

Causal Fundamentalism in Physics

Henrik Zinkernagel

Department of Philosophy I, Campus de Cartuja,
18071, University of Granada, Spain.

zink@ugr.es

Forthcoming in M. Dorato, M. Rédei and M. Suárez (eds) *EPSA 2007: Launch of the European Philosophy of Science Association* (Springer 2009)

1. Introduction

Causality in physics has had bad press in philosophy at least since Russell's famous 1913 remark: "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, p. 1). Recently Norton (2003 and 2006) has launched what would seem to be the definite burial of causality in physics. Norton argues that causation is merely a useful folk concept, and that it fails to hold for some simple systems even in the supposed paradigm case of a causal physical theory – namely Newtonian mechanics.

The purpose of this article is to argue against this devaluation of causality in physics. I shall try to defend that Norton's charges against causality in Newtonian mechanics are flawed, and I will also suggest how the central causal message of Newtonian mechanics may proliferate into its supposed successor theories, namely special (and to some extent general) relativity and quantum mechanics. My main argument is that Norton's (2003) alleged counterexample to causality (all events have causes) within standard Newtonian physics fails to obey what I shall call the causal core of Newtonian mechanics (essential parts of the first and second law). In particular, I argue, Norton's example is not in conformity with Newton's first law – and his attempt to reformulate this first law (in order to make it conform to his example) results in an impoverished theory which lacks important features of Newtonian mechanics (in particular, in this 'Nortonian mechanics', the notion of inertial frames lacks a physical justification, and the close connection between the first law and the notion of time is lost).

2. The dome and the alleged failure of determinism (and causality)

Norton (2003) provides an interesting example of a system – a mass on a dome – which seems to conform to Newtonian mechanics, yet fail to be deterministic. This indeterminacy comes about due to an uncaused change in the state of motion of the mass on the dome, and can therefore also be seen as a failure of causality (all events or, more precisely in Newtonian mechanics, all changes of states of motion have a cause). As Norton points out, this example is striking as no reference to exotic features, such as space invaders appearing with unbounded speed from spatial infinity or an infinite number of interacting particles, are needed to produce the indeterminism.¹ If Norton is

¹ Both space-invader examples and the so-called supertasks (involving an infinite number of particles) may be argued to be less troublesome for determinism in Newtonian mechanics e.g. since they involve

right that his example is an example of a Newtonian system then not all Newtonian systems are causal and, as Norton emphasizes, Newtonian mechanics cannot therefore (in general) license a principle or law of causality.

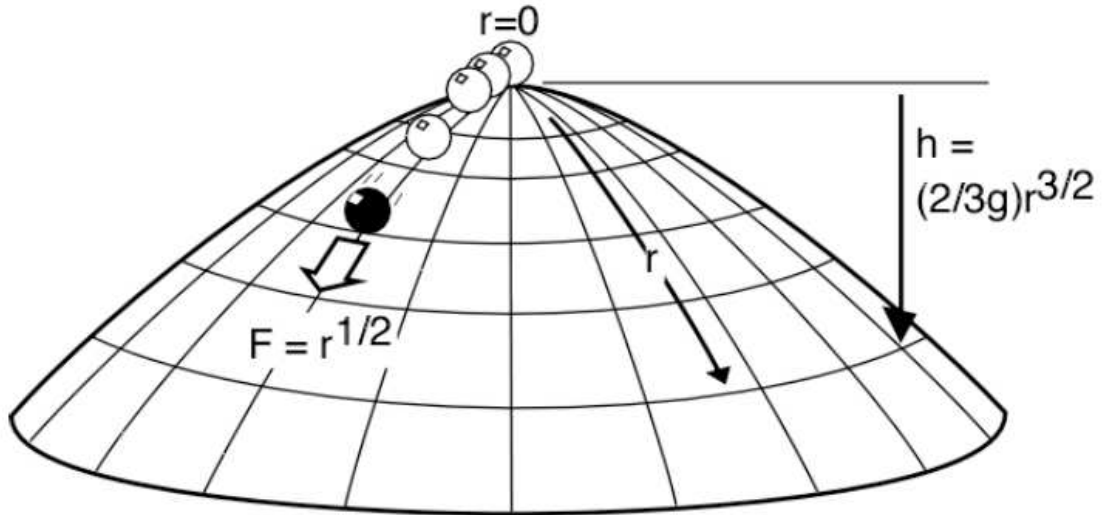


Figure 1: The mass on the dome (taken from Norton 2006).

In the example, we are invited to consider a unit mass point which, under the influence of gravity, can slide frictionlessly over the surface of a dome in which a radial coordinate r is inscribed (Figure 1). The shape of the dome is given by the height function $h=(2/3g)r^{3/2}$ which specifies the vertical distance of each point below the apex at $r=0$. At any point on the dome, the gravitational force tangent to the surface (equal to the net force acting on the mass point) is $F=r^{1/2}$ which implies that Newton's second law for the mass point takes the form:

$$(1) \quad F = ma \Rightarrow d^2r / dt^2 = \sqrt{r}$$

The remarkable feature of the example is that if one starts with the mass at rest at the apex (that is, with $r(t=0) = r'(0) = 0$), equation (1) does not only have the expected solution in which the mass remains at the apex forever ($r(t) = 0$ for all t), but also solutions corresponding to the mass taking off spontaneously at an arbitrary time T in an arbitrary radial direction:

$$(2) \quad \begin{aligned} r(t) &= (1/144)(t - T)^4 \quad \text{for } t \geq T \\ r(t) &= 0 \quad \text{for } t \leq T \end{aligned}$$

There is no doubt that Norton has found an interesting example of a differential equation without a unique solution. It is interesting because it apparently corresponds to a physical (even if idealized) situation within Newtonian mechanics. But the question is:

non-conservation of energy (see e.g. Alper and Bridger cited in Norton 2006, p. 13) and, in the space-invader case, non-conservation of particle number (and hence an ambiguity in the very specification of the physical system), see e.g. Malament (2007).

is the mass on the dome a genuine Newtonian system – that is, a system which satisfies Newton’s laws of motion?²

2. Is Newton’s first law satisfied for the mass on the dome?

As Norton recognizes, the spontaneous motion of the mass described by equation (2) may make one wonder whether Newton’s first law is satisfied at the crucial moment $t=T$ (the last moment at which the mass is at rest). Nevertheless, Norton (2003 and 2006) argues that the mass on the dome example does indeed satisfy this law both in an ‘instantaneous’ form (if $F=0$, then $a=0$) and in the ‘non-instantaneous’ form which Newton originally proposed (see below). This is so, Norton says, since for any moment up to and including $t=T$, there is no acceleration of the mass point, the net force acting on the mass vanishes, and the mass can be said to be in ‘uniform motion’ (in a state of rest). As far as I can see, however, this conclusion is much too fast.

Newton’s original formulation (in the 1729 translation of the *Principia*) of his first law is as follows:

(N1): Every body continues in its state of rest, or of uniform motion in a right line, unless it is compelled to change that state by forces impressed upon it.

There are at least two problems in maintaining that Norton’s mass on the dome satisfies (N1). First, take a look at the word “continues” in (N1): If a body is not compelled to change (at t) then it *continues* in its state of motion. Can that mean anything but: if it is not compelled to change at some moment t , then at least in the very next moment (even if that moment is infinitesimally close) the state is unchanged? If this is right then the motion in Norton’s system does not conform to Newton’s first law at $t=T$, since – for any $\varepsilon > 0$ – it is not true that the mass *continues* in its original state at $t+\varepsilon$. This issue is closely related to Norton’s own, and in my view correct, suspicion that the phrasing “compelled to change” in (N1) “...suggests that changes of motion must be brought about by forces acting at the same time as the change, if not even earlier” and so it suggests that “...forces must be first causes” (Norton 2006, p. 6).³ There is no first cause which forces the particle to move and so, according to (N1), it remains at the apex.⁴

The second (though related) problem concerns Norton’s claim that the mass at $t=T$ is in a state of rest (a particular case of a state of uniform motion). Clearly, Norton needs to say that the mass is moving uniformly (being at rest) at $t=T$ since, by Newton’s first law, it must be so since it has been at rest for all $t<T$ and since there are no forces acting on the mass at this moment. But the question here is what exactly is meant by uniform motion. Uniform motion means travelling equal (for instance, zero) distances in

² For a detailed mathematical analysis of the dome example, see Malament (2007). Norton (2006) discusses a list of possible objections to the dome being a Newtonian system but rejects each of these. As will be clear in the following, I disagree with some of his arguments.

³ See also the *Note added in proof* at the end of the paper.

⁴ As far as I can see, this point also accounts for why Norton’s “time reversal trick” (Norton 2003, p. 16) does not support that the acausal mass on the dome is a Newtonian system. The ‘reversed motion’ in which a mass point with a precisely adjusted initial velocity slides up the dome and halts exactly at the apex *is* consistent with Newtonian mechanics. In this reversed case, at $t = T$ the net force on the mass point vanishes, and for all $t > T$ the mass point is at rest and in uniform motion, and so no conflict arises with (N1). But insofar as forces should be first causes we cannot generate Norton’s motion from this allowed reversed case (when time is “run backwards” from some $t > T$, (N1) – understood as including the constraint that forces must be first causes – demands that the mass remains at rest for all t).

equal times – so that moving uniformly at some instant t seems to require considering, at least small, intervals of time at both sides of the instant t . Another way to state this point is that whereas one may speak of instantaneous velocity, or velocity at a specific moment (by the usual limiting procedure, $v = \Delta x / \Delta t$, $\Delta t \rightarrow 0$, the limit taken either from above or below), it makes little sense to speak of *constant* velocity at some specific moment unless reference is made to moments or time intervals on both sides (before and after) the moment in question.

As an illustration of this point, compare Norton’s mass on the dome with a standard example in Newtonian physics in which a white ball at rest on a billiard table is hit at $t=T$ by a red ball. At any moment $t < T$ the white ball is ‘moving uniformly’ (being in a state of rest). Moreover, for any moment $t < T$ we can find a successive moment $t + \varepsilon (< T)$ in which the ball is also moving uniformly (so that intervals on both sides of any moment $t < T$ can be found in which the velocity is zero). But even though the velocity of the white ball is still zero at the time of collision ($t=T$), we are not saying that it is moving uniformly at this moment. Of course, in this case we can attribute the non-uniform motion to the fact that a force acts on the ball – but it is equally true that the velocity at any moment $t > T$ is non-zero and therefore that the velocity is not *constant* at $t=T$.

Norton notes that if one were to insist (as I do) that $r'(T) = r''(T) = 0$ is not a sufficient condition for identifying a ‘state of rest’ or ‘uniform motion’, one would be “creating difficulties with other canonical examples” (2006, p. 6). For instance, Norton alludes to a harmonic oscillator, like a mass attached to a spring, passing through the origin: “We normally think of the mass at just that one moment as moving inertially—there is no net impressed force, so the velocity is constant, in the sense that the acceleration vanishes.” (2006, p. 7). But Norton does not explain (nor cite evidence for) why this is supposed to be an instance of ‘normal thinking’. Indeed, I believe that a more natural reaction to such examples is to say that the mass is *not* moving inertially when it passes through the origin (nor at any other moment), precisely because it is only at that moment that velocity, acceleration and net-impressed force vanish.⁵

3. Is Newton’s first law really needed? Inertial frames and the notion of time

In Norton (2003), it is argued that we do not really need Newton’s first law in order to do Newtonian physics or, at least, that we only need it as a special case of the second law. Thus, even if my critical remarks in the last section – concerning whether Newton’s first law in its original form is satisfied – are accepted, Norton could respond that the mass on the dome is nevertheless a Newtonian system satisfying Newton’s laws. More specifically, Norton claims that all a system needs to comply with in order to be Newtonian is an *instantaneous* version of Newton’s first law:

(NI*): In the absence of a net external force, a body is unaccelerated.

This law is indeed satisfied for the mass on the dome since for all $t \leq T$ (including $t=T$), $a(t) = r''(t) = 0$. Now, (NI*) amounts to reducing Newton’s first law to a special case of the second, $F = ma = 0$, so the question is whether this move – making Newton’s first law redundant – is legitimate within Newtonian physics (even if it is often done in

⁵ Note that the third derivative of the position is non-vanishing at the moment when the harmonic oscillator passes through the origin. In the mass on the dome case, the fourth derivative is ill-defined at $t=T$. Uniform motion would seem to require, then, the vanishing also of higher (>2) order time derivatives of position.

physics textbooks). I see at least two (related) reasons to resist the idea that (NI*) is enough for Newtonian physics, and therefore two reasons to resist that the first law is redundant in the sense of being merely a special case of the second law.⁶ The first reason has to do with inertial frames and the second with the notion of time in Newtonian mechanics.

Inertial frames

According to a tradition going back to Neumann and Lange (see e.g. DiSalle 2008) Newton's first law is closely related to the notion of an inertial frame of reference. It is only in such inertial frames that Newton's second law can be applied in its standard form, $F=ma$, in which F is an impressed (and not a fictitious) force. And it is precisely this standard form of Newton's second law that Norton invokes on the dome, so his example implicitly assumes that the dome is neither rotating nor accelerating in some direction. More generally, the use of the standard form of Newton's second law in the dome example presupposes that an inertial frame of reference exists. Now, one way to understand Newton's first law in a non-redundant way is precisely to construct it as an existence claim (see e.g. Friedman 1983, p. 117). On this account, Newton's first law implies that inertial frames (in which any putatively free particle will move in a straight line or stay at rest) exist.⁷

Newton himself did not need the first law to secure the existence of an inertial frame as he independently (of the laws) assumed the existence of absolute space. But if the notion of absolute space is rejected (either due to its inobservability, on some relational account of space – holding e.g. that the notion of space is necessarily bound up with that of laws – or with a view to relativity), Newton's first law in its original (non-instantaneous) form seems to be the only physical justification for assuming the existence of an inertial frame – and this justification is absent in 'Nortonian mechanics' without the full version of the law. In 'Nortonian mechanics', the existence of inertial frames will have to be postulated on an *ad hoc* basis and it may therefore be argued that the transition from Newtonian to Nortonian mechanics implies a loss of explanatory power.

However, it has been claimed, e.g. by Earman and Friedman (1973, p. 337), that a four-dimensional approach to Newtonian mechanics does indeed make (N1) redundant insofar as inertial frames – and, correspondingly, the trajectories of freely moving particles – are determined directly by the (affine) structure of spacetime. In this approach, then, it is thus not necessary to postulate the existence of inertial frames independently. Nevertheless, this objection might be countered by questioning whether the postulation of absolute spacetime structure explains the existence e.g. of inertial frames and the trajectories of force-free bodies. As Brown (2006, p. 24) notes:

⁶ Newton himself, of course, stated (N1) as an independent law and some modern authors, e.g. James L. Anderson (1990, 1192), likewise consider it erroneous to regard the first law as a special case of the second. See also discussion and references in Brown (2006, p. 15 ff.).

⁷ One could perhaps argue that Newton's second law could likewise be understood as an existence claim concerning inertial frames or that both laws of motion (or all three) jointly assert the existence of inertial frames (see e.g. DiSalle 2008, p. 6). Still, the connection between the first law and inertial frames might be more fundamental e.g. because free particles can be used, at least in principle, to construct such frames ("We must define an inertial system as one in which at least three non-collinear free particles move in noncoplanar straight lines; then we can state the law of inertia as the claim that, relative to an inertial system so defined, the motion of any fourth particle, or arbitrarily many particles, will be rectilinear" (DiSalle 2008, p. 5)).

In what sense is the postulation of the absolute space-time structure doing more explanatory work than Molière's famous dormative virtue in opium? ... It is simply more natural and economical – better philosophy in short – to consider absolute space-time structure as a codification of certain key aspects of the behaviour of particles (and/or fields).

The notion of time

The second reason to question whether Newton's first law is redundant in mechanics has to do with the role and notion of time in the theory. Recall first Newton's famous distinction:

Absolute, true, and mathematical time, of itself, and from its own nature, flows equably without relation to anything external, and by another name is called duration: relative, apparent, and common time, is some sensible and external (whether accurate or unequable) measure of duration by the means of motion, which is commonly used instead of true time; such as an hour, a day, a month, a year. (Newton 1729, p. 6)

As argued in detail in Zinkernagel (2008), it is reasonable to question whether this notion of absolute time is at all intelligible. In particular, it is very hard, if not impossible, to specify the meaning of 'flow' (or 'equable flow') without relation to anything external. Indeed, 'equable flow' involves the idea of steady or uniform motion, which more than suggests a reference to the motion of a physical system. In turn, the notion of uniform motion is instantiated by the free particles described by Newton's first law of motion. Now, as Newton himself notes in the Scholium, there might not be any real systems in uniform motion (due to friction and the universality of gravitational attraction) – hence uniform motion is an idealization. But just as Newton's 'common time' refers to real physical systems which can be used as clocks, it can be argued that for his absolute time to make sense, it must be seen to refer to idealized physical systems which can be used as perfect clocks, such as a free particle in uniform motion. In this way, (N1) and freely moving particles provide an implicit definition – and thus a sensible notion – of absolute time in Newtonian mechanics.⁸ Insofar as a sensible notion of absolute time is part of the theory, it is not clear that we can do Newtonian mechanics without (N1).

Of course, the motions of ideal clocks in Newtonian mechanics may be either uniform or non-uniform (accelerated) so one might well ask whether Newton's second law – and associated idealized physical systems such as the mathematical pendulum – could not, after all, be sufficient for an implicit definition of time in the theory. However, I think there are reasons to believe that this suggestion for making redundant (N1) will not work. First, as argued above, Newton's second law can only be applied in its standard form (and so provide a measure of absolute time via e.g. a mathematical pendulum) if the system is at rest in an inertial frame, and such a frame may need (N1) for its justification. Second, and in accordance with the above discussion, the notions of an "equable flow of absolute time" and of "equality of time intervals" can be seen to be interdependently defined by (N1) since equal intervals of time are those in which a free

⁸ This is consistent with a version of relationism about time in which time is dependent upon, but not reducible to, physical systems which can serve as clocks (see Zinkernagel 2008 and Rugh and Zinkernagel 2009a for detailed discussions of this version of relationism).

body moves equal distances.⁹ In other words, absolute time and uniformly moving systems (including uniformly rotating systems; see below) may be taken to be mutually dependent notions.

However, it could still be argued that absolute time might just as well be implicitly defined through Newton's second law and systems involving forces – e.g. systems involving uniform circular motion, such as a satellite orbiting a planet, in which equal time intervals can be defined as those in which the satellite moves equal distances (or through equal angles). Nevertheless, uniform circular motion is not uniform motion and the latter notion may be argued to have conceptual priority regarding the definition of time. In part because force (and forced motion) is defined by Newton (1729, p. 4) in terms of deviations from uniform motion; and in part because application of the second (instantaneous) law of motion seems to require a prior specification or notion of time (e.g. in order to use the orbiting satellite as an ideal clock one presupposes that the magnitude of the centripetal force is constant in time).¹⁰ This suggests that the relation between (N1) and time may be more fundamental than that between time and Newton's second law.

4. Causality in Newtonian mechanics and beyond

The notion of time (and that of ideal clocks) in Newtonian mechanics is also central to understand the 'causal message' of the theory. Norton's overall idea is that causality plays no fundamental role in modern physics and, in particular, no such role in Newtonian mechanics. By contrast, my view is that there is a clear 'causality content' in Newton's first two laws which can be captured as follows:

'Causal core' of Newtonian mechanics.

A body in uniform motion continues its motion unless the body is caused (by a force) to change its motion (accelerate). The same causes (forces) acting in the same circumstances will have the same effects.

The first part of this causal core is slightly more general than Newton's first law since uniform motion is not restricted to be straight line motion. This is in conformity with Newton's own remarks in the *Principia* (after stating his first law) which allude to uniformly rotating systems such as spinning planets.¹¹

The second part of the causal core is an abstract way of formulating the causality content in Newton's second law, which does not presuppose the precise form of this law. This second part is particularly important when discussing non-uniform periodic systems as it implies that a body in non-uniform periodic motion repeats its earlier

⁹ The close relation between (N1) and absolute time in Newtonian mechanics has been emphasized also e.g. by Barbour: "...the law of inertia [N1] itself has two quite distinct parts: the rectilinearity of the motion and the uniformity of the motion. These correspond, respectively, to absolute space and absolute time" (Barbour 1989, p. 28).

¹⁰ This point is closely related to one made earlier, namely that Norton's instantaneous form of Newton's (first or) second law in which $F=ma=0$ is not a sufficient condition for uniform motion. In consequence, one cannot define absolute time from this special case of Newton's second law (since a time interval, in which the acceleration is constantly equal to zero, must be presupposed when integrating up $F=ma=0$ to get uniform motion).

¹¹ Right after stating the first law, Newton writes: "A top, whose parts by their cohesion are perpetually drawn aside from rectilinear motions, does not cease its rotation, otherwise than as it is retarded by the air. The greater bodies of the planets and comets, meeting with less resistance in more free spaces, preserve their motions both progressive and circular for a much longer time" (Newton 1729, p. 14).

states of motion unless caused to change by additional forces. Thus, for instance, whenever the ‘mathematical pendulum’ reaches its top position, it will start falling. One direct consequence of the causal core is thus that it guarantees the continuity of a system in motion – and in particular the continuing motion of an ideal clock – and therefore also the continuing ‘flowing’ of absolute time.

Norton (2003) claims that a causal fundamentalist is confronted with the dilemma that either a causal principle restricts the content of our science or it does not. This is a dilemma according to Norton since no restricting causal principle which holds for all of science is forthcoming whereas a causal principle which does not imply such restrictions would be an “empty honorific”. But, at least as far as Newtonian mechanics is concerned, the causal core does impose restriction on science: It rules out non-continuous motion in the sense described above and, in particular, it rules out examples like Norton’s mass on a dome to be admitted into Newtonian mechanics.

If the relationism about time mentioned above – which tie the notion of time to that of physical processes which can be used as clocks – is on the right track, Newtonian mechanics embodies a close link between time and the causal core. This close link can be argued to proliferate into Newtonian mechanics’ supposed successor theories and in this sense pave the way for a more general (i.e. beyond Newton) causal fundamentalism. There is no room here for arguing this point in detail so what follows is merely a few notes and references suggesting how such an argument might be constructed.

In special relativity, the causal core of Newtonian mechanics is valid as it stands, and, just as in Newtonian mechanics, the physical basis for the notion of time in this theory may be argued to rest on physical processes which can be used as clocks. Indeed, Einstein himself argued that the physical meaning of the (space and) time coordinate(s) is (are) given in terms of (measuring rods and) standard clocks. A standard clock in an inertial system in special relativity is just like the (idealized) clocks in Newtonian mechanics and such standard clocks obey the causal core.

In general relativity things are less clear: On the standard interpretation, gravity is not seen as a force (but rather as curvature of space-time) and (N1) does not hold in general so one might well ask whether there is room in the theory for the causal core. Still, the physical interpretation – and the correspondence with empirical tests – of the theory is established in terms of (rods and) standard clocks. In particular, a relationist account of (space-)time in GR – in which a sensible notion of time is (non-reductively) dependent on physical clock systems may be argued to be more satisfactory than substantialist alternatives (see Rugh and Zinkernagel 2009a).

As regards quantum mechanics, the principle of ‘same causes, same effects...’ in the causal core does not hold in general due to the probabilistic nature of the theory. Nevertheless, the time dependence of wave functions, and hence the reference to the evolution of a quantum system, refers to the ordinary (classical) conception of time. So if the suggested relationism for time is accepted then a case might be made for the necessity of classical described clocks in (the interpretation of) quantum theory (see Rugh and Zinkernagel 2009b). A similar point can be made for quantum field theory in which both time and space are treated as classical background variables. Indeed, from a relationist premise, Teller (1999, p. 321) argues in the context of quantum field theory that “...space-time facts are facts about actual or potential space-time relations between physical bodies. The presupposed space-time relations are classical, exact valued; so the presupposed physical bodies between which the relations do or would have to be taken, in this respect, are classical too”.

Summary and conclusions

I have argued that Norton's mass on the dome system fails to obey Newton's first law in its standard formulation. Moreover, Norton's instantaneous formulation of Newton's first law seems insufficient as a replacement for the original version since the notion of inertial frames in such a modified theory lacks a physical justification and since an intelligible notion of time in Newtonian mechanics appears to be closely tied to Newton's first law in its standard form. I therefore claimed that the mass on the dome is not an acausal Newtonian system because it is not a Newtonian system. I suggested (but in no way proved) how the causal content (or causal core) of Newtonian mechanics – given its close connection to the notion of time – may play a central role also in relativistic and quantum theories. In this sense, the possibility remains that a principle of causality (captured by the causal core) – *pace* Norton – plays a fundamental role in physics.

Note added in proof (footnote 3)

Carl Hoefer notes (in private communication) that my understanding of (N1) would seem to rule out also standard examples in Newtonian mechanics in which a force is "turned on smoothly". In such cases, $F=0$ up to some moment $t=T$, and $F>0$ for $t>T$, so (N1) is apparently violated at $t=T$ (since the system does not continue in its original state for any $t > T$ even though $F=0$ for $t \leq T$). Moreover, as Norton (2006, 8) hints, the standard use of continuously varying forces (and continuously varying trajectories) in Newtonian mechanics seems to weaken the demand for forces being first causes. However, in standard examples involving continuously varying forces the physical situation may just as well be described via a sequence of discrete (first cause) forces, as can be seen e.g. in Newton's analysis (mentioned by Norton) of planetary orbits using polygonal trajectories in which a series of discrete forces act momentarily at the beginning of each segment (and the limit of vanishing segment size is taken at the end). This equivalence between a continuously varying force and a sequence of discrete forces is *absent* in the dome case: Physically, we cannot attribute any first cause to the dome motion whereas a smoothly turned-on force can be described in first cause terms (the "turning on" can be seen as a first cause of the change of motion, even if it can be modelled also by a continuously varying force).¹² Mathematically, the lack of equivalence in the dome case is reflected by the fact that the Newtonian *difference* equation (for a polygonal path of motion) corresponding to Norton's *differential* equation (1), in contradistinction to this latter equation, does have a unique solution, namely $r(t) = 0$ for all t .¹³ Thus, as no first cause can be associated with Norton's particle at the apex it stays put.

¹² When the "turning on" is modelled by a series of discrete forces, the first (or indeed any) non-vanishing discrete force in the series can be identified with a first cause, and my point is that in standard cases – but not in Norton's – we can use both the continuously varying force approach and that of a series of discrete (first cause) forces.

¹³ If one takes $r(t+\Delta t) = r(t) + v(t)\Delta t$; $v(t+\Delta t) = v(t) + r(t)^{1/2}\Delta t$; and $r(0) = v(0) = 0$, the difference equation for the mass on the dome is found to be $(r(t_{n+2}) - 2r(t_{n+1}) + r(t_n))/h^2 = (r(t_n))^{1/2}$ in which $h = \Delta t$ (step size) and $t_n = nh$. Since $r(t_0) = r(t_1) = 0$, this equation has the unique solution $r(t_n) = 0$ for all n (and thus the same solution in the limit $h \rightarrow 0$). I first saw this equation on an anonymous internet forum post by a user named "jason1990". Hans Henrik Rugh points out (in private discussion) that there are many ways to "discretize" a differential equation, and in this sense the differential equation quoted here is not unique for the dome. However, the difference equation approach is arguably the more fundamental one for Newton and, as far as I can see, the demand of forces as first causes acting at the beginning of each segment does in any case single out the quoted difference equation.

Acknowledgements

It is a pleasure to thank Carl Hoefer and Svend Rugh for comments and discussion. I also thank the audiences at the “Singular causality, counterfactuals and mental causation” workshop in Granada and at the I EPSA Madrid conference for comments during presentations of this work. Financial support from the Spanish Ministry of Education and Science (project HUM2005-07187-C03-03) is gratefully acknowledged.

References:

- Anderson, J. L. (1990) “Newton’s first two laws of motion are not definitions”, *American Journal of Physics*, 58, 1192-1195.
- Barbour, J. (1989) *Absolute or Relative Motion? Vol. 1: The Discovery of Dynamics*. Cambridge, CUP.
- Brown, H. (2006) *Physical Relativity*. New York: Oxford University Press.
- DiSalle, R. “Space and Time: Inertial Frames”, *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, E. N. Zalta (ed.),
URL = <<http://plato.stanford.edu/archives/fall2008/entries/spacetime-iframes/>>.
- Friedman, M. (1983) *Foundations of Space-Time Theories*. Princeton: Princeton University Press.
- Malament, D. (2007) “Norton's Slippery Slope”, manuscript.
<http://philsci-archive.pitt.edu/archive/00003195/>
- Norton, J. D. (2003) “Causation as Folk Science”, *Philosophers' Imprint* Vol. 3, No. 4
<http://www.philosophersimprint.org/003004/>; reprinted in H. Price and R. Corry (eds.),
Causation and the Constitution of Reality. Oxford: OUP, 2007.
- Norton, J. D. (2006) “The Dome: An Unexpectedly Simple Failure of Determinism”, manuscript. philsci-archive.pitt.edu/archive/00002943.
- Newton, I. (1729) *Mathematical Principles of Natural Philosophy*. Trans. by A. Motte and F. Cajori. Berkeley: University of California Press, 1962.
- Rugh, S. E. and Zinkernagel, H. (2009a) “On the physical basis of cosmic time”. *Studies in History and Philosophy of Modern Physics* 40, 1-19.
- Rugh, S. E. and Zinkernagel, H. (2009b) “Time and the cosmic measurement problem”. In preparation.
- Russell, B. (1913) “On the Notion of Cause”, *Proceedings of the Aristotelian Society* 13, 1-26.
- Teller, P. (1999) “The ineliminable classical face of quantum field theory”. In T.Y. Cao, (ed.) *Conceptual Foundations of Quantum Field Theory*, Cambridge: Cambridge University Press, 314-323.
- Zinkernagel, H. (2008) “Did time have a beginning?”, *International Studies in the Philosophy of Science* 22:3, 237-258.