

TIME AND BACKGROUND-FREE THEORIES

1

In the paper I wish to begin to explore the consequences for metaphysics of thinking that a good physical theory should be background-independent. More generally I want to ask whether the conception of time not as a background but as an active component of the physical universe has any significant consequences for metaphysics.

I think that a natural conception of space and time is to regard them as a (possibly infinite) container or stage for the events that make up the history of the universe. They are not part of the contents of the container nor are they actors or props in the action on the stage. They are an inert but necessary background. This conception plays a part in metaphysical argument. And reasons to doubt that conception may undermine some of those arguments. In this paper I would like to examine this conception and the possible consequences of so doing for certain metaphysical questions.

2

The first piece of metaphysics I will look at concerns the nature of fundamental natural properties. Brian Ellis (2001) takes the view that some of these are essentially dispositional whereas others are not; they are essentially categorical. The essentially dispositional properties are those whose essential natures are to be understood in terms of the dispositions they confer on the objects possessing them. For example, it might be that charge is such a property. If so, the property of possessing a charge of value q_1 confers on any object with that charge a disposition to exert a force on other charged objects and to experience a force, in accordance with Coulomb's law:

$$\mathbf{F} = \epsilon_0 \frac{q_1 q_2}{r^2} \quad (1)$$

and to generate a magnetic field when in motion according to the Biot–Savart law:

$$d\mathbf{B} = \frac{\mu_0}{4\pi} \frac{(\mathbf{J}dV) \times \hat{\mathbf{r}}}{r^2} \quad (2)$$

While there are many questions to be asked of such a view, let us assume that it is correct, for current purposes. What is interesting is that Ellis does not think that what goes, in this example, for charge, goes for all fundamental properties and relations. He does not think that they apply to spatial and temporal relations. Thus there is no disposition that reflects the essential nature of a spatial distance nor of a temporal duration nor even of a relativistic space-time interval.

Why should Ellis deny this? Stephen Mumford (2004: 188) suggests what might seem an obvious riposte. Transposing Mumford's response to our example, we may

observe that spatial separation is a factor in Coulomb's law just as charge is, likewise for the Biot–Savart law. Mumford's suggestion is in effect that we can see these equations as reflecting a dispositional consequence of the nature of spatial separation, just as we see them reflecting the dispositional nature of charge.

A parallel riposte for time may be less obvious, but is available nonetheless. In the Biot–Savart law, charge is hidden in the quantity \mathbf{J} , which is the current density, i.e. the density of charge flowing per unit time. Likewise time is of course implicated in many laws, such as Newton's second law:

$$\mathbf{F} = m\mathbf{a} \tag{3}$$

The contribution of time may be more directly perceptible in derived laws such as:

$$\mathbf{F}t = \Delta m\mathbf{v} \tag{4}$$

which tells us that the change in momentum due to a constant force is equal to that force times the duration for which it is applied, and:

$$\mathbf{s} = \frac{\mathbf{F}t^2}{2m} \tag{5}$$

which tells us how far a mass m will be moved, from stationary, by a force \mathbf{F} applied for a duration t .

So there is no lack of laws that refer to temporal intervals as crucial determinants of various outcomes. Thus Mumford's response looks entirely feasible: these laws show that certain quantities are nomically dependent on time. We can regard these laws as reflecting the dispositional nature of temporal relations.

3

While I agree that Mumford's response is basically correct, it does also seem to me that there is an intuitive asymmetry between the contribution of charge and of distance in Coulomb's law and between charge and time in the Biot–Savart law or between time and force or mass in the Newton's third law and its derived laws. Considering Coulomb's law, it is natural to think of the force as being generated by the charge, and the role of the distance as secondary, as moderating the effect of the charge. It is thus plausible to think of the charge as the disposition, but not the distance. Likewise, in (4) it is tempting to think of the force as the causal power here, and the duration as playing a lesser role. If one thinks that a turned on tap is disposed to fill the sink, that quantity of water in the sink is determined by how much the tap is on and how long it is on for. The disposition here concerns the tap and rate of flow of water; the duration of the flow is the junior partner in this arrangement. Likewise in (5) the distance moved by the mass is a manifestation of the force and of the mass, which resists the force. The time t in a sense measures how much the force disposition is applied but is not itself a cause, and so cannot be attributed a disposition.

In the last paragraph I have tried to express the reluctance we might feel towards endorsing Mumford's proposal. I am very far from suggesting that we should allow this reluctance to move us to reject the proposal—on the contrary, I support the proposal. However, I do suggest that this reluctance explains Ellis's view that certain

properties and relations are not to be regarded as essentially dispositional, and that these include spatial and temporal relations.

If that is correct, then in order to assess Mumford's proposal we need to identify the source of the intuitive reluctance we feel towards it and then we can come to a view as to whether this intuition may or may not be justified. I suggest that the following reasons might occur to one, *prima facie* at least, as explanations for this intuition:

- (i) Very many laws refer to temporal intervals; they cannot all be regarded as manifestations of a temporal disposition.
- (ii) Durations are relations and so cannot be dispositions.
- (iii) Where some other nomically relevant property is clearly dispositional, we readily identify that property as the active disposition and so can exclude the temporal relation.
- (iv) Time is a metaphysical background and so is non-causal; as such it cannot be dispositional.

I shall now consider these in turn.

(i) It is true that temporal relations occur in many laws of nature, and so if they are dispositions, they are multi-track dispositions. I myself am reluctant to think of fundamental properties as multi-track dispositions. To do so is to hide a mystery. For example, in classical physics at least, it is I think misleading to think of there being a single multi-track property mass, which manifests itself both in Newton's second law and also in the law of gravitation. Rather, it seems to me as to others, more perspicuous to think of there being two single-track properties, inertial mass and gravitational mass, which as a matter of nomic fact are perfectly proportional. That nomic proportionality is of course a fact that needs explanation, presumably in dispositional terms. In my view, it is correct that multi-trackness might be a reason for denying fundamental dispositionality, unless it can be explained away. I doubt, however, that this thought is responsible for the intuitive reluctance we have identified. The view just sketched, that multi-track dispositions are not fundamental, is a highly contentious one. Its truth, if it is true, is unlikely to have much of an influence on our intuitions and so does not explain the target phenomenon, our intuitive reluctance to regard temporal relations as dispositional. After all, charge is involved in several laws of nature—above we have mentioned two, Coulomb's law and the Biot–Savart law, but that does not significantly dampen the intuitive plausibility of the idea that charge is dispositional.

(ii) One might think that dispositions have to be unary properties, intrinsic properties in particular, and so temporal relations cannot be dispositions. This might be a source of the target intuition. But if so it is a poor basis for that intuition. Dispositions can certainly be extrinsic, as Jennifer McKittrick (2003) shows. For example, this key is disposed to open my front door, but that disposition depends on the lock with which that door is fitted. The weight of some object is the force it experiences in its local gravitational field and as such implies various dispositions, such as the disposition to depress a spring balance to a certain degree, that are not intrinsic. Correspondingly, there are relations that confer dispositions: key x matches lock y implies that y is disposed to be opened by x . x is heavier than y implies that x is disposed to displace y in a pair of scales. Abelard loves Eloise implies that Abelard is disposed to behave in certain ways towards Eloise. It *may* be that our intuitions that relations

cannot be dispositional is stronger when it comes to fundamental properties—these should all be intrinsic, monadic properties. Nonetheless, I am not sure that such an intuition is either that strong or correct. Whether it is correct or not rather depends on what advances physics reveals to us, but it is entirely plausible that mass, which is dispositional, will turn out to be a relation between a particle and the Higgs field.

(iii) The third response suggests that we have a strong intuition that an effect can be the outcome of one disposition only, and so where two properties appear to be involved, only one of them can be regarded as the disposition whose manifestation is this outcome. If some force is the result of a charge at some distance, then since the charge is clearly a causally efficacious disposition, the distance cannot be. Likewise, if an object changes its momentum as a result of some force being applied for a time, since the force is clearly dispositional, the time cannot be. Again, I doubt that our intuitions are especially strong in this regard; furthermore if this were the source of the target intuition, it would be a poor ground for it. Typically a disposition will require a stimulus in order to be manifested. This stimulus will itself typically be the instantiation of some disposition. So one would expect many cases of the manifestation of one disposition to involve a second disposition. For example, the breaking of a fragile glass will involve the stimulus of a stress of a certain kind being applied to it. But that stress will itself be a disposition, and so this case ought to be viewed as the interaction of two dispositions. Consider an air-filled rubber balloon that is in a deep water bath. By controlling the depth of the balloon (and hence the pressure on the balloon) and the temperature of the water (and hence the temperature of the air in the balloon), we can use the ideal gas law to calculate the volume of the balloon:

$$V = \frac{nRT}{p} \tag{6}$$

It seems correct to think of the volume of the balloon as the result of the interaction of two dispositions, the pressure on the balloon, p and its temperature T . There seems no inclination to regard one as not being dispositional just because the other is.

(iv) The last suggestion is that our intuition that temporal intervals are not dispositions arises from the fact that we intuitively think of time as a background and so not as a causal property and so not a disposition. I shall explore this proposal in the next section.

4

In suggesting that we have an intuition that temporal relations cannot be proper causes, I am not suggesting either that we do not think of such relations as explanatory or that we never refer to temporal relations in causal contexts. Clearly the fact that temporal relations appear in equations such as (4) and (5) means that they will figure in explanations. Nonetheless, when attributing causation we do tend to think of time as subsidiary in such contexts, as a modifier of the cause rather than as a cause itself. It is as if time is thought of adverbially. It is true that we may refer to time in causal contexts: “We missed the train because Humphrey took so long polishing the shoes”. Even so, I think that we are inclined to regard time in a similar light to absences. We talk as if absences are causes, but we do not regard them as real, bona fide causes; the real cause is something else (my failure to water the plant is not the ‘real’ cause of its withering; that was the heat which dried it up). Likewise,

we may think of temporal relations as figuring in causal explanations in a secondary, derivative or partial way.

As mentioned, one hypothesis for explaining this is that we intuitively regard space and time as backgrounds and as such playing only a shadow causal role. This intuition is reflected in many of our theories and infects our metaphysics.

What is a background? According to John Baez (2000), a background structure is “any sort of structure appearing in a mathematical model of a physical system that is *fixed* rather than *dynamical*—i.e., which does not depend on the state of the physical system in question.” Baez goes on to give a heuristic characterization, “we can think of a background structure as something which affects the dynamics of the system while remaining unaffected by it.” Baez then points out that this violates the reciprocity of cause and effect, a relationship that Harvey Brown calls the ‘action–reaction principle’. Baez gives a simple illustration, the problem of a bead on a wire. In solving this problem one will treat the wire as a background. The curve of the wire

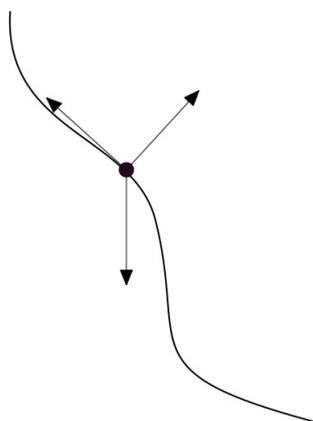


Figure 1:

is fixed and does not change and this allows, depending on the equation of the curve, a relatively straightforward means of calculating the equation of motion of the bead, via the forces of gravity and friction. However, this ignores the fact that the bead is exerting corresponding forces on the wire, which would thus deform or move. If one did take the forces on the wire into account, then the position of the wire would no longer be a background, because it would not only affect the dynamics of the bead but is also affected by the dynamics. (We would probably then treat the endpoints of the wire as a background.)

Time and space in Newtonian mechanics are treated as a background. The temporal and spatial metrical structure, i.e. the relationships among space-time points, are fixed. In this case not *everything* that is part of or a consequence of the structure affects the dynamics of physical objects without being affected by them. For Newtonian space-time determines an absolute position and velocity for objects, but these play no part in affecting the dynamics of those objects. They are nonetheless part of the background structure. In this sense redundant structure is also background—we might distinguish between active and redundant background. A redundant background structure is not strictly in conflict with the action–reaction

principle, although it would conflict with the Eleatic principle, that only what is causally (nominally) efficacious exists. Active background does play a part in determining dynamics, and so is in conflict with the action–reaction principle. Moving from classical Newtonian space-time to neo-Newtonian space-time eliminates its redundant background. On the other hand it does not eliminate its active background, i.e. those elements of the classical picture that violate the action–reaction principle.

One might wonder, from the perspective of a dispositional essentialist view of all properties, why the action–reaction principle is important. After all, if the properties and relations in question can be part of an ‘action’, why does the lack of a ‘reaction’ matter? My view is that the action–reaction principle plays a powerful role in our understanding of causation. Basic causal action is always interaction. Interestingly this is a point that Ellis emphasizes very strongly. If that is right, then the fact that time is treated as a background implies that time and temporal relations are not *real* causes. Or if they do exist they are pseudo-causes or modifiers to the real causes but not real causes themselves. Thus in a realistic physics, time as a background structure occupies an uncomfortable and unstable position. It is part of the structure, playing a role in nomological explanation; but also not truly real, not a full cause. So in summary:

time is a background structure \wedge the action–reaction principle is true \Rightarrow
time is non-causal (non-dispositional)

In which case Ellis is right about time (and likewise about space).

5

One possible response to the uncomfortable position of time as a background structure is to deny that time really exists at all. The idealist interpretation of Leibniz suggests that this is his solution. Of course, such a response does nothing to rehabilitate time as causal or dispositional.

The alternative is to seek a physics which eschews background structures. Indeed relationalism of the kind Leibniz promoted and the elimination of background structure are related, but not quite the same. I take relationalism to be the claim:

(R) the existence of space and time are dependent on the existence of material objects and relations among them.

Note that Leibniz’s arguments against Newton’s absolutism are directed against the *redundant* background structure, not against background structure per se. Thus his arguments from the principle of sufficient reason are effective (if effective at all) only against classical Newtonian space-time with its absolute times, positions, and velocities, but not against neo-Newtonian space-time, even though the latter has plenty of (active) background structure.

Still there is a slightly different, more general, sense of ‘absolute’ as used by Einstein:

If Newton called the space of physics ‘absolute’, he was thinking of yet another property of that which we call ‘ether’. Each physical object influences and in general is influenced in turns by others. The latter,

however, is not true of the ether of Newtonian mechanics., The inertia-producing property of this ether, in accordance with classical mechanics, is precisely not to be influenced, either by the configuration of matter, or by anything else. For this reason, one may call it 'absolute'. (Einstein 1999: 15)

All background structures, being immutable, are absolute in this sense.

The debate between absolutism and relationalism and that between substantivalism and relationalism are typically taken to be more or less the same debate. Part of the problem here is that even once we distinguish between philosophical uses of 'substance' and everyday uses, the philosophical uses are still varied. Principal among the features attributed to substances are the following:

(A) Substances are genuine particulars. This is the Aristotelian idea that substances are the subjects of predication but cannot be predicated of other things. This makes them particulars in modern parlance. I think it is implicit in most Aristotle-inspired discussions that substances are normal 'things', but not gerrymandered combinations of them: Julius Caesar, Brutus's knife, the Temple of Jupiter, are all substances, but their mereological sum is not.

(B) Substances are substrates. This is the notion of substance we find in Locke, and relates closely to the Aristotelian notion: substances are what support of bear properties

(C) Substances are unchanging. This conception we find in Hume (1978: 220), when denies that there really is any substance in normal cases, since it is an illusion bred by the smooth changes we perceive in things, that we infer that there is some unchanging thing that "continue[s] the same under all these variations".

(D) Substances independent fundamental entities, on which others depend (a) for their existence and/or (b) for their properties.

Let us assume that time and space (or space-time) are objects, and that these not identical with matter. Then they are (A) genuine particulars. Whether they are (B) substrates depends on the precise relationship between space, time and matter. The Boscovich view would make space-time a kind of substrate. It seems difficult to see how something could be a substrate if it were not fundamental, so if (D) is false of space-time (B) seems false too.

If relationalism is true, then the existence of space-time is dependent on matter, and so (D) is false of space-time, and so, as stated (B) is false too. Thus for Leibniz the substrate role, indeed the true substance role is reserved for monads.

Whether (C) must be false of space-time if relationalism is true is unclear. Perhaps space-time could be a dependent entity, but insofar as it exists, is an unchanging substrate. I presume that Leibniz thought that although time and space were dependent on relations between objects, he did not think that the structure of space and time change as objects move and accelerate. So space and time might be dependent but unchanging.

How does the rejection of absolutism and background structures fit into this picture? The key feature of a background structure is its immutability. And so if we eliminate background structures, space and time cannot have the feature (C). But as just suggested, relationalism does not directly imply the negation of (C), only

of (D), and presumably (B). The adoption of relationalism does not suffice to ensure background-freedom. Whether some proposal for relationalism in fact secures background-freedom depends on the details of that proposal. Thus it seems to me that Mach's programme, radically relationalist though it is, does not imply the falsity of (C) in any obvious way, and so does not imply that we have removed space-time as a background.

6

A number of philosophically-minded physicists have argued that a satisfactory theory of quantum gravity must be background-free (Smolin 1991; Rovelli 1997). If they are correct, then that provides welcome if indirect empirical support for the application of dispositional essentialism to spatio-temporal properties. If space and time are not backgrounds, then that removes the motivation for regarding them as sufficiently acausal to be not essentially dispositional.

I shall not review the arguments for background-independence arising from consideration of the prospects for quantum gravity. Rather, I shall remark on the consequences for the issues I have been addressing of the well-known fact that Einstein's general theory of relativity is a major advance in the direction of background-independence, since that advance is sufficient for our purposes. The important development is that the space-time metric in general relativity is dynamical. That is, it is determined by the specific solutions to the equations of the theory and as a consequence its structure is determined by, for example, by the matter field.

With this in mind we can see not only that the structure of space-time is changeable as a consequences of the distribution of matter, but also that this structure can indeed be a genuine cause and can be conceived of in dispositional terms. The well-known rubber-sheet illustration of general relativity emphasizes this point. Not only does a massive object cause space-time to change its structure but also that structure influence the behaviour of matter, for example, the motion of a test particle. It is thus plausible to see the structure of space-time in dispositional terms. One way to see this is to consider that in Newtonian physics, gravitational mass is a simple disposition: a mass exerts a force directly on another. However, in general relativity, this disposition is no long a simple disposition but a compound one. The gravitational effects of a mass on a test particle are mediated by the distortion of space-time. That is the gravitational disposition is made up of a disposition to distort space-time *plus* what must be a second disposition, the disposition of space-time to influence the motion of the test particle. The point can be made in more detail by considering the fact that I can distort space-time by subjecting some elastic material object to a stress. That action will also have the effect of creating gravitational waves, which may in turn bring about stresses in some distant object (Nerlich 1994: 183).

To some extent we have moved away from our original question, which was how to characterise temporal duration and spatial displacement in dispositional terms. Instead it looks as if we are attributing dispositions to regions of space-time. Are we now answering a different question from the one we started with? If the question is different, then that is what one would expect as a consequence of the change in the science. The dispositional essentialist view is concerned with natural properties and relations, and in particular with fundamental ones. And so, if our preferred current theory tells us that local space-time metric is what is explanatorily fundamental

rather than point to point spatial and temporal intervals, then we need to shift our focus from the dispositionality of the latter to the dispositionality of the former.

7

To conclude: temporal relations and spatial relations present the monistic dispositional essentialist with a problem, because it is difficult to see how these relations can be conceived of as dispositional at all. While pointing to their roles in laws might seem to offer a route to understanding their as dispositional, it remains the case that our intuitions do not endorse a conception of space and time as causal, and that fact is a block on our seeing them as dispositional. The view that space-time is explanatory but not causal in classical physics is one endorsed by other writers on this topic, such as Graham Nerlich (1994). My own diagnosis of this is that the classical view violates the action–reaction principle, and consequently it is implausible to regard space-time as causal. Thus in order to regard space-time as causal and hence potentially dispositional, we need to remove that feature of the classical view that violates the action–reaction principle. That feature is the fact that classical space and time are background structures. And, thankfully (from the point of view of dispositional essentialism) we are not obliged to maintain that feature. Indeed general relativity removes it for us, showing how the variable geometry of space-time is both disposed to undergo changes and is disposed to bring about changes. We are thus entitled not to regard space-time and its properties as purely categorical, but can treat them as we do other fundamental natural features of the world.

References

- Baez, J. 2000. What is a background-free theory? Internet publication URL = <http://math.ucr.edu/home/baez/background.html> accessed 06.06.2009.
- Einstein, A. 1999. On the ether (1924). In S. Saunders and H. R. Brown (Eds.), *The Philosophy of Vacuum*, pp. 13–20. Oxford: Oxford University Press.
- Ellis, B. 2001. *Scientific Essentialism*. Cambridge: Cambridge University Press.
- Hume, D. 1739-40/1978. *A Treatise of Human Nature*. Oxford: Clarendon Press.
- McKittrick, J. 2003. A case for extrinsic dispositions. *Australasian Journal of Philosophy* **81**: 155–74.
- Mumford, S. 2004. *Laws in Nature*. London: Routledge.
- Nerlich, G. 1994. *What Spacetime Explains. Metaphysical Essays on Space and Time*. Cambridge: Cambridge University Press.
- Rovelli, C. 1997. Halfway through the woods: Contemporary research on space and time. In J. Earman and J. Norton (Eds.), *The Cosmos of Science*, pp. 180–223. Pittsburgh, PA: University of Pittsburgh Press.
- Smolin, L. 1991. Space and time in the quantum universe. In A. Ashtekar and J. Stachel (Eds.), *Conceptual Problems of Quantum Gravity*, pp. 228–91. Boston, MA: Birkhauser.