are disguise the fundamental difference: extrinsically-defined symmetries, viewed from an active, operational point of view, transform real physical properties and relations, whereas symmetries that are intrinsically-defined never do. Our concern is with the latter. Only equations and transformations that can be given an intrinsic physical meaning can be used to model the world as a whole.

The consequences of the PII for such transformations are then immediate. These are intrinsically defined symmetries; they therefore leave all the real physical quantities unchanged. The world thus arrived at does not differ, with respect of any real physical property or relation, from the world with which one begins. So they are numerically the same.

In the case of diffeomorphic spacetime models in GTR, this thesis has been called *Leibniz Equivalence*. But the thesis is quite general. It applies equally to any symmetry of a physical theory, when applied to the world as a whole, and to any transformation that can only be intrinsically defined. Some of these applications remain controversial. In the case of gauge theory, and specifically electromagnetism, the invariant quantities are the electromagnetic fields: the scalar and vector potentials A_μ, that are transformed by the gauge symmetry, do not directly correspond to any real physical magnitudes. Local (differential) relations among them are real - as exhibited by the gauge-invariant 2-form $\partial_\nu A_\mu - \partial_\mu A_\nu$ - but not the potentials themselves. This case is controversial because because an explanation of certain effects - notably the Aharonov-Bohm effect - in terms of gauge invariant quantities must be non-local, in contrast to an account of it in terms of the potentials. Another controversial case is the canonical approach to GTR (the constrained Hamiltonian formalism). There

[17] The contrast has been put in a rather different way by Stachel (1993), who invokes the distinction between a theory interpreted in terms of a non-dynamical individuating field, and those interpreted in terms of a dynamical one. According to Stachel, GTR is unique in requiring the latter. It is true that there are special reasons why the symmetries of GTR must be intrisically defined, but the option is important elsewhere as well; certainly the use of a dynamical individuating field was historically important to NTG (Saunders 2002. There I spoke of 'internally' defined symmetries, but since 'internal symmetry' is already in use and means something quite different, 'intrinsic' is better).

the invariant quantities, the orbits of the group of transformations generated by the Hamiltonian constraint, are equivalence classes of three-geometries. In fact it follows that no quantities preserved by the constraints can be functions of the time (there follows the 'frozen time formalism', as advocated by Rovelli 1991).

Of the discrete symmetries, consider first permutation symmetry. This is as much a symmetry of classical statistical mechanics as quantum theory, although in the former context it has received very little attention. By Leibniz Equivalence, permutations act as the identity - classical atoms are just as 'indistinguishable' as quantum ones, in the usual physicists' sense of the term.[18] In both cases, it is to this that the extensivity of the entropy function can be traced, as required of the classical thermodynamic entropy.[19] The failure of extensivity was a puzzle in the early days of classical statistical mechanics, when it was thought that permutations of particles should yield a physically distinct state of affairs, Gibbs' protests to the contrary notwithstanding. The puzzle was quickly overwhelmed by another, the discovery of quantum statistics. It was only much later that it was realized that permutation symmetry could be treated in exactly the same way in classical statistical mechanics as in quantum mechanics.[20] The point has yet to achieve broad acceptance by the physics community. let alone among philosophers, but it is a direct consequence of Leibniz Equivalence.[21] (An obvious question then arises: whence then the difference between classical and quantum statistics? But the answer here is clear enough, at least from the point of view of phase-space methods: all the differences can be traced to the use of a discrete measure on phase space rather than a continuous one.)

Finally, consider spatial inversion. Were this a symmetry, then applied to the world as a whole it would follow that only quantities and relations invariant

[18] I hestitate to call either classical particles or fermions 'non-individuals', in the light of French and Rickles' use of the term (for such particles satisfy at least one of their criteria for individuals, namely option (d), that a version of Leibniz's principle applies to them - this volume, p.14, and Section 5.3).

[19] In particular, to yield zero entropy of mixing for samples of the same gas. Against this, van Kampen (1984) has claimed that the extensivity of the classical entropy function is only a convention. Taking this line, Huggett (1999) has concluded that nothing physical hangs on Leibniz Equivalence in the case of the permutation group (that, as he put it, there is no physical basis to favour the abandonment of 'haecceistic' phase space, a view subsequently endorsed by Albert 2000 p.45-7). It may be that the entropy of mixing for samples of the same gas is not directly measurable (although it may be viewed as the limiting case of entropies that are), but it hardly follows from that that the issue must be settled by *convention.*

[20] For a history of this early controversy, see Jammer (1966). For an account of the role of permutation symmetry in classical statistical mechanics, see Hestines (1970).

[21] French and Rickles (p.13, this volume) are clearly sympathetic with the view that Leibniz Equivalence, as applied to permutations, is incompatible with classical physics (equivalently, that classically one is committed to the use of haecceistic phase space); that Ehrenfest was right to criticize Planck's removal of the factor $N!$ as *ad hoc.* But I say that the use of the PII here, as applied to possible worlds, has an obvious pedegree in classical physics and classical metaphysics, and that there is nothing *ad hoc* about Leibniz Equivalence as I have derived it. Why insist that what was wanted was new physics, rather than a better interpretation of the old? (The same applies in the quantum case. The central question that they raise in Sec.5 - What is the ground of Permutation Invariance? - is answered the same: it is the consequence of Leibniz's principles.)

under the transformation would be real: two spacetime models, the one the spatial inversion of the other, would describe the same world, and the same handed objects within it. The hand considered in itself, in an otherwise empty space, would be neither left nor right handed.[22] This interpretation of global spatial inversion was first advocated by Weyl (1952). It was defended, with qualifications, by Earman (1989), and it has more recently been argued for by Hoefer (2000). Of course it has turned out that spatial inversion is *not* a symmetry of the standard model, but an argument to a similar effect remains: the combination of space, time, and matter-antimatter inversion (TCP symmetry) is demonstrably a symmetry of any relativistic quantum theory.[23] From Leibniz Equivalence, it follows that the world does not have one TCP orientation rather than the other. Its mirror image, on inverting matter and antimatter and the arrow of time, is one and the same.[24]

5 The Principle of Sufficient Reason

I have had more to say about the PII than the PSR; let me close with a remark on their relation: despite the changes in the former principle, they still function in tandem.

Recall Clarke's criticism of the PSR:

> Why this particular system of matter, should be created in one particular place, and that in another particular place; when, (all place being absolutely indifferent to all matter,) it would have been exactly the same thing vice versa, supposing the two systems (or the particles) of matter to be alike; there could be no other reason, but the mere will of God. (Alexander 1984, p.20-21).

Leibniz did not respond to this directly; he surely agreed with Clarke that atomism is inconsistent with the PSR - but only given the further presupposition, common to them both, that it must be possible to refer to a substance uniquely, independent of its relationships with other substances. It would be odd to make

[22] Kant clearly saw this implication of Leibniz's principles, although initially he thought it confined to Leibniz's views on the nature of space (Pooley, this volume, Sec.2). Two years later, in the *Inaugural Dissertation* of 1770, he had rejected the generalist picture as well, and with that much broader principles of Leibniz's philosophy.

[23] Pooley (this volume, Sec.4) denies that PCT symmetry is of any relevance to a relational account of handedness. But my claim is that it tells against absolutism: as I understand it, *only* transformations that are symmetries provide a sufficient condition for a quantity to be counted as unphysical. It is only insofar as absolute spatial orientation is changed by the PCT transformation, if a symmetry, that one can conclude it is not physically real. In the absence of such a symmetry I see no reason to suppose that absolute 3-dimensional spatial orientation is not a perfectly respectable real physical quantity in its own right, Pooley's arguments to the contrary notwithstanding.

[24] Of course there remain a number of open questions about the arrow of time, which cannot be addressed independent of the interpretation of quantum mechanics. (For example, TCP inversion can hardly remain a symmetry of state-reduction theories.)

such a claim today: it is a piece of metaphysics without any basis in modern logic; there is no reason to believe in it from the point of view of any physical theory; it is more contentious than any of the principles so far considered.

Abjuring metaphysics of this order, there is no longer a conflict with the PSR. Leibniz Equivalence applies here as to any other symmetry - for given Clarke's assumption that the particles are exactly alike, particle permutations are surely symmetries. The permutation does not therefore lead to a possible world numerically distinct from the actual one. No decision as to which particle is to be placed in which position needs to be made.

A second example shows better how the novel features of the PII - the existence of relative and weak discernibles - works in tandem with the PSR. Consider the location of a material system in space. Were space a real entity, then, according to Leibniz, it would again follow that its parts must be individuated uniquely, without reference to anything else. But since qualitatively alike, there could be no reason to situate the material system in one place rather than another - a problem for the PSR. How does this case fare under the PII?

The points of space, independent of their relations to matter, unlike particles of matter, independent of their spatial relations, are in fact discernible. If, now, it were possible to refer to one point of space rather than another (without reference to matter), it would make sense to ask at which of the two the material system is to be placed, leading to the same difficulty with the PSR. But in fact the points of space are only *weakly* discernible, so we cannot refer to any one point rather than another, and the difficulty does not arise. Evidently insofar as we can view the parts of a highly symmetric entity, such as a homogeneous space, as objects in their own right (as discernibles), without reference to anything else, it is essential - consistent with the PSR - that they not be absolutely discernible from one another.

It is also worth remarking that it is only in the generalist picture that there is no conflict with the PSR. For let us introduce into this space Adams' pair of identical globes, each with identical observers. The material system is to be placed adjacent to one globe rather than the other - which? In the generalist picture, there is no choice to be made, for neither can be referred to uniquely. From the standpoint of the two observers, one will see the system appear nearby, but not the other - again, this fact can be reported without any difficulty. But the one who sees it appear nearby will consider the event to be wholly arbitrary, and contrary to the PSR. Here, one might say, is chance, where in the generalist picture everything is determined.

6 Close

I have spoken throughout of the interpretation of theories in terms of objects, but it is object in the logical sense, the sense that Frege was concerned with. And, whilst I maintain that certain doubts have been laid to rest concerning it - for example, as expressed by Quine, on whether the concept of object must crumble altogether in the face of particle indistinguishability in quantum mechanics (Quine 1990 p.35-36) - it is object in a very thin sense that is secured. In strongly interacting high-energy physics, it is doubtful that objects as individuated (using the PII) by the invariant properties and relations definable in quantum field theory will be quanta at all (although, in the light of asymptotic freedom in QCD, in the ultra-relativistic limit quarks as objects presumably remain). In classical field theories, obviously, one obtains little more than field values (or, given diffeomorphism symmetry, relations between field values). These are objects as events.

Coincidences of field values, and complexes of relations among them - this is a world understood in terms of structural descriptions, a world as graph, not a collection of things that evolve in time. It is a structuralist account, too, by virtue of its reliance on mathematical form, as the key to the distinction between the physical and the merely mathematical. But I see no reason, deriving from special relativity and quantum theory, to deny that the descriptions one ends up with, in the interpretation of physical theories, gives the properties and relations among objects, in the logical sense of the term. The conservative notion of object that I have been concerned with, thin as it is, is not in jeopardy. It is not so thin as to be governed by an unanalyzable notion of identity.

What is not so certain is that the current framework will be preserved in quantum gravity. In the quantum canonical approach, unlike in the classical theory, it is not so clear that quantities preserved by the constraints are sufficient to build up an account of change (not on any approach to the problem of measurement). And here there are avenues being explored - causal sets, for example - that, if successful, may well lead to a more direct account of identity than the logical one that I have given. They are speculative, and if contrary to the PII, they will have violated the methodology that I have been advocating; but the PII as I understand it is no *a priori* truth.

References

Adams, R. (1979), 'Primitive Thisness and Primitive Identity', *Journal of Philosophy*, **76**, 5-26

Albert, D. (2000), *Time and Chance*, Harvard University Press, Cambridge.

Alexander, H, (1956), *The Leibniz-Clarke Correspondence*, Manchester University Press, Manchester.

Barbour, J. and B. Bertotti (1982), 'Mach's Principle and the Structure of Dynamical Theories', *Proceedings of the Royal Society (London)* **382**, 295-306.

Belot, G. (2001), 'The Principle of Sufficient Reason', *The Journal of Philosophy*, xxxxx

Black, M. (1952), 'The Identity of Indiscernibles', *Mind*, **61**, 153-64.

Butterfield, J. (1993) ‘Interpretation and Identity in Quantum Theory’, *Studies in History and Philosophy of Science*, **24**, 443-76.

Castellani, E., and P. Mittelstaedt (2000), “Leibniz’s Principle, Physics, and the Language of Physics”, *Foundations of Physics* **30**, xxxx.

Cortes, A. (1976), ‘Leibniz’s Principle of the Identity of the Indiscernibles: A False Principle’, *Philosophy of Science* **45**, 466-70.

Earman, J. (1989), *World Enough and Space-Time,* MIT Press, Cambridge.

French, S. (1995), ‘Hacking Away at the Identity of Indiscernibles: Possible Worlds and Einstein’s Principle of Equivalence’, *Journal of Philosophy*, **91**, 455-66.

French, S., and M. Redhead, (1988), ‘Quantum Physics and the Identity of Indiscernibles’, *British Journal for the Philosophy of Science,* **39**, 233-46.

Gödel, K. (1930), ‘Die Vollständigkeit der Axiome des logischen Funktionenkalküls, Monaatshefte für Mathematik und Physik’, **37**, 349-360. Translated as ‘The Completeness of the Axioms of the Functional Calculus of Logic’, in From Frege to Gödel, J. van Heijenoort, ed., 1967, Harvard University Press, Cambridge, Mass.

Hacking, P. (1975), ‘The Identity of Indiscernibles’, *Journal of Philosophy*, **72**, 249-56.

Hestines, D. (1970), ‘Entropy and Indistinguishability’, *American Journal of Physics*, **38**, p.840-45.

Hilbert, D., and P. Bernays (1934), *Grundlagen der Mathematik*, Vol.1, Springer, Berlin.

Hoefer, C. (2000), ‘Kant”s Hands and Earman’s Pions: Chirality Arguments for the Substantival Space’, *International Studies in the Philosophy of Science*, **14**, 237-56.

Huggett, N. (1999), ‘Atomic Metaphysics’, *Journal of Philosophy*, XCVI, 5-24.

Ishiguro, H. (1972), *Leibniz’s Philosophy of Logic and Language*, Cornell, Ithica.

Jammer, M. (1966), *The Conceptual Development of Quantum Mechanics*, Mc.Graw Hill, New York.

Leibniz, G. W. (1714), ‘Monadology’, in *Monadology and Other Philosophical Essays*, trans. P. and A. Schrecker, Bobb-Merrill, Indianapolis, 1965.

Margenau, H. (1944) ‘The Exclusion Principle and its Philosophical Importance’, *Philosophy of Science* **11**, 187-208.

Mermin, D. (1998), ‘What is Quantum Mechanics Trying to Tell Us?’, *American Journal of Physics* **66**, 753-767.

Massimi, M (2001), 'Exclusion Principle and the Identity of Indiscernibles: a Response to Margenau's Argument', *British Journal for the Philosophy of Science*, forthcoming.

O'Leary-Hawthorne, J., and J. A. Cover (1996), 'Haecceitism and Anti-Haecceitism in Leibniz's Philosophy', *Noûs*, **30**, 1-30.

Pooley, O. (2002), *Understanding Spacetime: Recent Relationalist and Substantivalist Themes in the Interpretation of Physics*, D.Phil Thesis, University of Oxford.

Quine, W. V. (1960), *Word and Object*, Harvard University Press, Cambridge.

Quine, W. V. (1974), *The Roots of Reference*, La Salle, Open Court.

Quine, W. V. (1976), 'Grades of Discriminability', *Journal of Philosophy* **73**; reprinted in *Theories and Things*, Harvard University Press, Cambridge (1981).

Quine, W. V. (1990), *Pursuit of Truth*, Harvard University Press, Cambridge.

Redhead, M., and P. Teller (1992), 'Particle Labels and the Theory of Indistinguishable Particles in Quantum Mechanics', *British Journal for the Philosophy of Science*, **43**, 201-18.

Renn, J., T. Sauer, M. Janssen, J. Norton, and J. Stachel (2000), *The Genesis of General Relativity: Sources and Interpretation*. Vol.1, *General Relativity in the Making: Einstein's Zurich Notebook*, Kluwer, Dordrecht.

Rovelli, C. (1991) 'What is Observable in Classical and Quantum Gravity', *Classical and Quantum Gravity*, **8**, 297-304.

Saunders, S (2002), 'Indiscernibles, General Covariance, and Other Symmetries: the Case for Non-Eliminativist Relationalism', in *Revisiting the Foundations of Relativistic Physics: Festschrift in Honour of John Stachel*, A. Ashtekar, D. Howard, J. Renn, S. Sarkar, and A. Shimony, (eds.), Kluwer, Dordrecht.

Schrödinger, E. (1926), 'Zur Einsteinshen Gastheorie', *Physicalische Zeitschrift*, **27**, 95-101.

Stachel, J. (1993), 'The Meaning of General Covariance', in *Philosophical Problems of the Internal and External Worlds: Essays Concerning the Philosophy of Adolf Grünbaum*, A. Janis, N. Rescher, and G. Massey, eds., University of Pittsburgh Press, Pittsburgh.

Teller, P. (1991), 'Substance, Relations, and Arguments About the Nature of Space-Time', *The Philosophical Review*, **100**, 363-97.

van Fraassen, B. (1991), *Quantum Mechanics*, Clarendon Press, Oxford.

van Kampen, N. (1984), 'The Gibbs Paradox', in *Theoretical Physics*, W. Parry, ed.., Pergamon Press, Oxford, pp.305-12.

Weyl, H. (1952), *Symmetry*, Princeton University Press, Princeton.

Wiggins, D. (2001), *Sameness and Substance Renewed*, Cambridge University Press, Cambridge.