Reduction, Autonomy, and Causal Exclusion Among Physical Properties

Alexander Rueger

Department of Philosophy
University of Alberta
Edmonton, Alberta
T6G 2E5
Phone: (780) 492-9028

e-mail: alex.rueger@ualberta.ca

Abstract

Reduction, Autonomy, and Causal Exclusion Among Physical Properties

Is there a problem of causal exclusion between micro- and macro-level physical properties? I argue (following Kim) that the sorts of properties that in fact are in competition are *macro* properties, viz., the property of a (macro-) system of 'having such-and-such macro properties' (call this a 'macro-structural property') and the property of the *same* system of 'being constituted by such-and-such a micro-structure' (call this a 'micro-structural property'). I show that there are cases where, for lack of reducibility, there is a *prima facie* intra-level causal competition between the two kinds of properties. The problem can be resolved without giving up on the causal efficacy of the macro-structural properties if we understand instances of macro-structural properties to be *parts* of micro-structural property instances. The parthood relation between both kinds of property instances can be mapped onto the way physical theory deals with the relation of their descriptions in the framework of perturbation theory. The application of this framework to the problem of emergent properties is discussed.

Reduction, Autonomy and Causal Exclusion
Among Physical Properties

1. Introduction

Nonreductive physicalists have forcefully argued that if the 'problem of mental causation' has no other solution than epiphenomenalism, then we have more difficulties than we may have bargained for. A seemingly straightforward generalization of the problem, so they claim, will show that not only are mental properties causally impotent but so must be chemical, geological, biological, etc., properties. Since this is an absurd consequence, the initial form of the problem must be ill-posed. In Van Gulick's formulation, if the argument for epiphenomenalism of the mental were correct, it would show that mental "properties are epiphenomenal [in the same] sense in which chemical and geological properties are also epiphenomenal." We thus don't have much reason to be concerned: mental properties "seem to be in the best of company [because] no one seems worried about the causal status of chemical properties."[1]

Against this type of rescue attempt on behalf of nonreductive physicalism Kim has recently claimed that, properly understood, the problem of mental causation does not generalize and thus does not have any obviously absurd consequences for the other 'special sciences' (1997; 1998). The only (physicalist) alternative to epiphenomenalism that he finds viable, however, is abandoning nonreductive physicalism and adopting a form of reductionism.

I would like to show that the original point made by van Gulick and others has some validity and that it is, *pace* Kim, relevant to the problem of mental causation. The problem *does* generalize; in fact, it generalizes so as to apply even within physics itself. That is the domain in which I shall discuss the issue: I'll argue that we can resurrect the problem in the relation between macro and micro level physical properties — not, however, as an *inter*-level problem but as an *intra*-level problem. This will involve an argument that certain macro level regularities and properties are not reducible to micro-based regularities and properties on which the former supervene. What results is an apparent *causal competition* between such properties — a competition analogous to the problem of mental causation. Do we therefore have to be epiphenomenalists about macro physical properties? It seems that physicists solve the problem in a different way, by treating — as can be shown in the formalism of setting up the problem — instances of the macro properties as *parts* of instances of the micro-based properties. Since parts and wholes do not stand in competition for causal sufficiency, the analogon of the problem of mental causation does not arise. Although I do not make any claims about mental causation, the implication of this discussion is that the strategy that succeeds in physics may well work in the mental causation problem too. And since the need to distinguish between causal sufficiency and causation therefore arises already within physics, we can defend nonreductive physicalism against those critics who have argued that making such distinctions violates the very idea of physicalism (Crane 1995).

2. Generalization of the Exclusion Problem

Here is a nutshell formulation of the problem of mental causation as I understand it for the purposes of this paper: As nonreductive physicalists we want to see the world structured in a hierarchy of levels of objects and properties, including the physical and the mental level. We seem, however, to be prevented from holding simultaneously that (i) the mental (higher) level is autonomous or nonreducible to the physical (lower) level, (ii) that the physical (lower) level is causally closed — that is, every physical event has a complete physical cause, and (iii) the properties at the mental level are causally efficacious, in addition to (and not overdetermining) the causally efficacious properties at the physical level. We seem prevented from accepting all these claims if we also hold the *principle of causal exclusion*: If an event or property *x* is causally sufficient, in the circumstances, for an event or property *y*, then no *z*, wholly distinct from *x*, can be causally relevant (and hence not sufficient) for *y*.[2] Given this principle, affirming the autonomy of the mental appears to drain it of causal influence in the world; securing the causal efficacy of mental properties undermines their autonomy and collapses nonreductive physicalism into reductive physicalism. Let's call this, more generally, the problem of causal exclusion.

It seemed clear that such problems, if they are genuine problems, would not only affect the relation of mental and physical properties but more generally the relation between any properties at any pair of higher/lower levels. The problems would arise for biological (higher level) and chemical (lower level) properties as well as for chemical and physical properties. But this kind of list still does not exhaust the scope of the problems:

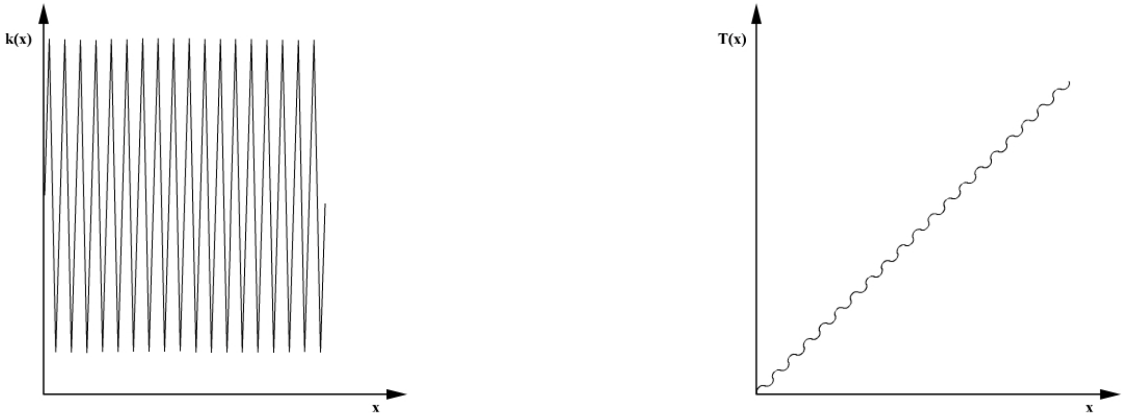
nothing should prevent the difficulties to crop up as well *within* the discipline of *physics itself*, if we picture the physical world as layered into, e.g., micro and macro levels, and assume a kind of autonomy or nonreducibility for the levels.

But does the problem of causal exclusion really generalize in this way? Kim has recently argued that the reasoning behind the generalization attempts is flawed (1997; 1998, 78-87, 112-118). Under Kim's construal of the problem, we have causal competition only between properties at the same level but not between properties at different levels. The micro properties of a system are properties of the (micro) constituents of the system, the macro properties are properties of the system itself. To point to a property of "having such and such a micro constitution (constituents plus their relations)" is to point to a macro property, a "micro-based" or "microstructural property" in Kim's terms — a property that the system *qua* macro object has, not a property of the micro constituents of the system. The causal powers of such properties will trivially be different from the powers of the constituents and therefore there cannot be competition. Only when we are dealing with properties at the same level, e.g., a mental property *of me* and a physical property *of me*, can the problem of causal exclusion arise.

Let's look at an example where we can actually compare a macroscopic description of a system with a microscopic one. I choose an example of the simplest possible form, that is, a case where quantum mysteries or complicated interactions don't play a role. Nonetheless, in the formal treatment the case is typical for the whole class of problems involving different levels or scales. Consider the treatment of steady state heat conduction in a one-dimensional rod of length L (Hinch 1990, pp. 126f.; Kevorkian/Cole 1996, 614-17). This system is described in terms of its temperature T(x) and its thermal conductivity k(x) which both vary in dependence on the spatial variable x. We assume there to be no heat sources and scale the spatial variable so that the length of the rod L = 1. The boundary conditions at the ends of the system are T(x = 0) = 0 and T(x = 1) = T_1 (see Figure 1b). The equation for heat conservation in this system is

$$(d/dx)\ [k(x)dT(x)/dx] = 0, \tag{1}$$

We assume a 'microscopic picture' of the rod: it consists of individual 'atoms', separated by empty space — a periodic lattice with a period of length $l = \varepsilon L$, with $\varepsilon \ll 1$. The conductivity k(x) will then be a rapidly oscillating function of position (see Figure 1 a). At the macroscopic level, the level at which we make measurements about the system's



(a) **(b)**

Figure 1 (a),(b).

temperature, etc., no such fast variations are observed: T and k are slowly changing function of (macro) position.

The task now is to establish the form of the heat equation at the macro level as well as its solution. This macro description of the system should be obtained from the microscopic description, Eq. (1), in the limit of $\varepsilon \to 0$, a kind of continuum limit. For very small ε, corresponding to almost vanishing micro periods l, we expect the behaviour of T to become effectively independent of the microscopic length scale; a description that satisfies this condition of being independent of the micro variables is what we call a macro description. Thus I label a description that refers to specifications of the system at the micro level, like Eq. (1), as D_{mic}; the macro description we are seeking is labeled D_{mac}. Given the widely different behaviour displayed by the macroscopic T and k (slow variation) and the corresponding microscopic quantities (rapid variation), it is — as we shall see in the next section — a highly non-trivial problem to show that we can in fact arrive at D_{mac} starting from D_{mic}, that is, to justify the macro level regularity from what we know about the micro-based regularity, Eq. (1).

I suggest that something like D_{mic} is what nonreductive physicalists like van Gulick have in mind when they talk about the microscopic level of a system and when they compare the mind-body relation with the relation of two descriptions of the same system at different scales or levels. D_{mic} is a description of the micro components of the system together with their relations; it refers, therefore, to a *micro-structural* property of the rod, the property of having such and such a micro constitution which is itself a macro level property. The sought-for macro description, D_{mac}, then refers to properties at the *same* level as D_{mic}. In analogy to D_{mic} I call D_{mac} the *macro-structural* description.

More precisely stated: D_{mic}, or Eq. (1), together with the boundary conditions specific to the rod, describes a regularity concerning the connection between the distributions of the properties k(x) and T(x) of the system: a distribution of properties k(x) in the rod (which is a *macro* property of the rod, even though local values of k(x) are specified at the micro level) 'causes' a property distribution T(x) (which again is a macro property even though individual values of T(x) are properties at the micro level).[3] Analogously for D_{mac} where a regularity involving distributions of the macro level conductivity and temperature is described. Therefore, both the micro-structural description D_{mic} and the macro-structural description D_{mac} refer to *macro* level regularities or causal relations.

D_{mac} supervenes on D_{mic}. There can't be changes in the macroscopic temperature distribution of the rod without corresponding changes in the microscopic distribution. Hence, the exclusion problem should arise in the form Kim acknowledges — as an *intra*level problem, a competition between the properties described by D_{mac} and those involved in D_{mic}. I take it that van Gulick and others were just speaking loosely in their arguments; what they really had in mind was what Kim, speaking more precisely, identifies as an intralevel problem of causal exclusion.[4]

According to Kim causal exclusion does not arise

> "for micro-based properties in relation to their constituent properties because the former do not supervene on the latter taken individually or as a group. Rather, they supervene on *specific mereological configurations* involving these microproperties — for a rather obvious and uninteresting reason: they *are* identical with these micro-configurations." (1998, 117f.)

So if the exclusion argument applies to our case, there still is the possibility that it can be rendered harmless by showing that the properties in D_{mac} supervene in the "rather obvious and uninteresting" sense of identity on those in D_{mic}. This is how Kim ultimately suggests to solve the intralevel problem: reduce mental properties to physical ones (in the functionalization sense of reduction; cf. Kim 1998, ch.4.) Thus, only if we can make it plausible that D_{mac} does not reduce to D_{mic} , i.e., that the macro-structural properties have some degree of autonomy from the micro-structural properties, have we created a serious instance of the causal exclusion problem within physics.

3. Non-Reducibility

Can D_{mac} be reduced to D_{mic}? And in what sense of 'reduction'? I suggest adopting a notion of reduction that is most natural in the kind of context we are discussing because it reflects the technical way in which physicists investigate the explanatory relation between D_{mic} and D_{mac} (Nickles 1973; Berry 1994; Batterman 1995; Rueger 2000a,b; 2001). We define:

> A theory Θ_0 reduces to a theory Θ just in case there is a uniform limit in a suitable parameter of Θ in which the solutions of Θ, u(x), go over into the solutions of Θ_0, $u_0(x)$.

'Uniform' means that this limit should hold for *all* x (for which solutions are defined), not just for a subset. In other words, given a perturbation expansion for the solutions of Θ

$$u(x) = u_0(x) + \varepsilon u_1(x)...,$$

in some 'small' parameter ε, we require that

$$\lim_{\varepsilon \to 0} u(x) = u_0(x) \text{ for all x.}$$

This is the sense of reduction in which, for instance, (some parts of) Special Relativity Theory (Θ) go smoothly over into Newtonian Mechanics (Θ_0), and hence Newtonian Mechanics reduces to Special Relativity. Intuitively, Θ_0 reduces to Θ in this sense if Θ is just Θ_0 plus some 'small' corrections; the behaviour of the system described by Θ becomes gradually and smoothly indistinguishable from the behaviour described by Θ_0 when we go to the limit $\varepsilon = 0$. A failure of reduction, by contrast, means that the Θ-behaviour is quite different (at least in some parts of the solution domain) from the behaviour described by Θ_0, however small we let ε become (different from 0 though). Note how this notion of reduction relates to the notion of Nagel-type reductions where we require that Θ_0 be *derivable* from Θ. If Θ_0 is not the uniform limit of Θ, then Θ_0 , a fortiori, can't be derived from Θ either; if Θ_0 and Θ are logically contradictory, however, they are not Nagel-reducible but can still be standing in a limit relation. So Nagel reducibility implies the limit sense of reducibility but not vice versa.

In the heat conduction problem, the question is whether D_{mic} (Eq. 1), in a suitable limit, uniformly converges to D_{mac}; that is, whether, in this limit, the macro-structural description becomes indistinguishable from the micro-structural description of the system and hence could be replaced by, or eliminated in favour of, the latter. A first attempt at a perturbative solution of Eq. (1) would define a parameter ε such that the expansion

$$T(x) = T_0(x) + \varepsilon T_1(x) + \varepsilon^2 T_2(x) + \ldots \qquad (2)$$

would reduce, in the limit $\varepsilon \to 0$, to the macroscopic solution $T_0.(x) = T_0$ (which should be independent from the micro variable x). The 'macroscopic limit' of the micro-structural description is the limit $\varepsilon \to 0$, with ε as the ratio of the length of the rod L and the micro period l (as defined before: $\varepsilon = l/L << 1$). This limit, however, turns out *not* to be uniform: T(x) does not uniformly converge to $T_0(x)$. For large x, the higher-order terms in the series (2) grow faster than the lower-order terms, thus destroying the asymptoticness of the expansion, the basic requirement on any reasonable expansion, namely that terms of higher order should be smaller than terms of lower order. For large x, the expansion ‘blows up’, the temperature of the rod increases without bound, which is clearly unphysical. This unfortunate behaviour of T(x) indicates that the micro behaviour of the system cannot be represented as essentially the macro behaviour, described by $T_0(x)$, plus small corrections. T(x) and $T_0(x)$ differ considerably, even for the smallest nonzero values of the parameter. Thus, $T_0(x)$ does not reduce to T(x) in our sense. D_{mac} retains some sort of autonomy from D_{mic}.

This autonomy of the macro-structural description can be further illustrated by looking at the technical remedy that has been developed to deal with problems in which the perturbation expansion of the solution, analogous to Eq. (2), breaks down. One can in such cases still construct a *uniformly valid approximation* in the form of a power series, starting with the macro solution $T_0(x)$ (although these series will not usually be convergent). The way to do this is to explicitly introduce *two* length scales in the micro description, the microscopic scale x and a macroscopic scale $\xi = \varepsilon x$. Thus we consider an asymptotic expansion of T as a function of the two independent variables:

$$(3) \qquad T(x, \xi) = T_0(x, \xi) + \varepsilon T_1(x, \xi) + \varepsilon^2 T_2(x, \xi) + \ldots$$

Inserting this expansion into Eq. (1) and replacing the derivative d/dx by $\partial/\partial x + \varepsilon\, \partial/\partial\xi$, (because we treat x and ξ as independent) gives, at the lowest order of the expansion:

$$(4) \qquad \partial/\partial x\, [k(x)\partial T_0(x, \xi)/\partial x] = 0,$$

which shows that $T_0(x, \xi)$ is indeed independent of the microscopic variations measured by x; it depends only on the macro level position ξ, as a macro level quantity should:

$$T_0(x, \xi) = T_0(\xi).$$

To determine the form of T_0 further, we have to go to higher orders of the expansion. We find that $T_2(x, \xi)$ can be prevented from increasing without bound for large x, and thus from wrecking the asymptoticness of the expansion, *only* if $T_0(x)$ satisfies the ("solvability") condition

(5) $$\partial/\partial\xi\ [k_{eff}\ \partial T_0(\xi)/\partial\xi] = 0,$$

where

$$k_{eff} = [\frac{1}{l}\int_0^l \frac{dx}{k(x)}]^{-1}$$

is the harmonic mean of the microscopic conductivity k(x), averaged over a micro period which makes it effectively independent from x and thus transforms it into a macro parameter.[5] Eq. (5), however, is precisely the desired regularity connecting macro conductivity and temperature; Eq. (5) is D_{mac}. Because we introduced the macro scale (ξ) as a second independent variable, we have gained the freedom to force the expansion Eq. (3) to be asymptotic by imposing Eq. (5) on it which effectively eliminates those terms in the expansion that grow too fast with increasing x. Thus, we have recovered from D_{mic} the macro level description D_{mac} in the limit $\varepsilon \to 0$ (see Figure 2).

$$\begin{tikzcd} D_{mac}: & k_{eff} \arrow[r] \arrow[d, dash] & T_0(\xi) \arrow[d, dash] \\ D_{mic}: & k(x) \arrow[r] & T(x) \end{tikzcd}$$

Figure 2.

The important point in this calculation is that the ‘derivation’ of D_{mac} from the micro-strcutural description inevitably involves quantities at *both* length scales. You can’t go in the limit from the micro-structural description Eq. (1) to the macro-structural description Eq. (5) without formally introducing two independent spatial scales at which the quantities change. Once these scales have been introduced, the macro-structural description itself has to be invoked to ensure that the perturbation expansion of the micro-structural description remains asymptotic. Despite the fact that there is no uniform convergence of D_{mic} to D_{mac}, and hence no reduction of the macro- to the micro-structural description, we can construct a uniformly valid approximation of the micro in terms of the macro. The price to be paid for this is the inevitable autonomy of the macro-structural description — you can’t get rid of it in favour of the micro-structural description alone.

Introducing the macro scale into D_{mic} reveals features characteristic of macroscopic bodies, e.g., properties like the measured temperature T or the ‘effective’ elasticity of a medium with a specific micro structure. To use a phrase that now has some currency in Philosophy of Mind: Different scales allow us to ‘see’ different patterns in the distribution of physical quantities; a behavioural pattern may be pertinent in a description at the macro level, but may be lost in a micro level description of the same system.[6] Recall Putnam’s famous case of the round peg of 1 inch diameter which does not fit through a square hole in a board of 1 inch diagonal extension (Putnam 1975, 295ff.; slightly modified). Putnam compares two possible explanations of this fact, a microscopic one in terms of the arrangement of molecules in peg and board, etc. (“the microstructural deduction”), and a macroscopic one in terms of the geometrical relations between the macroscopic objects involved. The ‘microstructural deduction’, if one were to actually perform it, would be an instance of the kind of asymptotic analysis considered above. We can expect a nonuniform limit relation between the micro- and macro-structural descriptions and thus the introduction of a macro scale into the lower level description will be required, confirming the autonomy or indispensability of the higher level.

The nonuniformity of the limit is the real reason for why the macro-structural description tells us ‘something different’ than the micro-structural description, and for why the former is not reducible to the latter. One extremely important service that the asymptotic expansion of D_{mic} in terms of two different scales does for us is to make it clear to what extent the macro-structural description, the behaviour of the system at the macro scale, is independent of, or insensitive to, the details of the system’s behaviour at the micro scale — what we saw for the measured temperature of the rod (cf. Eq.(4)). In the terms that are familiar from the philosophical discussion of mental properties: the 'multiple realizability' of the macro description, which Putnam mentions as its main advantage over the micro story, is an immediate outcome of the asymptotic procedure described (cf. Batterman 2000).

It is worth emphasizing an important feature of the limit notion of reduction applied to the relation of macro- and micro-structural descriptions. There is *no* claim here that the former are always and in general irreducible to micro-structural descriptions; whether or not we have non-reducibility is an empirical matter, to be decided case for case for given descriptions. In those cases where a uniform limit exists, the descriptions will be reducible; there are no *a priori* reasons for assuming that the appropriate limits are always non-uniform (see Berry 1994 or Batterman 1995 for examples). Intuitively, a nonuniformity will occur only when the behaviours of the system described by D_{mic} and D_{mac} are very different, as illustrated in our case of the temperature distribution.

There is an important worry about claims of the sort I made so far — that they are 'merely epistemic' results, results that depend completely on our human (or otherwise) interests and cognitive capacities. Thus what I called before the autonomy of the macro-structural properties is, on this view, an autonomy of description only, not a 'metaphysical' independence. We would then be led back to Kim's recent view that one has to "give up [the higher order property] E as a genuine property, only recognizing the expression 'E' or the concept of E." Such concepts are merely ways of "picking out first-order properties in terms of certain causal specifications that are of interest to us."

One cannot have it both ways: retaining E as an irreducible property and nonetheless have E play “an important role in a special, ‘higher-level’, science.” (Kim 1999, 17f.)

One way of arguing for the reality of macro-structural properties is to show that they have causal powers that are different from the causal powers of the underlying micro-structural properties. But any such attempt, of course, has to get around the argument of the causal exclusion problem which denies that macro-structural properties can be causally efficacious.

4. The Causal Efficacy of Macro-structural Properties

If macro-structural properties are ‘real’ and if we accept the basic claim of physicalism that the micro level is causally closed, then it seems mysterious how the former properties can have causal influence or efficacy in addition to the causal efficacy possessed by the micro-structural properties. Suppose a proper cause has to be sufficient, in the circumstances, for bringing about its effect. Any given effect would then have at least *two actually occurring sufficient* conditions, one micro- and one macro-structural one. Unless we are prepared to admit causal overdetermination as a regular and ubiquitous feature into our view of the world, we cannot tolerate such proliferation of causally sufficient conditions. Notice that this an *intra*-level formulation of the causal exclusion problem which usually is set up as an inter-level competition between micro- and macro level properties. I am concerned with competition among macro level properties, one kind of which are micro-based properties (described in D_{mic}) while the other sort consists of macro-structural properties that do not refer to micro constituents (described in D_{mac}).

The majority of the responses to the problem of mental causation as a special case of the causal exclusion difficulty (understood as an inter-level problem) try to resolve the issue by distinguishing, in more or less detail, between two kinds of causal relation. One kind is supposed to characterize the causal interaction at the physical (lower) level and the other is employed to secure a causal role for the mental (higher) level properties in addition to the causal relation at the lower level. Distinctions between "causal efficacy" (for the lower level) and "causal relevance" (for the higher level) (e.g., LePore/Loewer 1987) belong here as well as between "triggering" and "structuring causes" (Dretske 1993). If mental properties, for instance, are causally relevant to physical effects, they do not obviously compete in the role of cause with any physical properties that may either be also relevant or causally efficacious; mental properties as structuring causes of some effect E, i.e., as bringing about conditions that have to be in place for E to occur, do not compete with physical properties as triggering causes, i.e., as properties that bring about E in the conditions secured by the structuring cause. Against such attempts at "denying [the] homogeneity" of the causal relation Tim Crane (1995) has argued that the nonreductive physicalist who distinguishes different sorts of causation is effectively giving up one of the central motivations for physicalism itself: If mental properties cause in a different sense than physical properties, what is physicalism all about? Wasn't physicalism motivated by the hope that one could reconcile — somehow! — the reality of mental causation and the completeness of physics?

"[T]he problem with denying homogeneity is that it is now impossible even to

> state the original motivation for physicalism: the conflict between mental causation and the completeness of physics. So there is no clear reason for saying that these mental phenomena are ‘constituted by’ or ‘realized by’ physical phenomena. Physicalism has lost sight of its motivation.” (Crane 1995, 235)

Nevertheless, I think we need some such distinction. But it is *not* that we need it because otherwise there could be no mental causation. The distinction between ways of causing is not an *ad hoc* maneuver to fight off the threat of epiphenomenalism. Rather, the point I want to make is that we need a distinction between ways of causing already *within physics itself*.

I want to show this on the basis of the nonreducibility of macro-structural properties established above. The fact that the exclusion problem *does* generalize (albeit as an intra-level problem), requires us to adopt different ways of causing — and since we need to do this within physics itself, it should hardly be regarded a strategy that is detrimental to physicalism.

Kim has criticized attempts to establish the desired distinction by invoking the principle of causal exclusion: if one cause (at the lower level) is, by assumption, in the circumstances *sufficient* for the effect, what role other than a nominal one does the other presumed cause (at the higher level) still have to play?[7] Consider Block’s famous case of the red cape that angers the bull. There is, suppose, a higher-level property in the cape, besides the property of being red, viz., ‘provocativeness’ which supervenes on the cape’s redness. What causes the bull’s anger?

“[I]f the color of the cape is, in and of itself, a sufficient cause of the anger (at

> least, sufficient in the circumstances), what *further* causal work is left for its provocativeness? What special contribution of its own can the cape's provocativeness make in the causation of the anger? The answer obviously is none: given the color of the cape as a full cause, there is no *additional* causal work left for its provocativeness, or anything else." (Kim 1998, 53)

We can rephrase this, as an intra-level problem, for the comparison of the properties of the heated rod described by D_{mic} and D_{mac}. If the micro-structural property of having a certain distribution of micro conductivity k(x) over the whole rod, in the circumstances, is sufficient for bringing about a certain distribution of the micro temperature T(x) in the rod, what further causal work is left for the further macro-structural property of the rod, viz., the property of having a certain distribution of the macro conductivity k_{eff}, given that the macro-structural temperature distribution $T_0(\xi)$ supervenes on the micro-based T(x) distribution? After all, supervenience means that once the k(x) distribution causes the T(x) distribution, the values of $T_0(\xi)$ over the rod are fixed. But, according to D_{mac}, the k_{eff} distribution is also sufficient, in the circumstances, for bringing about the $T_0(\xi)$ distribution.

The problem then is: how can two *actually existing* sets of conditions, the set C_{mic}, referred to in D_{mic}, and the set $C_{mac,}$ referred to in D_{mac}, both be sufficient for some effect E? The only way this is possible is if one set, say C_{mac}, is *contained within* the other, C_{mic} (or, of course, the other way around). The distribution of k(x) is sufficient, through the supervenience of $T_0(\xi)$ on T(x), for the $T_0(\xi)$ distribution and so is the k_{eff} distribution. This is possible only if either the property of having a certain k(x) distribution is *part of* the property of having a certain k_{eff} distribution, or the other way around. Since properties themselves cannot stand in part-whole relations, we have to say more precisely: Either an *instance* of the property of having a certain k(x) distribution is part of an *instance* of the corresponding property involving k_{eff}, or *vice versa.* Understanding the seemingly competing sufficient conditions as standing in the relation of part and whole allows us to avoid the principle of causal exclusion and therefore the original exclusion problem. The principle ruled that if a property P is causally sufficient, in the circumstances, for a property Q, then no property Z, wholly distinct from P, can be causally relevant (and hence not sufficient) for Q. This obviously does not apply to those properties (property instances) where P is contained in Q, or *vice versa*.

Take the case where the k(x) distribution is sufficient for the distribution of macro temperature ($T_0(\xi)$). What makes the former sufficient for causing the latter is the fact that it contains, as a part, the macro conductivity distribution (k_{eff}) which, by itself, is also sufficient for the macro temperature distribution. The micro conductivity 'profile' of the rod, in other words, contains, besides the sufficient macro distribution, 'too much other stuff' — stuff that isn't really required for causing the $T_0(\xi)$ distribution. This 'other stuff', plausibly, consists in microscopic details the presence or absence of which makes no difference to the resulting macro distribution. The rod's macro conductivity profile, by contrast, is sufficient as well as required for the $T_0(\xi)$ profile. This seems like a metaphysical advantage of the k_{eff} distribution over the k(x) distribution, and Stephen Yablo (1992; 1997) has promoted this sort of advantage to a characterization of what a cause should be: sufficient *and* required for its effect, that is, “commensurate” (or “proportional”) to its effect.

> “... [C] can be causally sufficient for [E] although it incorporates indefinite amounts of causally extraneous detail, and causally relevant to [E] even though it omits factors critical to [E’s] occurrence. What distinguishes causation from these other relations is that causes are expected to be commensurate [or proportional] with their effects. [...] Thus faced with a choice between candidate causes ... , the more proportional of the two is, other things equal, to be preferred [as cause].” (Yablo 1992, 434f.)

A more technical characterization of the notions of ‘requiredness’ and ‘sufficiency’ goes like this: For property instances C, C*, C’, and E: C is *required for* E iff there is no $C^* \subset C$ such that E would have occurred if C* had occurred without C. And C is *enough* (or sufficient) for E iff for all C’such that $C \subset C'$, E would have occurred if C had occurred without C’ (after Yablo 1997, 266f.).[8]

For example, instances of scarlet and red are both sufficient, suppose, to bring about the bull’s anger. This is possible without competition because it is part of being scarlet to be red. But it may be that only red is a cause of the anger because being scarlet is ‘more than enough’ for bringing about the bull’s anger; being scarlet is being red plus something else, not required for the bull’s state of mind. Any other shade of red would have had the same effect. This can also be rephrased for the case of redness and the second-order property of provocativeness. The requirement of proportionality would now have to single out provocativeness as sufficient and required and hence as the proper cause of the anger; red, in turn, would contain ‘too much’ for the effect, namely provocativeness plus something else, not required for the bull’s anger.[9]

So it turns out that we need a distinction, although not really between ways of causing but rather between causal sufficiency and proper causation, the latter being characterized by the requirement of proportionality. The properties of having a certain distribution of micro conductivity and a certain macro conductivity profile in the case of the heated rod do not compete for causal sufficiency with respect to the resulting macro temperature distribution because parts and wholes cannot so compete. But they do compete, according to Yablo’s suggestion, for the role of cause. Whichever establishes itself as sufficient *and* required, will come out as *the* cause. Thus it is the macro conductivity distribution which properly causes the macro temperature distribution in the rod: The macro conductivity profile is sufficient for bringing about the temperature distribution $T_0(\xi)$ but the former is also required for the latter since different distributions of k_{eff} will lead to different temperature distributions. The micro-structural property of having a certain micro conductivity profile, although sufficient for $T(x)$[10], and hence sufficient for the supervening $T_0(\xi)$, is *not required* for the $T_0(\xi)$ distribution because the same $T_0(\xi)$ profile will be brought about by any micro distribution that is characterized by a small value of ε in Eq. (3). Each non-zero value of ε indicates a different $k(x)$ distribution, causing a different $T(x)$ profile; all of these distributions, however, are equivalent with respect to leading to the same distribution of $T_0(\xi)$.

The part-whole relation, of course, is asymmetric but we have not so far given a reason for why the k_{eff} distribution, for instance, should be contained in the $k(x)$ profile rather than the other way around. Both ways would formally solve the exclusion problem. To see the reason for the asymmetry of the relation, compare the description of our system in macro-structural terms in Eq. (5) with the description that results from understanding the problem as a perturbation of a description in micro-structural terms, i.e., after inserting the expansion $T(x, \xi) = T_0(x, \xi) + \varepsilon T_1(x, \xi) + \ldots$ into the original problem. The macro-structural description results from setting $\varepsilon = 0$ in the formulation of the perturbation problem; D_{mic}, with $\varepsilon > 0$, is a perturbation of D_{mac} with $\varepsilon = 0$. This is a natural indication that the (instances of) properties referred to by D_{mac} are contained in those referred to by D_{mic}.[11] D_{mic} is 'richer' because it contains the 'stuff' introduced by setting $\varepsilon > 0$; but it is precisely this additional 'stuff' that is not required for the connection between the macro-structural distributions of conductivity and temperature (k_{eff} and $T_0(\xi)$), described by D_{mac}. Another way of looking at the comparison of the macro- and micro-structural descriptions is to note that the properties referred to by D_{mic} ($\varepsilon > 0$) could be said to 'realize' the properties referred to by D_{mac} ($\varepsilon = 0$). Without suggesting a detailed explication of the realization relation in terms of perturbed and unperturbed systems, I just want to point out that an instantiation of a (macro-structural) property of the $\varepsilon = 0$-system, say a certain k_{eff} distribution, is *necessitated* by, as well as *explained* by the instantiation of a (micro-structural) property of the $\varepsilon > 0$-system, i.e., a certain $k(x)$ profile.[12] Many different systems, characterized by different values of ε or different factors of ε, can be understood as perturbations of the same $\varepsilon = 0$-system. Thus we have multiple realizability.

John Heil (1999) has criticized such attempts to understand the realization relation in terms of the part-whole relation for mental and physical properties because he thinks these attempts lead to a "swallowing up" of the mental by the physical:

> "Suppose that an agent, *a*, possesses N_1 [a physical realizer of the mental property *M*]. In what sense does *a also* possess *M*? The causal powers of *M* are described as a subset of the causal powers of N_1. But it is not as though the presence in *a* of N_1 includes two clusters of causal powers: those bestowed by N_1 and those bestowed by *M*. The supposed presence of *M* in *a* appears to be entirely absorbed by the presence of N_1 ... [I]t is hard to see what more there is to *a*'s possessing *M* beyond *a*'s possessing N_1. [...] The worry here is that *M* is swallowed up by its realizers." (Heil 1999, 194)

Surely, by possessing N_1, or in our case, a k(x) or T(x) distribution, the system necessarily also possesses M or distributions of k_{eff} and $T_0(\xi)$. But there is a non-arbitrary way of demarcating, say, $T_0(\xi)$ from T(x) — in terms of unperturbed part of a system and the perturbed system itself with their associated, qualitatively different behaviours. The fate of the realized property, the macro temperature distribution, of being "swallowed up" by the realizing property, a micro temperature profile, is avoided by demonstrating that the former is not reducible to the latter.[13] Nonreducibility in our sense implies that certain 'patterns' in the distribution of qualities of a system can be 'seen' only in the macro-structural description and are lost in micro-structural descriptions. This is the sense in which the properties realized maintain a degree of autonomy from their realizers.

There is, then, a distinction to be made between physical properties that are sufficient for a given effect property and physical properties that are causally responsible for the effect. Such a distinction appears to violate Crane's requirement of homogeneity of the causal relation — a requirement he regards as central for any sort of nonreductive physicalism. Since we find that this distinction is already needed in physics itself in order to deal with causally efficacious and irreducible macro-structural properties, it should obviously not be taken to threaten the very motivation of physicalism. Still, nothing I have argued should be taken as *positive* support for the claim that macro-structural properties are real. All I have tried to show is that the problem of causal exclusion should not be regarded as a problem that casts doubts on the reality of such properties. In other words, *if* these properties are real then they are able to perform their own sort of causal work, work that is not preempted by the jobs done by the associated micro-structural properties.

5. Emergent Properties

A view that takes macro-structural properties to be (often) irreducible to more basic micro-structural properties has to explain how it relates to the doctrine of emergentism. Emergentism, roughly, claims that given certain complexes of basic properties, other properties with 'novel' causal powers come into existence. If the basic level is assumed to be the physical level, this doctrine has often been seen as in conflict with any form of physicalism because the existence of irreducible, non-physical properties with their own distinctive causal powers would disrupt the causal closure of the physical, one of the fundamental tenets of physicalism.

Kim has argued that most varieties of nonreductive physicalism are, in fact, indistinguishable from emergentism and are therefore incoherent (e.g., Kim 1992). He summarizes the emergentist doctrine in three theses:

> (1) "[Ultimate Physicalist Ontology] There are basic, nonemergent entities and properties, and these are material entities and their fundamental physical properties."
>
> (2) "[Property Emergence] When aggregates of basic entities attain a certain level of structural complexity ..., genuinely novel properties emerge to characterize these structured aggregates."
>
> (3) "[The Irreducibility of Emergents] Emergent properties are 'novel' in that they are not reductively explainable in terms of the conditions out of which they emerge." (Kim 1992, 122-24)

Claims (1) and (3) are obviously shared by nonreducive physicalists and emergentists. Claim (2) is closely related to the physicalist's view that higher level properties (in particular, mental properties) have to be realized in lower level physical properties. The conjunction of the theses — whether in emergentist or physicalist formulation — , together with the principle of causal exclusion, inevitably leads to the problem of causal competition (see also Kim 1999).

One usually thinks (cf. thesis (2)) of emergent properties as properties possessed by a 'whole' or configuration of parts but not by any of its parts alone or in other configurations than this whole. But Kim has reminded us that this is a trivially satisfied characterization of properties of wholes and parts; no distinctive feature of emergence can be attached to it because the most trivial aggregative properties of objects would fall under this definition. Recall the notion of a micro-based or micro-structural property: P is a *structural* (higher-level) property iff P is the property of having proper parts a_1, a_2, ..., a_n, such that there are lower-level properties $P_1(a_1)$, $P_2(a_2)$, ..., $P_n(a_n)$ and the parts stand in relation $R(a_1, a_2, ..., a_n)$ (Kim 1998, 84). Such a property P is 'new' in the sense that none of the properties of the parts, P_i ($i = 1, ..., n$), is identical with it; P is characteristic of the configuration of the parts a_i because it depends on the specific relation R in which the parts stand. Furthermore, P can obviously be equipped with causal powers which are different from the causal powers of any of the P_i . Putting three 3kg pieces of metal together results in a configuration with the 'new' property of weighing 9kg with different causal powers than the individual pieces had.

Emergentists do not tolerate cases like this as examples of emergent properties. Usually such a micro-based property would be labelled a '(merely) resultant' property. How then could we distinguish emergent from resultant properties? One way is to spell out the 'novelty' of emergent properties as their being unpredictable, unexplainable, or irreducible in some sense: a description of the base properties and their configuration will not be sufficient to allow us to predict the new emergent properties, or explain how they resulted from their bases, or reduce them to these bases. I have elsewhere tried to explicate, within the framework of dynamical systems theory, a physicalistically respectable notion of emergence that preserves the central idea that emergent properties should be irreducible to their base properties and 'novel' with respect to them (Rueger 2000a,b). Irreducibility is understood in the way it is used above, as breakdown of uniform limit relations between descriptions of a system's behaviour (section III). 'Novelty' of a system's behaviour, compared to, say, the behaviour of its subsystems is characterized as *qualitatively* different behaviour; the inequivalence of topological features of the system's and the subsystems' phase space portraits captures the sense of 'qualitative' required here.

The point of this explication is that it allows us to meaningfully distinguish between emergent and non-emergent ('merely resultant') properties *within* the category of structural properties of a system; emergent properties in this sense do not stand in contrast to structural properties of an assembly of parts but they are still different from 'merely resultant' properties of the assembly because they are 'novel' in the sense outlined. 'Weakly emergent' seems an appropriate label for this class of properties[14] in order to distinguish them from a putative sort of 'strongly' emergent properties, a sort that many traditional emergentists have tried to characterize as properties which "introduce ... novel causal influence over the behavior of the objects bearing them" and which have to be "distinct from any structural property of the object." (O'Connor 1994, 94f.)

It should be plausible, even without going into further details of how to define weak emergence (Rueger 2000a), that the relation of macro- and micro-structural properties analyzed above can be the relation of (weakly) emergent properties to their base properties if the conditions of irreducibility and novelty are satisfied. In our case, the macro temperature distribution is qualitatively different from the micro temperature profile and we can say that the macro distribution is a (weakly) emergent feature of the system, a feature which arises in the limit $\varepsilon \to 0$.[15] The problem of causal exclusion for weakly emergent properties is avoided in the way outlined above: these properties, or rather their instances, are contained in the instances of base properties possessed by the whole. Thus the distribution of $T_0(\xi)$ over the heated rod, which corresponds to the $\varepsilon = 0$ case, is *contained in* the distribution of T(x) over the rod, corresponding to $\varepsilon > 0$: the smooth macro distribution is nothing but a pattern visible in the rapidly changing micro distribution. In terms of the causal powers characterizing the properties in question, the powers associated with the macro distribution form a subset of the powers defining the micro distribution. The causal powers of the emergent property, therefore, are not new but nevertheless *different* from the causal powers characterizing the micro temperature profile. This difference in causal powers accounts for the irreducibility and novelty of the properties with respect to each other.

If many (although not necessarily all) macro-structural properties turn out to be (weakly) emergent properties on this view, one might worry that the notion of emergence is trivialized: there are just too many emergent properties. On the other hand, proposals for ‘strong’ emergence usually suffer from being unable to provide convincing examples for their doctrine. Whatever side our sympathies in this dilemma may tend to, the problem that Kim diagnosed for nonreductive physicalism on the grounds that any such position would be committed either to epiphenomenalism or to downward causation and thus to a violation of the causal closure of the physical, is not really a threat after all.

Acknowledgements

For helpful criticisms of earlier versions I am grateful to participants of the Inter-University Conference in Philosophy of Science at Dubrovnik (April 2002), especially Stephan Hartmann, as well as to Li Li, Patrick McGivern, and John Symons.

References

Antony, L.:1999, 'Multiple Realizability, Projectibility, and the Reality of Mental Properties', *Philosophical Topics* **26**, 1-24

Batterman, R.W.:2000, 'Multiple Realizability and Universality', *British Journal for Philosophy of Science* **51**, 115-145

Batterman, R.W.: 1995, 'Theories Between Theories', *Synthese* **103**, 171-201

Bedau, M.A.: 1997, 'Weak Emergence', *Philosophical Perspectives* **11**, 375-399

Berry, M.:1994, 'Asympotics, Singularities and the Reduction of Theories', *Logic, Methodology and Philosophy of Science,* Vol. 9. **, 597-607

Clapp, L.: 2001, 'Disjunctive Properties: Multiple Realizations', *Journal of Philosophy* **98**, 111-36

Crane, T.: 1995, 'Mental Causation', *Proceedings of the Aristotelian Society*. Suppl. **69**, 211-236

Dretske, F.: 1993, 'Mental Events as Structuring Causes of Behaviour', in: J. Heil and

A.Mele (eds), *Mental Causation,* Oxford University Press, Oxford, 121-136

Heil, J.: 1999, 'Multiple Realizability', *American Philosophical Quarterly* **36**, 189-208

Hinch, E.J.: 1990, *Perturbation Methods,* Cambridge University Press, Cambridge

Kevorkian, J. and J. D. Cole: 1996, *Multiple Scale and Singular Perturbation Methods,* Springer, New York

Kim, J.:1989, 'Mechanism, Purpose, and Explanatory Exclusion', *Philosophical Perspectives* **3**, 77-108

Kim, J.: 1992, ''Downward Causation' in Emergentism and Nonreductive Physicalism', in A. Beckermann et al. (eds), *Emergence or Reduction?* DeGruyter, Berlin, 119-138

Kim, J.: 1997, 'Does the Problem of Mental Causation Generalize?' *Proceedings of the Aristotelian Society* **97**, 281-297

Kim, J.: 1998, *Mind in a Physical World,* MIT Press, Cambridge (Mass.)

Kim, J.: 1999, 'Making Sense of Emergence', *Philosophical Studies* **95**, 3-36

LePore, E. and B. Loewer: 1987, 'Mind Matters', *Journal of Philosophy* **86**, 630-642

Nickles, T.: 1973, 'Two Concepts of Intertheoretic Reduction', *Journal of Philosophy* **70**, 181-201

O'Connor, T.: 1994, 'Emergent Properties', *American Philosophical Quarterly* **31**, 91-104

Putnam, H.: 1973, 'Philosophy and our Mental Life', in H. Putnam, *Mind, Language and Reality,* Cambridge University Press, Cambridge, 291-303

Rueger, A.: 2000a, 'Physical Emergence, Diachronic and Synchronic', *Synthese* **124**, 297-322

Rueger, A.: 2000b, ‘Robust Supervenience and Emergence’, *Philosophy of Science* **67**, 466-489

Rueger, A.: 2001, ‘Explanations at Multiple Levels’, *Minds and Machines* **11***,* 503-520

Shoemaker, S.: 1999, ‘Realization and Mental Causation’, *Proceedings of the 20th World Congress in Philosophy,* Vol. 9, 23-33

Teller, P.: 1995, *An Interpretive Introduction to Quantum Field Theory,* Princeton University Press, Princeton

Yablo, S.: 1992, ‘Cause and Essence’, *Synthese* **93**, 403-49

Yablo, S.: 1997, ‘Wide Causation’, *Philosophical Perspectives* **11**, 251-81

Endnotes

---

1. Quoted after Kim 1997, 288.

2. See, e.g., Kim 1989 for motivation and discussion of the principle.

3. If you find it implausible to think of the conductivity k as the 'cause' of the temperature distribution T, imagine the problem (Eq. 1) with a heat source included:

$$(d/dx)\ [k(x)dT(x)/dx] = Q(x).$$

Whatever I say about k as the cause could then be rephrased, perhaps more intuitively, in terms of the source strength Q.

4. Kim does mention this question at one point in the context of "the doctrine of property emergence"; he announces a program for further work: "we need to deal with the question whether some micro-based properties supervene on other, more basic, micro-based properties,and the attendant question of their respective causal powers." (1997, 297 fn.16)

5. Thus the the microscopic conductivity k(x) and the macroscopic quantity k_{eff} are not just related to each other like a sequence of numbers and its arithmetic mean — one of the features

---

that indicate that the transition from a micro- to a macro-structural description is by no means trivial.

6. See Antony 1999 for considerations along similar lines.

7. Excluding, as before, systematic overdetermination of effects by their causes.

8. Yablo actually formulates these definitions not in terms of parts and wholes of property instances but in terms of 'determinables' and their 'determinates'. But the intended meaning is that a property instantiation Y necessitates the instantiation of a property X "because X is immanent in or included in Y. This is all it takes to kill the appearance of causal competition." (Yablo 1997, 275, n. 22)

9. On this view of the relation between mental and physical properties, see also Antony 1999 and especially Shoemaker 1999 and Clapp 2001.

10. Thus we preserve the causal completeness of the micro-structural properties (described in D_{mic}) by allowing these properties to be sufficient for their effects.

11. I am aware of the difficulties of interpreting the terms in a perturbation expansion as parts of a whole (cf., e.g., Teller 1995, 139-141). Nevertheless, I think the considerations adduced here

---

and below give some plausibility to my claim.

12. For the explanation issue see Rueger 2001; for multiple realizability Batterman 2000.

13. Suppose then that N_1 is "constituted by" M and some other property N_1*. Heil continues to worry: "...although M is not reducible to N_1, M is reducible to (in the sense of identifiable with) N_1', a material constituent of N_1 [where $N_1' = N_1 - N_1^*$] ... The temptation to regard M as a higher-level property vanishes" because M and N_1', if they are identical, are at the same level (Heil 1999, 195). In our case, however, the properties corresponding to Heil's M and N_1 are at the same level to start with.

14. For weak emergence see also Bedau 1997, 394f.; Bedau's definition of emergence, however, differs from mine.

15. That a property of a system is 'novel' or qualitatively different is obviously a relational claim which requires specification of a 'reference system' or property with which the new property can be compared. Thus choosing the micro temperature distribution as the reference or base property, we can say that the macro temperature profile is a novel feature of the system. Had we chosen the macro profile as our reference, the micro profile of the system would be classified as novel. See Rueger (2000 a, b).