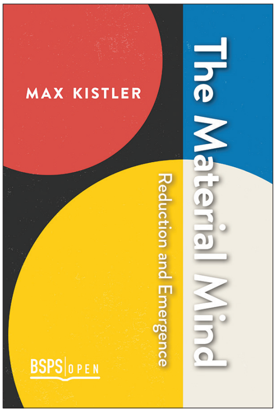
UNIVERSITY OF CALGARY
Press



THE MATERIAL MIND: REDUCTION AND EMERGENCE

Max Kistler

ISBN 978-1-77385-607-0

UNIVERSITY OF CALGARY
Press
press.ucalgary.ca

Acknowledgement: *We acknowledge the wording around open access used by Australian publisher,* **re.press**, *and thank them for giving us permission to adapt their wording to our policy* http://www.re-press.org

4

Emergent Properties

1. Introduction

The aim of this chapter is to justify the causal efficacy of mental properties by presenting the hypothesis that these properties are emergent. A concept of emergence appropriate in this context must satisfy the following requirements: it must apply to certain psychological properties; it must be compatible with contemporary science; it must respect the methodological requirement not to prejudge scientific discoveries, in particular with regard to the reduction of mental properties to neurophysiological properties; and it must justify the thesis that emergent properties are causally efficacious and not merely epiphenomenal. The guiding idea is that emergence characterizes certain properties of complex objects qualitatively different from the properties of the parts of these objects and that the concept of emergence applies to properties of structured objects at all levels: the emergence of mental properties from neurophysiological properties of the brain belongs to the same category of relation as the emergence of the chemical properties of molecules from the physical properties of the atoms that make them up.

Emergent properties belong only to complex objects. There are many examples of intuitively emergent properties, including the macroscopic properties of water and ice, emergent with respect to the microscopic properties of H and O atoms; the disposition of hemoglobin molecules to bind and discharge oxygen, emergent with respect to the properties of the atomic components of these molecules; the disposition of living organisms to transmit their hereditary characteristics, emergent in relation to the properties of certain parts of their bodies and in particular in relation to the properties of DNA molecules and the genes that they contain; and, finally, the collective behaviour of a colony of ants, emergent in relation to the behaviour of the individuals that it contains (see Holland 1998; Johnson 2001).

A condition traditionally imposed on emergent properties is that of being *new*: the property E of the structured object s is emergent only if none of its parts $p_1 \ldots p_n$ has E. However, we will see that it is not easy to determine a sufficiently restrictive sense of what "new" means in this context to prevent the result that *all* properties of s count as emergent. I will propose the following criterion: a property is new if it possesses new *causal powers* that correspond to new laws.

The need to be compatible with science and not to prejudge reducibility means that we have to abandon the predominant criterion in the traditional conception of emergence: the *irreducibility* of emergent properties. Within the framework of the conception of emergence that I am going to develop, emergence is compatible with reducibility: reduction requires the discovery of a set of laws that guarantees the existence of an instance of the reduced property each time the set of reducing properties is present. This is compatible with the emergence of the reduced property: the properties of the parts of the complex object s, together with the relations existing between these parts, give rise, according to what I will call a "law of composition,"[1] to emergent properties of s. These properties are emergent in the sense that they give the whole causal powers that its parts do not have. In fact, the nomological origin of emergent properties guarantees their reducibility.

1 I am moving away from the terminology of Cummins (1983), which has the merit of insisting on the need to recognize non-causal forms of explanation. According to the terminology proposed by Cummins, a "law of composition" determines the analysis of a system: it is therefore the type of system, a property of the whole, that determines the type of parts. I have two objections to this interpretation. The first objection is that it invites confusion between the direction of investigation and the direction of objective determination: first we know the property of the whole; then we direct our investigations toward its components and the mode of their composition. But the concept of law requires that a law determine the objective relationships between properties, independently of epistemic priority and the direction of our investigation. If we accept that the properties of the parts determine those of the whole, then it is consistent with the ontological (i.e., realist) interpretation of laws to conceive of the laws of composition as the ontological basis of this determination: the laws of composition thus determine the whole on the basis of the parts, in a "bottom-up" direction, whereas Cummins conceives of the meaning of determination as being "top-down." The second objection is that Cummins excludes by definition the nomological analysis of multi-realizable global properties. There is no single "law of composition" (in Cummins's sense) of the property of being an eye. However, there are "laws of composition" in the sense proposed here for each type of eye.

This conception will allow us to make a qualified judgment of the possibility of *explaining* emergent properties. Reduction, always possible in principle, even if its complexity might put it practically beyond our reach, provides a scientific explanation of an emergent property. In this sense, we must reject the traditional claim that emergent properties are "inexplicable." However, it is possible to salvage the intuition that motivates that claim by applying the judgment of inexplicability, at least in part, to the laws of composition.

The new qualities of mental properties, in relation to the underlying neurophysiological properties, retain some of their intuitively mysterious character because the laws of composition that give rise to the existence of mental properties are a posteriori, and although they are necessary, like all laws (see above and Kistler 2002a, 2005a), it is *conceivable* that they are different from what they actually are. Even when we have discovered by virtue of which laws a phenomenon regularly occurs in specific circumstances, and when we have in this sense explained the phenomenon, nevertheless we can justify the intuition that we still have not *fully* explained the phenomenon, insofar as we have not explained the laws themselves that explain the phenomenon. The explanation of these laws has a limit in the axioms of the best theory,[2] which are always a posteriori and conceivably different from what they are in reality. However, the mystery of the quality of mental properties is no greater *in principle* than the mystery of the quality of chemical properties. If it is greater for us today, then it is only because we are further away from the discovery of the laws of composition that give rise to mental properties, based on neurophysiological properties, than we are to the discovery of the laws of composition that give rise to the chemical properties of molecules, based on the properties of their atomic components.

2. Minimal Conditions and Weak Emergence

The concept of emergence is supposed to do justice to a twofold conviction. The first is that emergent properties, particularly cognitive properties, are real but distinct from the neuronal properties underlying them. According to the second, they are determined exclusively by the underlying material

2 According to the "best system account" of the laws of nature, due to Mill and Ramsey and more recently developed by David Lewis, "a contingent generalization is a *law of nature* if and only if it appears as a theorem (or axiom) in each of the true deductive systems that achieves a best combination of simplicity and strength" (1973, 73).

properties; in particular, cognitive properties are determined by the properties of the brain of the subject as well as those of the rest of her body and the environment with which she interacts. This conviction amounts to giving cognitive properties an intermediate status: they are distinct from the properties belonging to the parts of the subject, in particular the brain and the parts of the brain, as well as from those properties of the subject of the same quality as the properties belonging to its parts, such as her mass or volume. However, they are not *independent of* her brain and, more generally, of her body and environment. To say that the perceptual state of a person looking at a red tomato is an emergent property is to say that only a cognitive system obeying precise architectural constraints can possess it. Yet nothing simpler can have this experience; in particular, none of the parts of the system can have it. For example, the visual system alone, isolated from the rest of the brain and body, cannot have such experiences. This is what we mean when we say that the perceptual state is not a "resultant" property: a property of an object is resultant if it is not qualitatively different from the properties of its parts. Conversely, however, the perceptual state is determined exclusively by physical and physiological conditions of the brain, alongside the rest of the body and its particular environment. In this sense, emergentism is opposed to dualism: an emergent property is different from the properties of the substrate that brings it into existence, but it is this substrate and its physical environment that exclusively determine its nature.

The task of characterizing emergence can therefore be broken down into two parts. The first and easier part consists of distinguishing the emergent properties of an object from hypothetical "dualist" properties, whose existence and nature would be determined by something other than the material properties of the object's parts and their interactions. We can establish this distinction by means of three constraints that a property must satisfy in order to be emergent. In other words, the existence of a property that violates one of these constraints can be accommodated only within a dualistic framework.[3] Conversely, the fact that a property satisfies these constraints is not enough to make it emergent in the strict sense: these are necessary but not sufficient conditions. They are also satisfied by resulting properties and in particular, in the case of properties of people, by their purely physical and physiological

3 The satisfaction of these constraints corresponds to what Stephan calls "the weak theory of emergence" (1999, 66–67).

properties. The more difficult task will be to distinguish emergent properties from these resultant properties. The three constraints are as follows.

First, an "anti-dualist" condition that emergent properties must satisfy is that they belong only to physical systems (i.e., systems composed exclusively of physical parts). Everything that is emergent — whether property, process, or structure — belongs to a system made up exclusively of physical parts. This condition makes emergence compatible with physicalism.

Second, a property *P* of a complex system can be called "emergent" only if it does not also belong to parts of the system. Having a certain mass is a property that belongs to a person as a whole, but it also belongs to its parts. This is enough to justify the intuition that the property of having a mass does not "emerge" at the level of the cognitive system, because it is already present at the lower levels of complexity to which its parts belong. We can express this condition by saying that emergent properties are *systemic* properties of an object, which means that no part of the object possesses them (Stephan 2006).

Emergent properties are "anti-homeomerous." To take Armstrong's definition, "a property is homeomerous if and only if for all particulars, *x*, which have that property, then for all parts *y* of *x*, *y* also has that property" (1978, 2: 68). Emergent properties are anti-homeomerous in the sense that a necessary condition of emergence is that, for all *x* that possess the property, *none* of *x*'s parts *y* possesses it.

Third, the two previous constraints are not strong enough: if a cognitive system possessed a "spiritual" property imposed on it by God, either directly (as is the case according to occasionalism) or indirectly (as is the case according to parallelism), rather than through the material substrate of the system and its organization, then it would not be "emergent" from this material substrate. It would not be emergent even if God imposed the property in question only on systems with physical parts (so that the property satisfies the first condition) and if he imposed it only on the whole system but not on any of its parts (so that it satisfies the second condition). An emergent property must be determined exclusively by the parts making up the system: an emergent property is *determined synchronously* by the properties of the parts of the system and their organization. This determination requires *mereological supervenience.*[4] The emergent properties of a system *supervene* on the properties of its

4 See Kim (1984, 165; 1988b). I return to the concept of mereological supervenience in Chapter 5.

parts: there can be no change in the supervening properties without a change in the properties of the parts of the system.[5]

It is important to distinguish the *synchronic* determination of an emergent property from its *causal* determination. The metaphysical analysis of causality is controversial. Insofar as we conceive of cause and effect as particular events that occupy a delimited portion of space-time, cause and effect must be spatiotemporally distinct.[6] When we speak of a *property F* that causes another *G*, *F* must be the property of a cause event *c*, and *G* must be the property of an effect event *e*, where *c* and *e* are spatiotemporally distinct. In terms of causality between events, it is excluded conceptually that the cause is the same event as the effect; it is also excluded that cause and effect partially overlap. Causality can be said to exist only if the cause is spatiotemporally located elsewhere than the effect. On the basis of these conceptual constraints on causality, the emergence of *A* from *B* is incompatible with the existence of a *causal* dependence of *A* on *B*. According to the third condition, emergence is a synchronic relation. *A* is an emergent property of system *s* if *A* is determined by the properties *B* that the system possesses *at the same time*. Furthermore, *A* belongs to the whole system *s*, and the properties *B* belong to the parts p_i of the same system: the parts p_i possess *B* at the same time as the system *s* possesses *A*. Accordingly, there is partial spatiotemporal overlap between the carriers p_i of the properties *B* and the carrier *s* of the emergent property *A*; the emergence of *A* from *B* is a form of non-causal determination.

In such a conceptual framework, it is difficult to interpret Searle's assertion that "consciousness is a causally emergent property of systems" (1992,

---

5 However, synchronic determination goes beyond mereological supervenience: the latter is compatible with the possibility that systemic properties are determined by God rather than by the properties of the parts of the system (Kim 1993a). I will develop this point later in this chapter.

6 I cannot justify this thesis here. See Kistler (1999b, 2006d). Hume (1978, 76) justifies it as follows: if simultaneous causality were possible, then there could be no non-simultaneous causality. A cause that is sufficient for its effect could not precede it by a finite time because, since it is sufficient, no additional condition intervenes at the moment when the effect occurs. But in this case there is no sufficient reason for the effect to occur just then and not earlier or later. A different and less direct way of refuting the possibility of simultaneous causality is to argue that causes and effects are events: that is, particular entities that occupy particular areas of space-time. In this framework, simultaneous causality is impossible because the effect can occupy neither the same space-time zone as the effect nor a different one: the first case is impossible because the effect must be distinct from the cause; the second is impossible because the physics of relativity forbids simultaneous action at a distance.

112). Searle holds that all mental states and processes are at least potentially conscious, which leads him to defend more generally the thesis of the "causal emergence" of all mental phenomena. They are, he says, "caused by neurobiological processes" (1992, 1, 115, and passim).

Searle's thesis undoubtedly has its source in the traditional but erroneous identification of nomological determination with causal determination. Not every determination according to laws of nature is causal (see Humphreys 1989, 300–01; Salmon 1990, 46–50; Kistler 1999b, 2004a, 2006d). Only the neglect of this distinction can explain why Searle moves without justification from the claim that emergent properties are determined by causal interactions between the parts of a system to the claim that this determination itself is causal: emergent properties, he says, "have to be explained in terms of the causal interactions among the elements" of the system that possesses them, which he takes to justify calling them "causally emergent system features" (1992, 111). However, the fact that emergent properties are *synchronous* with the properties of the parts of the system that determine them forces us to conclude that this determination follows non-causal laws of determination. Emergent properties, therefore, are *not caused* by the properties that determine them through laws of composition.

In the same vein, E.J. Lowe justifies his assertion that mental properties have causal powers "not wholly grounded in . . . the causal powers of those elements of the system which produced [them]" (1993, 636–37) by comparing consciousness to a spider's web. Unlike the liquidity and transparency of water, causal powers based entirely on the powers of water's constituent molecules, the spider's web has "a life of its own" (636), an expression that Lowe borrows from Searle, who says that "consciousness gets squirted out by the behavior of the neurons in the brain, but once it has been squirted out, it then has a life of its own" (1992, 112). This analogy clearly reveals the confusion shared by Searle and Lowe between causal determination and non-causal nomological determination.[7] The spider's organs produce the web in a process spread out over time: it begins with events that take place in the spider and ends in the existence of the web. This means that the events that cause the web take place earlier than the events in which the web exercises its causal powers, for example, when it supports the spider passing across it.

7 I develop this criticism of Lowe's conception of mental properties in Kistler (2005b, 2022).

Along these lines, the analogy between the web and mental properties, in their respective relationships with the spider and the brain, is misleading: the subject possesses a mental property by virtue of a synchronic determination, which gives mental properties particular powers, but no independent existence in a substantial sense. Conversely, once created, the web no longer depends on the spider; it can continue to exist even if the spider disappears, whereas cognitive properties are permanently dependent on the underlying brain processes and events that give rise to their existence. As soon as brain activity ceases, the mental property ceases to exist. Synchronous nomological determination is not peculiar to the relationship between brain and mind. This is how, for example, the properties and processes taking place at the level of the electrons in a metal determine, in a non-causal way, the electrical and thermal conductivity of the metal. Without the microscopic processes involving the electrons, there is no electrical conductivity. This is also a case of non-causal determination, because the metal has the conductivity and the underlying microscopic properties at the same time.[8]

According to the concept of emergence defended here, emergence does not exclude reduction. However, an emergent property can only be reduced to its basis provided that we do not conceive of reduction as the discovery of the identity of the reduced property with its reduction basis. Insofar as one subscribes to the thesis[9] that reduction leads to the discovery of an identity, the compatibility of emergence with reduction can lead to strange conclusions. Searle, presupposing his thesis that emergence is a causal process, concludes that, once their reduction is complete, emergent properties "can . . . be identified with their causes" (1992, 115). This conclusion destroys the intelligibility of both emergence and causation. They would be reflexive relations, where things can cause themselves and emerge from themselves. It is possible to avoid such absurd consequences if we accept that neither the reduction of *A* to *B* nor the causality between *A* and *B* entails the identity of *A* and *B* and if we conceive of emergence as a relation of non-causal determination of the systemic properties of a complex object from the properties of its parts.

8 If we accept the idea that a "temporal part" of an object is an event, then these are properties of the same event. This appears to be a natural consequence of the conception according to which the content of a well-defined spatiotemporal zone always constitutes an event, even if it corresponds to no apparent change. See Kistler (1999b, 2006d).

9 It is defended in particular by Causey (1977). See Chapter 1.

3. Broad and the Epistemic Conception of Emergence

Properties that obey the three conditions formulated above can be said to be emergent in a weak sense. These are properties that belong only to exclusively physical objects, and they are determined exclusively by the properties of their parts. The weakness of this first concept of emergence lies in the constraint of "novelty." For example, without further clarification of what counts as "new" or "qualitatively different" in relation to the properties possessed by the parts of an object with emergent properties, the distinction between the "resultant" property of a human being of weighing 70 kg and her "emergent" property of hoping that the war will end soon has no rigorous basis. The general property of being massive is certainly not systemic because all of the parts of the body also possess it. However, weighing 70 kg is formally a systemic property because no proper part of the person's body possesses it. We have the intuition that this property is not "qualitatively different" or "qualitatively new": this intuition justifies the introduction of the distinction between strongly emergent properties that are not only systemic but also qualitatively new in relation to the properties of the parts and only weakly emergent properties, such as the property of weighing 70 kg.

The British emergentists at the end of the nineteenth century and the beginning of the twentieth century tried to make this distinction rigorous by analyzing it in terms of *explanation*: "resulting" properties are those whose presence can be explained on the basis of the properties of the parts, whereas an analogous explanation of the presence of emergent properties is possible only with the help of ad hoc postulates. However, this way of grounding the distinction between emergent and resultant properties is incompatible with physicalism. According to physicalism, every emergent property conforms to the third condition mentioned above: it is determined exclusively by the properties of the parts of its possessor. It is therefore sufficient to discover the laws underlying this nomological determination to explain (in the sense of deductive-nomological explanation) the presence of the emergent property. Within the framework of physicalism, all global properties can be explained in principle. Accordingly, this possibility does not give rise to any relevant distinction among such properties, between emergent and resultant.[10]

10 We cannot hope to construct a relevant ontological concept from the epistemic criterion according to which irreducible predicates and laws are emergent. However, the epistemic conception

The relevant laws that I call — in the words of the last great British emergentist C.D. Broad — "laws of composition" determine a property of a complex object on the basis of the properties of its parts and the interactions between these parts. Once the law of composition that determines a given emergent property is known, that property, and the laws in which it is involved, can be explained and reduced.

The physicalist framework — and in particular the third condition of the nomological determination of the properties of complex objects — thus imposes the possibility of emergent properties being subject in principle to reductive explanation. However, this constraint seems to clash with a central thesis of the emergentist tradition, according to which the distinguishing mark of emergent properties is the *impossibility* of explaining them completely. Broad (1925, 65) resolves the tension between the constraints of nomological determination and the impossibility of explanation by conceiving of emergent properties as nomological, in the sense that their presence is determined systematically by a law of nature. The presence of the emergent property can be predicted and even explained, therefore, at least in the minimal sense of deductive-nomological explanation. What cannot be explained in the case of an emergent property is the law that determines it itself. This law remains a "raw nomological fact."[11]

Broad gives the example of silver chloride. It is on the basis of the properties of its components, chlorine and silver, and their relationship as components of a molecule that has no other components, that a law determines the properties of silver chloride. But this law is an experimental law that is "brute" (Broad 1925, 55) in the sense that it cannot be derived within a more general theoretical framework. It can be discovered only experimentally by observing the properties of samples of this compound. Insofar as it is a purely experimental law in this sense, it is inexplicable. There are no answers to the questions why is silver chloride composed of equal parts of chlorine and silver, and why does it have such or such properties? The law of composition transfers its inexplicable, and in this sense mysterious, character to the

of emergence proposed by Hempel and Oppenheim (1948) is still widely accepted. Churchland expresses it by saying that "claims about the emergence of certain properties are therefore claims about the relative poverty in the resources of certain aspirant theories" (1985, 12). See also McLaughlin (1992). Malaterre (2010) shows that such an epistemic conception of emergence can be illuminating in analyzing the scope and limits of scientific explanation.

11 I take this expression from McLaughlin (1992, 81).

emergent properties whose existence it determines. As we will see, this thesis is responsible for the marginalization of emergentism in the philosophy of science. It can already be found in J.S. Mill, for example in his assertions that "it is impossible to deduce all chemical and physiological truths from the laws or properties of simple substances or elementary agents" and that "the laws of Life will never be deducible from the mere laws of the ingredients" (1843, Book III, Chapter VI, para. 2). Broad expresses it in these terms: "The characteristic behaviour of the whole *could* not, even in theory, be deduced from the most complete knowledge of the behaviour of its components" (1925, 59).[12] The properties of the compound AgCl, for example, emerge, according to this criterion, from the properties of its atomic components (Ag and Cl), insofar as the law of composition of AgCl is "a law which could have been discovered only by studying samples of silver-chloride itself, and which can be extended inductively only to other samples of the same substance" (Broad 1925, 65).

Broad proposes to ground the distinction between emergent and non-emergent properties on the distinction between two types of a posteriori laws, the criterion being their integration into a theory. Broad contrasts principles of composition (1925, 45, 66), which are explanatory, with "trans-ordinal laws" (77 ff.), which are "brute nomological facts" because they cannot in turn be explained on the basis of more general laws. Broad knows that the principles or laws of composition[13] are a posteriori, just as much as trans-ordinal laws, and that both types of law determine the properties of the whole entirely from the properties of the parts. According to Broad, the difference between the two types of law lies in the possibility of deducing them in turn

12 The rest of the quoted sentence contains a crucial qualification that makes it possible to reconcile emergentism with contemporary science and to which I will return. Broad anticipates my thesis that it is not possible to deduce the emergent properties of a complex object from complete knowledge of the properties of its parts, insofar as they are the properties that the parts possess *in isolation* or in other combinations. I say "anticipates" because he expresses this thesis in the language of causal determination, whereas it is a form of non-causal determination.

13 Broad seems to think that all laws of composition have an additive form but without explicitly affirming it (1925, 66). McLaughlin notes the existence of this "gap in the discussion" (1992, 77) but does not appreciate its importance. McLaughlin does not note that quantum mechanics refutes the emergentist thesis only if it is assumed that the laws used to deduce molecular properties from atomic properties are "compositional principles." Since these principles are not additive, the original conception leads us to consider this deduction as trans-ordinal. It is only because of the imprecision of what counts as a "compositional principle" in Broad that McLaughlin can conclude that the quantum deduction of the molecular state uses only compositional principles and thus succeeds in refuting the thesis that chemistry is emergent.

from more general laws. A trans-ordinal law allows predictions to be made and practical control to be exercised over the emergent properties that it determines. In these respects, it is no different from a law of composition. The distinctive property is that the trans-ordinal law that determines the global properties of a given sort of compound object is a primitive law that applies only to that type of object. It can be discovered only experimentally by examining samples of that particular compound (Broad 1925, 65).

This way of distinguishing between emergent and non-emergent properties, based on the distinction between primitive experimental laws and deducible composition laws, makes the concept of emergence an *epistemic* concept relative to a theory — in other words, relative to a given moment in the history of science. A property *E* is emergent only in relation to a given theory *T*. To say that *E* is emergent, relative to *T*, simply means that *T* does not have the conceptual resources to derive the law of composition that gives rise to *E*. As soon as a more powerful theory is constructed that has the means to produce this derivation, *E* ceases to be emergent. In other words, according to the conception of emergence developed by Mill and Broad, it seems to be reasonable to expect that no property is emergent in an absolute or ontological sense. For any property that appears to be emergent relative to the theories available at a given moment, it is only a matter of time before a theory is discovered that can explain the laws that give rise to its existence. Emergence is an effect of perspective: emergent properties are those whose scientific explanation is still unknown at a given moment.[14]

14 The epistemic conception of emergence was proposed by Henle (1942) and by Grelling in his correspondence with Hempel and Oppenheim. We could interpret Lewes's remark that we will "some day, perhaps, be able to express the unseen process [by which hydrogen and oxygen are transformed into water] in a mathematical formula; till then we must regard the water as an emergent" (1875, 414; cited in Stephan 1992, 28) as the expression of an epistemic concept of emergence. According to McLaughlin, Lewes does not maintain that the emergent character of a property depends on the progress of our knowledge. The passage quoted only expresses the idea, compatible with an ontological interpretation of emergence, that the evaluation of the hypothesis that a property is emergent depends on the progress of our knowledge. See Stephan (1992, 28n4). Similarly, Broad says that "within the physical realm it always remains logically possible that the appearance of emergent laws is due to our imperfect knowledge of microscopic structure or to our mathematical incompetence" (1925, 81). Broad does not express a relativist doctrine of emergence like that of Hempel and Oppenheim (1948) but merely the recognition that our hypotheses about the emergent or resultant nature of a given property are fallible, even though their truth conditions are objective and absolute.

In an epistemic or relativist conception of emergence, it could certainly be argued that, at a given moment, there are always laws that have the status of brute experimental laws, whose only justification comes from induction on the basis of observation or experimentation. For a trans-ordinal law that indicates the properties of a whole as a function of the properties of its parts, the former properties count as emergent but only in relation to this stage of scientific progress or in relation to the limited knowledge that characterizes it. According to the research strategy of seeking to integrate knowledge of isolated experimental laws within increasingly general theories, we can expect the law eventually to change its status in relation to Broad's classification. As soon as there is a theory that allows it to be deduced from more general principles, the experimental law *becomes* a law of composition. The properties of silver chloride that Broad (1925, 64) gives as an example are emergent only in relation to the scientific knowledge of 1925. Progress in chemistry has made it possible to deduce a certain number of properties of this compound from general laws that apply to all halogen and metal compounds and then from even more general laws that determine the properties of a compound from the structure of the electronic orbitals of its constituent atoms.

This observation does not settle the question, of course, of whether there are experimental laws that will definitively resist integration into a broader theory. But Broad's distinction seems to be capable of characterizing only properties emergent relative to a given level of scientific knowledge. Insofar as the strategy of scientific research is dominated by the "mechanistic" paradigm in Broad's sense, which seeks to deduce any law first discovered experimentally within a theory from more general laws, we can never infer from the "emergence" relative to a given stage of scientific knowledge to objective or ontological emergence, independent of the level of knowledge.[15]

Let us return once more to the crucial point, the possibility in principle of deducing the existence of a given property. The British emergentists are materialists who explicitly take a stand against dualist positions. For example, Broad explicitly rejects the doctrine of "substantial vitalism" defended by Driesch, who postulates the existence of entelechies to explain the phenomena

15 According to Hüttemann and Terzidis (2000), the properties that are emergent according to Broad's criterion are "anomalies." Insofar as scientists pursue a mechanistic strategy, they set out to forge a theory that transforms trans-ordinal laws into laws of composition and consequently deprives emergent properties of this status.

of life (see Broad 1925, 58, 69; McLaughlin 1992, 86). Therefore, phenomena that occur at levels of complexity higher than that of atomic physics are not independent of phenomena that occur at lower levels. To postulate substances specific to each level would be tantamount to considering that these phenomena are independent. This is exactly what Broad criticizes as "Theories of a Special Component" (1925, 60; see also 55–58), which explain the behaviour of complex objects on the basis of "the presence of a peculiar component which does not occur in anything that does not behave in this way" (55; see Beckermann 1992, 16). Science does not support the postulate of such substantial components, such as the *élan vital* or the entelechies of the vitalists. In this sense, classical emergentists accept the idea that complex phenomena are determined nomologically by the properties of lower levels:[16] trans-ordinal laws are laws. On the ontological level, they determine emergent properties; on the epistemological level, knowledge of a trans-ordinal law allows the deduction of the presence of an emergent property from the observation of the presence of the conditions that appear in the law's antecedent. Insofar as the Hamiltonian of a system composed of two hydrogen atoms determines nomologically the formation and properties of the molecular ion H_2^+, it is possible to deduce (using other laws and principles of quantum mechanics) the formation of the molecular ion H_2^+, even if this deduction involves a "trans-ordinal" law (because the form of the Hamiltonian has not yet been deduced from more general principles).[17]

There seems to be a contradiction, in the doctrine of Broad, between the postulate of the existence of trans-ordinal laws — which determine the existence of emergent properties — and the assertion that it is impossible, even in principle, to deduce the existence of these properties or to predict the situations in which they are exemplified. As Broad puts it,

> if the emergent theory of chemical compounds be true, a mathematical archangel, gifted with the further power of perceiving the microscopic structure of atoms as easily as we can perceive hay-stacks, could no more predict the behaviour of

16 Provided that we accept the thesis of the necessity of the laws that determine emergent and resultant properties, emergentism therefore leads to local strong emergence. I will come back to this later in this chapter.

17 I return to this example later in this chapter.

> silver or of chlorine or the properties of silver-chloride . . . than we can at present. And he could no more deduce the rest of the properties of a chemical element or compound from a selection of its properties than we can. (1925, 71)

But the contradiction is only apparent. The solution lies in the sentence omitted from the quotation above: it is not possible to deduce the properties of silver chloride "without having observed samples of those substances" (71). In other words, emergent properties can indeed be predicted and explained but only by means of trans-ordinal laws that are irreducible experimental laws. These laws are "brute nomological facts" that themselves cannot be explained.

We have already seen one reason why this criterion is not adequate to the ontological concept of emergence that we are seeking. The progressive construction of theories regularly leads to the integration of experimental laws into theories, which makes it possible to transform their status of "brute nomological facts" into theorems deducible from axioms or principles of a more general order.[18] A second reason is that there are usually several equivalent formulations of the same theory, which attribute the status of axioms to different statements.

McLaughlin (1992) proposes to interpret Broad's criterion ontologically: trans-ordinal laws are objectively and absolutely primitive in the sense that in principle they cannot be derived within a more general theory. This ontological interpretation runs the risk of rendering the concept of emergence empirically empty, whereby it would remain as a coherent concept but void of any positive reason to think that it has application.[19] According to McLaughlin's interpretation, the emergentist doctrine bases the distinction between compositional principles and trans-ordinal laws on the distinction between particle pair forces and *configurational* forces. According to emergentists, each level of organization is characterized by specific properties that are causal powers and obey laws specific to their level of organization. According to some emergentists, in particular Sperry, these causal powers and the laws that characterize them are based on "configurational forces"

18 This is why Hempel and Oppenheim take Broad's thesis that emergent properties are those determined by a trans-ordinal law — that is, a law that cannot be deduced but must be induced directly from experience — to be "untenable" (1948, 262).

19 This is also the conclusion of Hüttemann and Terzidis (2000).

(1964; cited in Sperry 1986, 266):[20] that is, "*fundamental* forces that can be exerted only by certain types of configurations of particles, and not by any types of pairs of particles" (McLaughlin 1992, 52).[21] Now the conception of the distinction between compositional principles and trans-ordinal laws, which identifies the former with laws governing forces between pairs of particles and the latter with laws governing configurational forces, amounts to a refutation of emergentism by stipulation. For it is then enough to note that contemporary science, and above all quantum mechanics, give us every reason to believe that there are no configurational forces, to deduce that there are no emergent properties (McLaughlin 1992, 89–91).

Nagel (1952), Feigl (1958, 411–13), and McLaughlin (1992) note that contemporary science — notably quantum mechanics and the theory of evolution by natural selection — has come to account for and explain some of the properties that the emergentists Mill and Broad considered to be paradigms of emergent properties. They draw the conclusion that the success of these scientific theories removes all credibility from the hypothesis of the existence of absolutely emergent properties. This argument effectively refutes the versions of emergentism defended by Bain, Lewes, Alexander, Lloyd Morgan, and Broad as well as those defended later by Popper and Sperry. However, we will see that it is still possible to salvage the intuition that, in complex systems, new properties appear or "emerge" with causal powers different from the powers of the parts of these systems. In other words, it is possible to construct an ontological concept of emergence that takes account of scientific progress.

To explore the possibility of a concept of emergence compatible with contemporary science, we will take a closer look at an example that is crucial both systematically and historically: the case of chemical bonding. At a systematic level, the determination of the chemical properties of molecules from the physical properties of the atoms that make them up is a paradigm of the relationship of determination between adjacent levels of compositional complexity. We can then examine the hypothesis according to which the determination of biological properties from the chemical properties of the

20 On Sperry, see Stephan (1992, 43n20).

21 McLaughlin attributes this interpretation to classical emergentism and to Broad in particular: "Broad's Emergentism commits him to configurational forces" (1992, 88). It does not seem to me that the text justifies this interpretation. McLaughlin does not give a textual justification, as he is careful to do for his other assertions. But this point of exegesis is of secondary importance.

components of living bodies, and the determination of psychological properties from biological properties, fall on the same side of the great division between resultant and emergent properties as the determination of the chemical properties of molecules from the physical properties of their atomic components. In historical terms, it was the advent of quantum mechanics, and in particular the explanation of the origin of chemical bonding on the basis of physical laws, that tolled the bell for the great era of British emergentism. In McLaughlin's words, it was "no coincidence that the last major work in the British Emergentist tradition coincided with the advent of quantum mechanics. Quantum mechanics and the various scientific advances it made possible are arguably what led to British Emergentism's fall" (1992, 54). The quantum-mechanical explanation of elementary chemical properties, therefore, can serve as a test case for concepts of emergence. An acceptable conception of emergence must be compatible with contemporary science in general and with quantum mechanics in particular. If such a conception exists, then the verdict of Feigl, Nagel, and McLaughlin is premature: only certain historical versions of emergentism are obsolete, notably Broad's, but not emergentism *simpliciter*.

4. Strong Emergence in Terms of the Impossibility of Deduction

Insofar as all properties of complex systems are determined by laws of composition, it seems to be promising to try to ground the distinction between emergent and resultant properties in the nature of laws of composition. For the moment, we can rule out one interpretation of what a law of composition is. According to a thesis of logical empiricism, also at the origin of an important interpretation of reduction, they are *metalinguistic* principles whose role is to *define one predicate* in terms of others.[22] This doctrine is rooted in

22 I have analyzed this conception of reduction — considered by Nagel (1961, 354–57) — in Chapter 1. Wimsatt (1976a, 221) explains that the *many-many* relationship (see also Endicott 1998) between micro- and macroproperties in the reduction of Mendelian biology to molecular biology (see also Hull 1974) refutes the conception of reduction as "translation." Armstrong's (1968) and Kim's (1998) theories of functional reduction share the conclusion that mental and neurophysiological predicates are simply two sets of predicates that designate the same properties. The dissimilarity is that Armstrong and Kim take the difference between these two sorts of predicates to be a logical difference: second-order predicates contain a generalization over first-order predicates.

a more general conception of science as a set of statements; according to it, laws are theorems deducible from axioms or fundamental principles, which are also statements. Nomological statements include causal laws, which are a posteriori and express factual regularities. Laws of composition, in opposition, are taken to be (conventional) rules of translation. In an extensionalist interpretation of the language of science, a universal law of biconditional form expresses not the co-extensionality of *two* properties but the extensional equivalence of two predicates designating *a unique* property. In this sense, laws of composition are taken to be metalinguistic statements.[23] In this metalinguistic conception of reduction, a property is emergent if there is no rule for translating the predicate that designates it into a predicate constructed exclusively from terms of a reductive science. Hempel and Oppenheim (1948) point out that this doctrine renders theses concerning the emergence of a property trivial. The fact that we have not yet discovered the laws of composition that determine certain biological or psychological properties is merely a sign of the terminological incompleteness of present-day science. In their words, "in this interpretation, the emergent character of biological and psychological phenomena becomes trivial; for the description of various biological phenomena requires terms which are not contained in the vocabulary of present-day physics and chemistry" (263).

There are more general reasons against applying a metalinguistic interpretation to laws of composition. Let me briefly mention some reasons for adopting a realist conception of properties, according to which laws are constraints that these properties exert on each other. Predicates designate these properties, and nomological statements designate the laws (see Armstrong 1983; Kistler 1999b, 2006d). Nomological statements are a posteriori even though they designate necessary relations (Kistler 2002a). Such statements, therefore, are neither conventional nor metalinguistic, both in the case of causal laws and in the case of compositional laws. Laws of composition, in particular, *determine* the properties (emergent or resultant) of a complex object from the structural properties of its components. However, the existence of such a determination relation does not justify the idea that the properties

23 The linguistic interpretation of the distinction between emergent and non-emergent properties can be found in Pap (1951–52), Tully (1981), and Teller (1992). In Tully's words, "if secondary quality terms are indefinable in terms of microscopic particles, then this is a logical fact about the language we use to describe the world" (266; cited in Stephan 1992, 40).

of the system are identical to the properties of the components. The non-identity of the properties renders inadequate the conception according to which the statement that expresses their nomological correlation is metalinguistic, in the sense of expressing the definition of one predicate by another, where the two predicates designate the same property.

With this in mind, let us return to the traditional conception of emergence in terms of the impossibility of deduction. I will evaluate this conception by examining the question of whether hydrogen's property of forming stable molecules is an emergent property, using the following criterion:

> (C1) *Criterion of deducibility*: a global property *G* of a complex object *s* is *(strongly) emergent* if and only if (it is weakly emergent,[24] and) it is impossible in principle to deduce (i.e., explain, predict) the fact that *s* possesses *G*, from *complete information about* the components of *s* and *the properties possessed by the parts of s when they are isolated.*

We will see that, in the light of contemporary science, the applicability of the concept of emergence on the basis of the criterion of deducibility depends crucially on the interpretation of the expression "complete information."

The advent of quantum mechanics seemed to refute directly the thesis that the chemical properties of molecules are emergent according to (C1), with respect to the physical properties of the atoms that make them up: in quantum mechanics, certain simple but fundamental properties of molecules can be deduced from general principles and certain premises concerning the component atoms.

It might come as a surprise to read about the debate on the emergence of chemical properties from physical properties in an inquiry into the status of psychological properties in relation to neurophysiological properties. However, this issue is of fundamental importance to the latter problem insofar as the hypothesis according to which the mental emerges from the physiological will gain credibility by virtue of the justification of a more general hypothesis. According to this general hypothesis, emergence makes it possible to characterize the nature of the relationship between the properties of

24 Earlier I defined the weak emergence of a property *E* of a system *s* in terms of three conditions: *E* belongs only to objects composed solely of physical parts; *E* is systemic; *E* is determined entirely by compositional laws, based on the parts of *s* and their interactions.

a domain of objects *x* in relation to the properties of the objects of which the objects *x* are composed. Conversely, if this general hypothesis proves to be untenable in the fundamental case of the relationship between physical and chemical properties, then the hypothesis according to which psychological properties are emergent in relation to the properties of the neurons in the brain is likely to appear ad hoc.

To evaluate the thesis that quantum mechanics refutes emergentism with regard to chemical properties (according to the criterion of deducibility), I will examine the most fundamental and simple deduction of all explanations of chemical properties, based on physics: the deduction of the stability of the molecule H_2 or, more precisely, of the molecular ion H_2^+. We will then see that the viability of the emergentist thesis depends on the interpretation of what counts as "complete" knowledge of the properties of the components. If we understand knowledge of the parts of the complex object H_2^+ — the H atoms — to be absolutely complete, then this knowledge contains knowledge of the law according to which these atoms form, in precise circumstances, H_2 molecules. The global properties of these molecules, therefore, will not be emergent according to the criterion of deducibility. If, however, we consider that knowledge of the parts is complete as soon as it contains knowledge of all the laws that determine the evolution of the parts — the H atoms — *in isolation*, then even quantum mechanics does not allow us, on the basis of such limited knowledge, to deduce the formation and properties of the H_2 molecule. In that case, the criterion of deducibility leads to the result that the global properties of molecules are emergent after all.

It therefore appears that the anti-emergent consensus in contemporary analytical philosophy is built upon a strong interpretation of *completeness* in the definition of the distinction between emergent and resultant properties. If we consider that complete knowledge of the properties of components is not limited to the properties that these components possess in isolation, but extends to the laws that govern their interactions, then the refutation of the thesis that the stability of the H_2 molecule is an emergent property becomes trivial: if knowledge of the law of formation of H_2 is already part of the "complete" knowledge of H atoms, then by stipulation we have transformed the property of stability of H_2 into a property deducible from physics.[25] I will

25 Hempel and Oppenheim (1948, 260) credit Kurt Grelling for first noting this point in the course of their correspondence. They point out that Grelling was murdered by the Nazis before

therefore adopt a more restrictive interpretation of the completeness of knowledge of the properties of parts.

Quantum mechanics allows us to deduce that molecules exist from the existence of atoms and their properties. The paradigmatic deduction concerns the prediction of the existence, or more precisely the permanent existence or stability, of the hydrogen molecule H_2. The limits of the possibilities of exact calculation are quickly reached, but the conceptual importance lies in the possibility, in principle, of deducing a property that belongs to the chemical level of the organization of matter, from knowledge of objects belonging to a lower level of complexity (e.g., atoms). According to this fundamental result, quantum mechanics makes it possible to deduce that two hydrogen atoms, whose nuclei are close enough to allow a partial superposition of the space occupied by their respective electrons, form a stable H_2 molecule. It is advantageous to study the nature of this deduction in the case of the molecular ion H_2^+, the simplest system of the molecular level existing in nature: whereas the molecule H_2 has two electrons and two atomic nuclei, each consisting of a proton, the molecular ion H_2^+ has only one electron for two nuclei. It is therefore by studying the explanation of the existence of this molecular ion H_2^+ that we can hope to isolate the features of a deductive explanation that crosses the boundary between different levels of complexity.

The deduction of the existence of the hydrogen molecular ion is based on three nomological presuppositions. The first presupposition concerns the structure of the system's Hamiltonian. The Hamiltonian is an operator that determines, via the Schrödinger equation, the energy levels of the system. Its form is characteristic of the system under study. The system of the molecular ion H_2^+ consists of two protons and one electron. Its Hamiltonian $\mathcal{H}$ contains a term corresponding to the kinetic energy of the electron ($p^2/2m$) (where p represents the momentum and m the mass of the electron) and three terms corresponding to electrostatic interactions among the three objects making up the system: the repulsive interaction between the two protons, separated from each other by R (e^2/R, where e represents the reduced charge,

he could publish these ideas (245n1). Including all second-order nomological properties in what is required to know a property perfectly is less radical, however, than including knowledge of all properties of everything. As Van Cleve points out, we could trivialize the doctrine of emergence even further: "I could maintain that all properties of everything in the universe are deducible from the properties of James Van Cleve, provided you counted among my properties such items as 'being such that the Eiffel Tower is 1,056 feet tall'" (1990, 223).

with $e^2 = q^2/4\pi\varepsilon_0$), and the attractive interaction between the electron and each of the two protons from which it is separated by r_1 (e^2/r_1) and r_2 (e^2/r_2) respectively.[26]

$$(*)\ \mathcal{H} = p^2/2m - e^2/r_1 - e^2/r_2 + e^2/R$$

The second presupposition concerns the fundamental law of quantum mechanics, according to which the energy levels of a system are the solutions of Schrödinger's equation

$$\mathcal{H}\,\psi = E\psi\,,$$

where $\mathcal{H}$ denotes the Hamiltonian, E the energy, and ψ the wave function, which characterizes the state of the system.

The third presupposition concerns the general applicability of the quantum mechanical formalism to solve this equation in the case of the system under consideration.

The law that determines the Hamiltonian according to equation (*) is a law of composition. It determines a characteristic property of the system as a whole, as a function of certain properties of its components and their relationships. The properties of the components involved, on which the properties of the whole depend exclusively, are the electric charges of the electron and the protons and the mass of the electron, m.[27] The relationships that determine the state of the overall system are the respective distances among these three components: R, r_1, and r_2. The law of composition (*) is empirical or a posteriori, determining the impact on the energy of the overall system of the individual properties of the components and especially their mutual relationships. The law (*) determines the structure and properties of the system as a function of the interactions among its components, themselves

26 The equation (*) should be interpreted as describing quantum operators.

27 The mass of the protons is not involved since the movement of the protons is neglected. The mass of the protons is much greater than that of the electron, which means that the electron's motion is much faster than that of the protons. This is "why, to a first approximation, the two motions can be studied separately" (Cohen-Tannoudji, Diu, and Laloë 1977, 1: 511). This is the so-called Born-Oppenheimer approximation (see also Cohen-Tannoudji, Diu, and Laloë 1977, 2: 1160).

determined by the properties of the individual components and their mutual relationships (in this case, spatial distances).

In contrast to classical emergentism, I consider the possibility that the properties of complex objects are determined by non-causal laws. Here we need to pay attention to a distinction that I mentioned earlier: the formation of a molecule from two originally isolated atoms is a causal process that evolves through time. We can consider that the two H atoms gradually come closer together until their electron orbitals partially overlap. But this causal process is not determined by the law (*) that determines the energy levels of the system as a function of the distance R between the nuclei and as a function of the Hamiltonian of the system, determined in turn by laws governing the interactions among the properties of the parts of the system. The determination of the system's energy levels and the distance R_0 at which this energy is minimal is not a process that evolves through time, in which case it would make it possible to justify the distinction between temporally separate cause and effect. The calculation proceeds in stages and takes time, but as Duhem (1906, 25) points out this does not warrant concluding that there is a real process whose stages correspond to the stages of the calculation. The calculation takes into account the electrical charges of the parts of the system as well as the fact that, when the electron orbitals partially overlap, a quantum "resonance" phenomenon occurs that leads to a decrease in the energy of the state of the system, thanks to the possibility that the electron occupies a hybrid state around the two protons.

To give just the general idea of this calculation, it is assumed that the state ψ of the electron must be an eigenstate of the Hamiltonian $\mathcal{H}$.[28] In the method of variations, it is assumed that this state ψ results from a linear superposition of the states ψ_1 and ψ_2, the stable states of an electron bound to one of the H atoms in the absence of the other. We know that the ground energy state is a minimum of the mean value of $\mathcal{H}$, where that minimum is a function of the distance R between the two protons, considered as a parameter. This minimum corresponds to the eigenvalue E of the energy of the system as a function of R. It turns out that the function $E(R)$ has a minimum. This can be explained by the presence in the energy term of a "resonance integral," which represents a lowering of the energy due to "the possibility

28 For details, see Cohen-Tannoudji, Diu, and Laloë (1977, 1: 409–10; 2: 1159–71).

for the electron to 'jump' from the vicinity of one of the protons to the other" (Cohen-Tannoudji, Diu, and Laloë 1977, 2: 1167). The existence of a value of R for which E is minimal corresponds to the existence of a stable ψ state resulting from a linear superposition of the two states ψ_1 and ψ_2, which correspond to the localization of the electron around one of the two protons. The stability of this superposed state, in turn, explains the existence of the molecule: the movement of the two protons away from each other, as well as toward each other, moves the system away from its state of minimum energy.

The stability of the molecular ion H_2^+, the distance R_0 (which characterizes the stable state), and the shape of the molecular electron orbital are deducible. But it is crucial to evaluate the information contained in these premises so as to discern whether this means that these global properties are emergent according to the criterion of deducibility. If that information bears only on the properties that the components possess when they are isolated, then the deduced global properties are not emergent; however, if the possibility of deduction requires that the premises contain information about the laws that determine the global properties from the interactions among the components, then these global properties are emergent according to criterion (C1) (although they are deducible).

The form of the deduction of the existence of a stable state in which the two H atoms are linked into a molecule corresponds to the second situation: according to (C1), the stability of the H_2 molecule is an emergent property. The possibility of deducing the stability of the molecule requires knowledge not only of the properties of the isolated atoms but also of the *laws of interaction* among the components of *different* H *atoms*. Insofar as the stability of the molecule cannot be deduced from the properties that its components possess *in isolation*, it is emergent.[29] Furthermore, the laws of interaction cannot be deduced from the laws determining the evolution of the components in isolation.

However, there are reasons to think that the criterion (C1), which made it possible to obtain this result, is inadequate. It appears that the condition expressed by (C1) is too strong when we reconsider the examples considered

29 Grelling and Oppenheim (1937–38, 1939) draw a parallel between the emergent properties of complex objects and the phenomena studied in *Gestaltpsychology*. The parallel lies in the fact that the prediction of a Gestalt as well as that of an emergent property "requires knowledge of certain structural relations among its parts" (Hempel and Oppenheim 1948, 261n18) and cannot be obtained from knowledge of the isolated parts alone.

above: (C1) makes all of the properties of complex objects emergent because knowledge of the isolated evolution of the parts never allows us to deduce the consequences of their interactions. The form of the laws of interactions is never an a priori consequence of the laws of the isolated evolution of the parts. As Broad says, "it is clear that in *no* case could the behaviour of a whole composed of certain constituents be predicted *merely* from a knowledge of the properties of these constituents, taken separately, and of their proportions and arrangements in the particular complex under consideration" (1925, 63). Predicting the result of an interaction always presupposes knowledge of an a posteriori law. Predicting the behaviour of the whole presupposes "that we have found a general law connecting the behaviour of these wholes with that which their constituents would show in isolation"; this law is the "law of composition" specific to the system (63).

Criterion (C1) must therefore be rejected. In fact, it does not offer the means to ground the intuitive distinction between emergent and resultant properties because it imposes such a strong condition on non-emergence that all global properties end up appearing as emergent.

Given this observation, Broad suggested an alternative criterion that takes account of the fact that deducing a global property of a system necessarily requires information about the laws governing the interaction of the parts. According to Broad's criterion,

> (C2) the global property *G* of a complex system *s* is (strongly) emergent if and only if (it is weakly emergent and if) the law of composition the knowledge of which makes it possible to deduce the presence of the property *G* is a *law specific to the type of system s which cannot be derived from laws applying to parts of s in isolation and from laws of composition specific to other types of system*.

(C2) differs from (C1) only by the reference to knowledge of the behaviour of the parts *in other combinations*. Consequently, the properties of AgCl would not be emergent according to (C2) if they could be deduced from compositional laws for *other* molecules containing Cl and Ag. However, it is doubtful that this corresponds to a real relaxation of the conditions of non-emergence. The possibility of using knowledge of the behaviour of Cl and Ag in other combinations, for example in sodium chloride NaCl (common salt), to deduce the behaviour of silver chloride necessarily requires knowledge of general laws that apply — in this example — to all molecules of which chlorine is a component. Furthermore, knowledge of the laws of composition concerning a set of types of molecules including the Cl component (NaCl, HCl, etc. but not AgCl) does not logically imply either the law of composition of compounds outside this set, such as AgCl, or general laws that apply to all Cl compounds. This is the problem of induction. The knowledge of the laws of composition of several types of molecules including Cl provides only the premises for an *inductive* argument leading to such a general law. Accordingly, it appears that the criterion of non-emergence (C2) proposed by Broad — despite the introduction of the clause "or in other combinations" — is just as strong as criterion (C1). Furthermore, (C2) is too weak because according to it all properties of complex systems come out as emergent, just as in (C1).

However, my reasoning indicates a criterion that is stronger than (C1) and (C2): we can hypothesize that the non-emergent properties are those that can be deduced from the complete information on the components of the system in isolation, as well as from *general laws* determining the interactions of these components, without recourse to laws specific to the type of system in question.

> (C3) A global property *G* of a complex system *s* is (strongly) emergent if and only if (it is weakly emergent and if) the law of composition — the knowledge of which makes it possible to deduce the presence of the property *G* — is a law specific to the type of system *s that cannot be derived from general laws applying to s as well as to other types of systems.*
>
> (C3 — non-emergence) A global property *G* of a complex system *s* is non-emergent if and only if the law of composition — the knowledge of which makes it possible to deduce the presence of the property *G* — can be derived from general laws applying to *s* as well as to other types of systems, without it being necessary to call on a law specific to the type of system *s*.

According to this criterion, which expresses a necessary condition for emergence, the property *G* is emergent only if the law that gives rise to *G* is a brute experimental law: it cannot be derived from more general laws (i.e., laws that apply to the properties of parts), even outside the type of system that possesses

G. On the contrary, a property of a complex object determined by *general laws* is not emergent. If the law of composition can be derived from general laws that apply to the properties of parts even outside the particular system under consideration, then the property is explicable to the same extent as the law. It is therefore non-emergent according to criterion (C3).

This new criterion of non-emergence is really weaker than criteria (C1) and (C2). According to (C3), the properties of the molecular ion H_2^+ are no longer emergent with respect to the properties of its components because their derivation only calls on general laws determining the components of the Hamiltonian of any system possessing components of the same types. However, (C3) still seems to be too strong as a criterion of emergence (or too weak as a criterion of non-emergence): it is true that many global properties of complex systems, including chemical properties, cannot be explicitly derived at present from knowledge of the parts and from general laws. But at least as far as chemical and biological properties are concerned, the successes of deductive explanations of certain paradigmatic properties — such as the stability of the H_2 molecule and the ability to transmit hereditary traits — constitute paradigms around which research programs aimed at reductive explanations in these fields have been built. From the point of view of such a research program, the systems that count as emergent according to (C3) appear to be emergent only provisionally or epistemically. (C3) seems to analyze a *concept of provisional or epistemic emergence*. From this point of view, properties that cannot (yet) be deduced appear as "anomalies" (Hüttemann and Terzidis 2000, 274) bound to disappear as soon as general laws and adequate deductions are discovered. From the point of view of the research strategy of reductionism that bets on the existence of such laws and deductions, it seems to be reasonable to presume that in the long run the chemical properties of molecules will prove not to be absolutely emergent according to (C3), compared with the physical properties of the atoms that make them up: today's quantum mechanics allows us to deduce certain simple but fundamental properties of molecules from general principles and premises concerning the component atoms. Other chemical properties seem to differ only in the complexity of their derivation, not by any fundamental ontological difference. Ontologically, (C3) is too weak a criterion for non-emergence because all complex properties satisfy it; in other words, it is too strong a criterion for emergence because no complex property satisfies it.

Another way of seeing that (C3) is too strong a criterion for emergence is this. We are looking for a criterion to justify the intuitive distinction between emergent properties that are qualitatively new and only appear in systems satisfying precise structural constraints and resulting properties that belong to complex systems but are only quantitatively different from properties belonging to their parts or to simpler systems. This distinction is ontological, not epistemic. Belonging to one or another of these categories must therefore be independent of our theories. We can be wrong, of course, about whether a property is emergent or resultant. However, our intuitive concept of emergence is incompatible with a conception that makes the emergent or resulting character of a property systematically relative to the scientific theories accepted at a given moment in history. Yet this is a consequence of (C3). Let us say that property *G* is determined by the law of composition *C*. *C* can be more or less integrated into a theoretical system. It might start out as a "brute" empirical postulate and later become derivable from other empirical laws, which in turn are brute regularities without explanation, before a way of deducing *C* from the most general axioms and principles of theory is finally discovered. As *C* is deductively integrated into an increasingly powerful body of theory, *G* gradually loses its emergent character and is transformed into a resultant property.

We must therefore find a criterion that does not categorize all macroscopic properties as emergent — as in the cases in (C1) and (C2) — but that also does not categorize them all as resultant — as in the case of (C3).

5. Emergence as Non-Aggregativity

William Wimsatt (1986) suggested a fruitful way of conceiving of systemic properties that, though explicable and therefore in a sense reducible to the properties of the system's components and their interactions, are nevertheless emergent, in the sense of appearing only in systems with a specific composition and organization. Wimsatt identified four features that distinguish the properties that he called "aggregative"[30] from emergent properties.

The *emergent* properties of a complex object depend on the organization of its parts. A property of a complex object is *aggregative* if it does not depend

30 The term "aggregative" is equivalent to the term "resultant," used to designate non-emergent properties in the tradition of British emergentism.

on the organization of its parts. The mass of a pile of stones is a paradigmatic example of an aggregative property. This mass does not depend on the spatial arrangement or interaction among the stones that make up the pile. The heap can be dismantled and reassembled at random without any change in the overall property of its mass.

Of course, we could have excluded the mass of the heap of stones from the emergent properties for the simple reason that it is not systemic: the mass of the pile is a property *of the same kind* as the mass of the individual stones. However, the application of this criterion remains intuitive until we have a clear criterion of what counts as "the same kind" of property. The criteria proposed by Wimsatt are intended to remedy this lack of clarity: according to his first criterion, a global property not qualitatively different from the properties of its components is recognized by the fact that its presence does not depend on the organization of its parts and therefore does not undergo any modification when these parts are swapped. The mass of a heap of stones is aggregative because it does not vary when two stones exchange their positions. The capacity of a given portion of DNA, conversely, is a non-aggregative property according to this criterion: the expression of a gene depends on its location downstream from a control sequence. Accordingly, substituting one sequence for another can modify the conditions of its expression. However, according to this criterion, non-aggregativity is not necessary for emergence. Wimsatt identifies four ways in which a property of a complex system can be emergent or non-aggregative. Each is sufficient, but none is necessary, for a property to be emergent. Invariance with respect to the permutation of parts is not necessary for emergence because there are emergent properties — such as the transparency or the rhombohedral shape of a quartz crystal — not modified by the permutation of parts. Permuting molecules or parts of the crystal does not affect these overall properties.

According to Wimsatt's second criterion, a property is aggregative if it changes quantitatively but not qualitatively when parts of the system are added or removed. The stability of an arch composed of stones with a trapezoidal cross-section is non-aggregative according to this criterion, insofar as the arch loses its stability if one of the stones in it is removed. However, this criterion is not a necessary condition for emergence either. The transparency and rhombohedral shape of a quartz crystal are emergent properties in the sense that the microscopic components of the crystal have neither transparency nor rhombohedral shape nor any qualitatively similar property. But the crystal remains transparent and rhombohedral even if some molecules are removed or added.

There is a third way in which a property can be non-aggregative: it is possible for an overall property to be modified without adding or subtracting parts (as in the second criterion) and without permuting components (as in the first criterion). It can be modified by changing the spatial organization of the parts. According to this criterion, the ability to lead to the expression of a phenotypical trait is a non-aggregative property of genes, insofar as a change in *the spatial arrangement* of genes — for example, because of recombination — affects the expression of a gene depending on the presence of the appropriate control units at precise locations upstream from the DNA sequence. This criterion is a generalization of the first. It does not provide a necessary condition for emergence for the same reason as the first; some emergent properties resist decomposition and rearrangement of parts.

According to Wimsatt's fourth criterion, a global property of a system is non-aggregative if its existence depends on interactions among the parts of the system. The interaction among the four subunits of a hemoglobin molecule, for example, reduces the energy required to bind oxygen; this capacity is therefore a non-aggregative, or an emergent, property according to this criterion. All emergent properties are undoubtedly dependent on such interactions. However, not all interactions give rise to emergent properties: the mass of the heap of stones remains an aggregative property of the heap even if the stones attract each other according to gravitation. This criterion is therefore necessary, but not sufficient, for emergence. Therefore, we need to add a condition to specify which interactions are sufficient for emergence. This is the purpose of the criterion that I consider in the next section.

6. Emergence in Terms of Non-Linear Interaction and Mill's Principle of the Composition of Causes

If we accept the idea that all global properties are determined by general laws from the properties of the components, then it seems to be promising to try to ground the distinction between emergent and resultant properties on the mathematical form of these general laws of interaction. We can draw inspiration from a fundamental distinction introduced at the origin of the emergentist tradition by John Stuart Mill. He distinguishes two ways in which two (or more) interacting causes can determine their common effect. In the mechanical mode of interactive determination, the effect is the mathematical sum — arithmetic or vector — of the effects that each cause would have had if it had acted alone. Mill characterizes effects determined in this way by saying that they "obey the principle of the Composition of Causes" (1843, Book III, Chapter VI, para. 2). He distinguishes them from effects determined according to a "second mode" that corresponds to "a breach of the principle of Composition of Causes" (1843, Book III, Chapter VI, para. 2). Using the terminology introduced by Lewes (1875), emergentists later called effects determined in the first additive way "resultant effects." If the effect of several causes is not the mathematical sum of the effects that would have resulted from the separate actions of the causes, Mill says, then the determination of the effect obeys a "heteropathic law" (Book III, Chapter VI, para. 2) and produces a "heteropathic effect" (Book III, Chapter X, para. 4). In the vocabulary of Lewes, we may speak of "emergent law" and "emergent effect."

We can use this distinction between two types of determination as a basis for the distinction between emergent and resultant properties. However, this presupposes that we dissociate the Millian distinction from the analysis of *causal* laws and apply it to compositional laws.

The law of composition (itself reducible to the laws determining the interactions among the parts of a complex object or system) expresses the dependence of the global properties of the complex object on the properties that the parts possess in isolation. The distinction between emergent global properties and resulting global properties can be grounded on the mathematical form of the interaction laws as well as the law of composition. If the interaction laws have an additive or linear form, then the law of composition will also be linear. According to the criterion inspired by Mill, the properties determined by a linear law of composition are resultant, and the properties determined by non-linear laws of composition are emergent (Wimsatt 1996, S374).

> (C4) A global property *G* of the complex object *s* is (strongly) emergent if and only if (it is weakly emergent and if) the fact that *s* has *G* is determined by the fact that *s* has the parts $p_1 \ldots p_n$ having the properties $P_{11} \ldots P_{nm}$ as well as by a law of composition that is not a logical consequence of the laws governing the properties $P_{11} \ldots P_{nm}$ of the parts in isolation. The non-linearity of the law of composition is itself a consequence

of the non-linearity of the laws of interaction applying to the properties of the parts.

Mathematically, the general form that characterizes different forms of addition is the linear form $y = ax+b$. The total force acting on a body is a paradigmatic resultant property. It is determined by a linear superposition of the component forces. This criterion, therefore, justifies the intuition that the total force acting on a body *results from* the parallel action of the component forces. (C4) also justifies the intuition that the stability of the hydrogen ion H_2^+ is emergent with respect to the properties of the atomic level:[31] the Hamiltonian that determines the state of the molecule does not result from the addition of the Hamiltonians determining the state of each of the component atoms. I will come back to this in a moment.

For a systemic property to be qualitatively different from the properties of the parts of the system, the law of composition determining it must have a non-linear form. In Holland's words, "emergence is above all a product of coupled, context-dependent interactions. Technically these interactions, and the resulting system, are *nonlinear.* The behavior of the overall system *cannot* be obtained by *summing* the behaviors of its constituent parts" (1998, 121–22). The "context" on which emergent properties depend might lie within the system: in this sense, the context in which the parts are located is constituted by the other parts with which that part interacts. It can also be the "extra-systemic" context with which certain parts of the system interact. However, the second case can be reduced to the first case by broadening the contours of the system so that it includes what was originally considered to be extra-systemic.

Criterion (C4) makes emergence compatible with physicalism. All properties of complex objects are determined by empirical laws. The difference between resultant and emergent properties does not concern the possibility

31 Curiously, McLaughlin uses two different criteria for emergence at two different points in the same article, to arrive at opposite results, without noting the contradiction. When he sticks to the criterion established by Mill, Lewes, Alexander, and Morgan, according to which the existence of a property is a heteropathic effect of lower-level properties (i.e., the property is emergent) when it is not determined by a law of additive composition, McLaughlin notes that "the Emergentists were right about there being emergents" (1992, 75); however, when he uses the criterion that he attributes to Broad, according to which a property is emergent if its causal power obeys a law of "configurational force," he arrives at the result that there are no emergent properties in this sense since there are no configurational forces (89–91).

of explaining and predicting them.[32] Provided that the laws that determine them are discovered, all of them can be explained and predicted in principle. The difference lies only in the form of the laws. Laws of composition that have the form of a linear function give rise to resultant properties; laws of composition that have a non-linear form give rise to emergent properties. The emergent properties of complex objects are qualitatively different from the properties of their parts because they are determined by non-linear laws of composition, but they are not irreducible.[33] By making emergence compatible with the possibility of reducing even emergent properties, and thus with physicalism, (C4) expresses a weaker concept than the traditional one that requires inexplicability, irreducibility, or unpredictability.

(C4) has retained Mill's distinction between linear and non-linear laws. However, there are some important differences between (C4) and Mill's original criterion. First, as we have seen, (C4) distinguishes between resultant and emergent properties as a function of the mathematical form of *non-causal* compositional laws, whereas Mill's distinction concerns causal laws and forms of *causal* determination. Second, Mill confuses the *ontological* distinction between laws of different forms (the distinction between linear and non-linear laws) with an *epistemic* distinction.

Let us consider the first difference. Mill distinguishes between two forms of causal determination: the difference is "between the case in which

32 Prediction and explanation are not always equivalent. Within the framework of the deductive-nomological model (and if we are dealing only with deterministic laws), explanation and prediction are analyzed in the same terms of the logical relationship (that of a valid deduction) between premises and conclusion. Only the direction in which knowledge is acquired differs. In explanation, we already know the conclusion, and we learn from which premises it can be deduced. In prediction, we learn which conclusions follow from premises already known. In the deterministic context, it is therefore legitimate to consider, as Broad does (see Stephan 1992, 38), the concepts of prediction and deduction (or deductive explanation) as equivalent. However, when the law is probabilistic (or statistical), it might be possible to explain an event because it can be deduced that it would happen with a certain probability. The question of whether this probability should be greater than a certain universal value is controversial (see Salmon 1990). Conversely, insofar as the probability is not 1, the event nevertheless can be considered unpredictable. It is conceivable that probabilistic composition laws exist (see Stephan 1992, 33). But this does not seem to be the case for the laws that interest us here. In general, therefore, it is not necessary to insist on the difference between prediction and explanation. This difference can be taken to be epistemic or pragmatic, in the sense indicated above.

33 Chalmers (1996, 378n41) and Bedau (1997, 375) call concepts of emergence of this kind "innocent."

the joint effect of causes is the sum of their separate effects, and the case in which it is heterogeneous to them" (1843, Book III, Chapter VI, para. 2). The former mode of composition of causes obeys the principle of the composition of causes, whereas the latter mode violates this principle. In contrast, (C4) uses the criterion of the mathematical form of laws to distinguish two forms of non-causal determination. The properties of complex objects are determined by the properties of their components and their interactions, according to non-causal laws of composition, specific to those components and interactions. Let us therefore ignore the fact that for Mill and other emergentists nomological determination is identified with causal determination, to reinterpret the terminology of Lewes (1875) in a non-causal way: properties determined by a linear law of composition are "resultant," whereas properties determined by non-linear laws are "emergent."

It is only because (C4) transposes the Millian distinction to non-causal laws that we can use it to ask whether the determination of the Hamiltonian of the molecular ion H_2^+ from the properties of its components obeys the "additive mode"[34] of (non-causal) determination or the "heteropathic mode." Stripped of its causal interpretation, Mill's question becomes that of whether the determination of the molecular properties of H_2^+ is obtained by means of an addition — vectorial or algebraic — or by means of a more complex law of composition. The form of

$$(*)\ \mathcal{H} = P^2/2m - e^2/r_1 - e^2/r_2 + e^2/R$$

shows that the interactions between the components of the molecule obey not an additive law but a "heteropathic" law. The Hamiltonian of H_2^+ is not the sum of the Hamiltonians of each separate body. For this reason, the properties of the molecule determined by this law are "emergent" rather than "resultant."

Second, criterion (C4) differs from Mill's in that it is *ontological*, whereas Mill gives his criterion an epistemic meaning. For him, the distinction

34 It is impossible to use the term "mechanical" in our conception of emergence. Being essentially a form of causal determination, a non-causal mechanical determination would be a contradiction in terms. This is not the case with the term "chemical," only contingently associated with a mode of causal determination.

between linear and non-linear laws is equivalent to the distinction between global properties that can be deduced a priori from knowledge of the parts in isolation and global properties whose existence can be discovered only a posteriori, through experience.[35] In his words, "it is impossible to deduce all chemical and physiological truths from the laws or properties of simple substances or elementary agents" insofar as the laws "of chemistry and physiology . . . owe their existence to a breach of the principle of Composition of Causes" (1843, Book III, Chapter VI, para. 2). In particular, "the Laws of Life will never be deducible from the mere laws of the ingredients" (1843, Book III, Chapter VI, para. 2). The thesis of the impossibility in principle to deduce chemical and physiological laws from physics makes Mill's conception incompatible with physicalism.

However, Mill is wrong to regard the ontological distinction between two forms of law as equivalent to the epistemic distinction between two forms of explanation. In reality, no law of interaction is a logical consequence, knowable a priori, of the set of laws governing the evolution of properties in isolation (apart from interactions). In other words, if we take as resultant only those properties *G* of *s* determined *logically* from the properties of the components of *s*, then all properties would be emergent. Without any empirical law, no property of *s* can be deduced. Indeed, even the additive composition of properties that gives rise to a resultant property is not a priori: it merely obeys a particularly simple law.[36] The mode of additive composition is not distinguished by its a priori character: we cannot know a priori how two bags of flour of 1 kg mass each will interact when the grocer places them together on the scale, any more than we can know a priori how the photons of superposed

35 Popper (1977) seems to confuse these two distinctions (between a priori and a posteriori and between linear and non-linear) insofar as he considers as emergent all properties *P* of complex objects *x* not logically (or a priori) deducible from the properties of the components of *x*. Conversely, Broad acknowledges that these distinctions are not equivalent. I will return to this issue in a moment.

36 It is perhaps this point that Stephan has in mind when he says that, "to deduce the weight of the whole, one must also invoke the principle of the additivity of weight. That principle, while nomological, is logically contingent" (1992, 35). According to the conception of laws that I have defended elsewhere (Kistler 2002a, 2005a), laws are metaphysically necessary. The difference between logical principles and laws of nature is therefore epistemic. We can know the former a priori but the latter only a posteriori.

rays of light will interact.[37] Knowledge of the laws of interaction is always a posteriori.[38]

Emergent properties appear if a non-linear law governs the interaction.[39] For example, the mass of a solid object is smaller than the sum of the masses of the atoms that compose it: the stability of a solid object originates from the fact that, in the solid state, the energy of the system is less than the sum of the energy of the atoms. This reduction in energy is equivalent to a small reduction in mass. Therefore, the mass — and consequently the weight — of a macroscopic body is emergent in relation to the masses of its atomic components.

Kronz and Tiehen (2002) criticize the idea of grounding the concept of emergence of a property *G* on the non-linearity of the laws of interaction that generate *G* by an analysis of the origin of the existence of entangled states in quantum systems, which are paradigmatic cases of emergent states. According to them, "the mark of a non-emergent property of composite systems in quantum mechanics crucially involves a *multiplicative* operation,

37 Nagel expresses this point as follows:

When the matter is viewed abstractly, the "sum" of a given set of elements is simply an element that is *uniquely determined* by some *function* (in the mathematical sense) of the given set. . . . [T]he question [of] whether such a function is to be introduced into a given domain of inquiry, and if so what special form is to be assigned to it, cannot be settled a priori. (1952, 23)

Nagel uses the example of the luminosity of a surface illuminated by two sources: "The physical brightness of a surface illuminated by two sources of light is sometimes said to have for one of its parts the brightness associated with one of the sources" (19). He then remarks (22) that one can speak sensibly of the sum of the two luminosities only if the light is monochromatic.

38 Broad (1925, 62 ff.) explicitly recognizes that principles of composition are logically contingent and a posteriori. Therefore, he does not explain the greater transparency of resulting properties compared with emergent properties, which corresponds to the possibility of explaining the former, but not the latter, by the particular (namely, a priori) epistemological status of principles of composition. Hempel and Oppenheim's (1948) argument — taken up by Van Cleve (1990, 224) — that Broad is wrong to consider the mass of a compound object as a resultant because the law of the composition of masses is contingent is therefore based on a misunderstanding (McLaughlin 1992n38). Hempel and Oppenheim wrongly think that Broad takes the principles of composition of resulting properties to be necessary. I can add that Hempel and Oppenheim and Van Cleve are wrong in their judgment of the modal status of the law of composition. Van Cleve argues that "the parallelogram law for the composition of forces is logically contingent" (224), concluding that the Broadian conception of emergence makes the behaviour of a three-body system emergent. The law of composition of forces is an empirical law of nature. If we substitute "a posteriori" for "logically contingent," then we obtain a correct judgment in line with Broad's doctrine.

39 Non-linearity is necessary for emergence. I do not claim that it is sufficient.

factorability into tensor product vectors (in the case of states) or matrices (in the case of properties), rather than an *additive* one" (2002, 333; italics added). Indeed, the mathematical operation that represents the transition from the description of two separate systems to the description of a complex system is the tensor product. If F_1 and F_2 are vectors, each of which represents a two-state system, then the vector representing the complex system formed from these two systems is $F_1 \otimes F_2$.

The entangled states of complex quantum systems, which, according to Kronz and Tiehen (2002) correspond to emergent properties, are distinguished by the impossibility of expressing them as tensor products of lower-dimensional states. This observation warns us against an abusive simplification of the criterion of emergence. It refutes the idea that any state of a complex system, whose description can be obtained from a multiplicative procedure based on the description of its constituent parts, is emergent. However, the objection raised by Kronz and Tiehen against such a simplistic conception does not call into question my criterion of emergence, which concerns the form of the law of composition and itself is a consequence of the set of laws that governs the interactions among the parts. In quantum mechanics, this law is represented by the Hamiltonian of the system, which determines its evolution and properties, in particular the eigenvalues of its observables. And it is indeed the presence of non-linear terms in the Hamiltonian that is responsible for its non-separability, which in turn is responsible for the non-separability of the system and the fact that the state of the intricate system is emergent. This means that we cannot represent the system and its evolution as a conjunction (whether by addition or by multiplication) of subsystems, each of which evolves as a function of its own Hamiltonian, which is "separable" (i.e., whose expression is independent of references to the other subsystems). The properties of the parts of such an intricate system cannot be represented without reference to the other parts. In quantum mechanics, a part p_1 of a system *s* is described by a "density operator" $\rho_1(t)$ obtained by forming the trace over the complementary part p_2 of the system. The mathematical development shows that the evolution of p_1 can be described by a separate operator of evolution $U_1(t)$, only if the Hamiltonian is separable. If it is not separable because it contains non-linear terms of interaction, then the evolution of p_1 is at all times dependent on the evolution of the whole system *s*. Mathematically, the operator U(t) of the evolution of the whole system is separable: that is, it can be represented as the product of two operators $U_1(t)$

and $U_2(t)$ so that $U(t) = U_1(t) \otimes U_2(t)$, if and only if the Hamiltonian of *s* is separable: that is, it can be written as the sum of two Hamiltonians $\mathcal{H}_1$ and $\mathcal{H}_2$ describing separately the subsystems p_1 and p_2. In mathematical terms, we switch from addition to multiplication because U is an exponential function[40] of $\mathcal{H}$: $U(t) = \exp(-i\mathcal{H}t)$.

Given the fundamental role of the Hamiltonian, this seems to show that the criterion of the additivity of laws of interaction is applicable in quantum mechanics. However, the criterion can be generalized outside that framework.[41]

7. Qualitative and Quantitative Difference

The qualitative "novelty" (Alexander 1920, 14n2) of the emergent properties of a whole in relation to the properties of its parts is one of the most important criteria of emergence for the classical emergentists. Novelty is also the fundamental feature of emergence in the definitions of emergence offered by Bunge (2003, 17)[42] and Sperry (1986, 267).[43]

In line with my thesis, Bunge argues that the emergence of a property is compatible with the possibility of explaining or deducing it: "It is mistaken to define an emergent property as a feature of a whole that cannot be explained in terms of the properties of its parts. Emergence is often intriguing but not mysterious: explained emergence is still emergence" (2003, 21).[44] However, the concept of novelty used by Bunge is too vague to support the distinction between emergent and resultant properties. According to him, a property *P* of a complex object *x* is resultant if there are components of *x* that also have

40 This presupposes that $\mathcal{H}$ does not depend explicitly on time.

41 Kronz and Tiehen give "some measure of plausibility" to the idea that the presence of non-linear terms in the Hamiltonian of a classical system can be responsible for emergent properties insofar as "a classical system can exhibit chaotic behavior only if its Hamiltonian is nonseparable" (2002, 332).

42 Bunge does not distinguish, in this definition, between synchronic and diachronic emergence. His conception of emergence makes any property acquired over time, even during a simple spatial shift, an emergent property.

43 In Popper's words, "there is the emergence of life . . . [that] creates something that is utterly new in the universe" (1977, 342). A little later, he says that "the fact of the emergence of novelty, and of creativity, can hardly be denied" (343).

44 Bunge also says that "every emergent property of a system can be explained in terms of properties of its components and of the couplings amongst these" (1977, 503; quoted by Stephan 1992, 31).

the property *P*. In other words, Bunge identifies emergence with what I called above the systemic character of a property. But we have seen that being systemic is not enough to be emergent. Without a distinction between systemic and emergent, Bunge cannot avoid considering all macroscopic properties as emergent. The property of weighing 100 kg is "new" compared with the property of weighing 50 kg, whereas it is clearly a resultant property. For his criterion to acquire a precise content, it is necessary to give a rigorous meaning to the distinction between qualitative novelty and quantitative novelty so that it corresponds to the intuitive difference between emergent and resultant properties.

Assumption (C4) suggests that the non-linear form of the law of composition is responsible for the qualitative difference between the emergent properties and the properties of the parts. Psychological research has established the existence of psychophysical laws that can be described via functions from stimulus space to representational space. As we saw earlier (Chapter 3.5), the metric properties of the latter can be determined empirically (Shepard 1962; Clark 1993). With knowledge of these psychophysical laws of composition, it becomes possible to predict the secondary quality or the qualitative appearance associated with a stimulus that has not yet been experienced (by some particular subject or even by anyone), for example the smell of a perfume made for the first time. As Feigl says, with knowledge of psychophysical laws and of the representational space of odours, "we should be able to predict the location of the quality in the topological space of odors" (1958, 416) by means of extrapolation.

The phenomena of colour vision can illustrate the appearance of new properties: the qualities of experience or "qualia." When we see light produced by the superposition of rays of different colours, the perceived colour is qualitatively different from the colour of the components. It is not the result of a simple superposition, or addition, of the sensations produced by each of the stimuli taken separately. The phenomenal appearance of the perceived colour is the result of a complex process. The signals received by the light-sensitive cells in the retina — the cones — are transformed by a set of "opponent processes," postulated in psychological terms by Hering (1920). The neurophysiological mechanism underlying this process was discovered by Hurvich and Jameson (1957). Through the interplay of reinforcement and inhibition, the three types of light-sensitive cones located in the retina, with maximum sensitivity respectively to long (*L*), medium (*M*), and short (*S*) wavelengths, give rise to three signals: an achromatic transmission channel made up of neurons that "add" the signals from the *L* and *M* receptors, giving rise to the perception of white when the signal is positive, and the perception of black when the signal is negative. A second, chromatic, channel is made up of neurons that "subtract" the signals from the *L* and *M* receptors. The signal from this channel produces perceptions of red, if the net signal is positive, or perceptions of green, if the net signal is negative. Subtracting the signal from the receptors sensitive to the shortest wavelengths (*S*) from the sum of the signals from *L* and *M* gives rise to a third signal that produces perceptions of yellow, if the result of this interaction is positive, or blue, if it is negative (see Lennie 2000, 577–78).

This mechanism explains why it is impossible to perceive a colour that is a mixture of red and green or of yellow and blue: when we are exposed to light composed of a superposition of rays of two exactly opposite colours (in the sense that their signals come from the same place in the visual field but are of opposite signs) and of the same intensity, the signal is zero. The opposite colours, rather than producing a mixed or combined appearance, cancel each other out. As Hardin puts it,

> since the neutral point of an opponent pair is achromatic, any stimulus that will put both chromatic opponent systems in balance will yield an achromatic perception. . . . If we are given a yellowish green light, we can render the appearance of this stimulus achromatic by adding enough blue to offset the yellow and enough red to balance the green. . . . Whiteness may be generated by as few as two wavelengths and in an indefinitely large number of ways. (1988, 38–9)

It is equally interesting to disregard the details of the complex mechanisms that lead from the absorption of light of different wavelengths to the perception of colours and instead to examine the result of this process: the structure of colour representations. As we have seen in Chapter 3, it is possible to reconstruct the structure of the cognitive space of colour representation, and in particular the minimum number of its dimensions, just from judgments of similarity between perceived colours. We find that the colour representation space has two dimensions and that the colour representations of the visible colours occupy points on the circumference of an approximately circular structure (see Figure 3.1).

The net result of the treatment of the physical stimulus by the colour vision system, therefore, is the transformation of a one-dimensional space into a two-dimensional structure located in a two-dimensional space. The simplest physical stimuli that give rise to the perception of colours are "monochromatic" rays in the physical sense (i.e., rays that contain only one wavelength). These coloured stimuli are ordered in a single dimension, which corresponds to a portion of the series of wavelengths, starting at a wavelength of around 400 nm and ending at a wavelength of around 700 nm.[45] The net result of the treatment of light signals by the visual system is the emergence of coloured perceptions.

The psychophysical laws that determine colour representations from physical stimuli are clearly non-linear since they associate representations located around the circumference of a circle with stimuli ordered along a single dimension. More generally, Shepard's work on the structure of the representational space of a certain number of phenomena belonging to different sensory modalities, shows that the transformation between stimuli and representations does not preserve relations of proximity. Intuitively, the projection of the space in which the stimuli are located onto the representational space is accompanied by a deformation that is generally inhomogeneous: the representations of two pairs of stimuli at equal distance from each other in physical space generally will not be at equal distance from each other in representational space.

However, the non-linear form of the law of composition is not a sufficient condition for qualitative novelty. According to the Weber-Fechner law, the intensity of sensations is a logarithmic function of the intensity of stimulations: $S = k \log I$, where S is the intensity of the sensation, I is the intensity of the stimulation, and k is a constant.[46] However, the intensity of the sensation caused by a stimulus of double intensity, which results from the superposition of two rays of the same colour, differs only quantitatively (by $k \log 2$)

45 Beyond these limits, electromagnetic waves do not give rise to coloured perceptual experiences and are therefore not visible. Rays with wavelengths just under 400 nm are called infrared; rays with wavelengths just over 700 nm are called ultraviolet.

46 The unit of measurement for S is the differential threshold (i.e., the sensation corresponding to the smallest perceived difference in intensity).

from the intensity of the sensation caused by a stimulus of single intensity. Accordingly, we need to supplement criterion (C4) with a condition guaranteeing the new quality of the emergent property.

Topological equivalence can be used as a mathematical criterion to account for the intuition of the qualitative difference between physical stimuli and sensations. The concept of topological equivalence between two spaces is formally defined by the existence of a homeomorphism between these spaces. If such a homeomorphism exists, then the spaces themselves are said to be homeomorphic, meaning that they are topologically identical. A homeomorphism is defined as a continuous bijection whose inverse is also continuous. It is not necessary to give the definitions of these mathematical terms in order to acquire an intuitive notion of topological equivalence. We can use, in fact, this intuitive criterion: two spaces are topologically equivalent if we can deform one into the other, without cutting it up or merging its different parts.

The topological non-equivalence of the one-dimensional space of physical monochromatic colours and the two-dimensional space of colour representations is manifested by a number of phenomena. The simplest of them are the complementary colours and the similarity of the colours corresponding to the stimuli at the extremes of the ordered series of physical stimuli, namely violet and red. These stimuli are separated by the greatest possible interval. In other words, violet and red are the most physically dissimilar of all the pairs of stimuli. However, their representations are no more dissimilar, for example, than the representations of yellow and green. Representations of colours located in diametrically opposed positions on the circular psychological space form complementary contrasts, giving rise to a number of well-known phenomena that have no equivalent in terms of physical stimuli: a ray composed of two rays of complementary colours of equal intensity is perceived as white (i.e., devoid of chromatic colour). Also, a small coloured surface surrounded by grey is perceived as surrounded by the opposite colour: a red disc is perceived as surrounded by green even though the background is perceived as grey if it is not juxtaposed to red.

It seems that the criterion of topological equivalence can be used to give a mathematically rigorous meaning to the idea of qualitative difference. The same criterion can be applied to physical systems. As far as the dynamic properties of a physical system (considered in classical mechanics) are concerned, a change between two dynamic states of a system is quantitative only when the trajectories of the system in its phase space — in these two states — are topologically equivalent; conversely, it is a qualitative change if the trajectories of the system are not topologically equivalent.

We can use the topological concept of novelty[47] in the definition of emergence as follows: when a purely quantitative change in microscopic parameters gives rise to a qualitative change in the trajectory of the system, this trajectory corresponds to an emergent property of the system.[48] When it is a resultant property of the system, a quantitative change in microscopic properties leads only to a topologically equivalent transformation of the system's trajectory.

However, the requirement of topological difference is too strong. The example of gases obeying the van der Waals equation shows that qualitative changes that do not correspond to topological differences can underlie emergence. Under certain conditions, these gases undergo phase transitions,[49] qualitative changes in their overall properties.

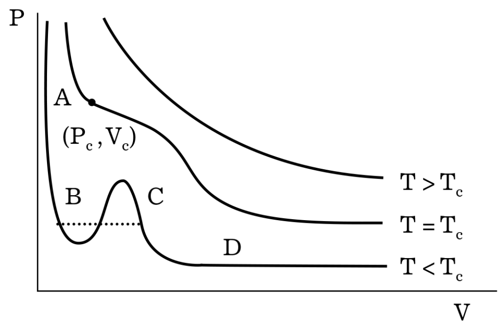


Figure 4.1 Graphs of three isotherms for a van der Waals gas. Critical point: critical pressure p_c, critical volume V_c, critical temperature T_c. Isotherms are functions of pressure, dependent on volume, at a fixed temperature. From Rueger (2000a, 484).

Figure 4.1 shows three curves, known as "isotherms," that exhibit the dependence of pressure on volume at constant temperature. The three isotherms

47 Anderson (1972) suggests using a criterion of symmetry breaking to characterize the emergence of qualitatively new properties in physical systems.

48 The trajectories of undamped and damped oscillators (Rueger 2000a) are not topologically equivalent.

49 Batterman (2002, Chapter 4) gives an accessible presentation of the reductive explanation of the phenomenon of phase transition.

shown in the diagram are qualitatively different because they do not have the same variations. Above a certain temperature, known as the "critical temperature" T_c, the volume of the gas is a strictly monotonic function of pressure; any decrease in pressure is accompanied by an increase in volume. At temperatures higher than T_c, the gas cannot be liquefied; the system remains in the gaseous state (also called gaseous "phase").

However, below the critical temperature, the relationship between pressure and volume is not monotonic. When we reduce the pressure of a liquid initially under high pressure (point *A* in Figure 4.1) at constant temperature, we reach a point (*B* in the figure) where, if we transfer heat to the liquid, the volume increases (up to point *C*) while the pressure remains constant. This process corresponds to a phase change from the liquid state to the gaseous state, which appears in the characteristic form of the isotherm for all values of $T<T_c$. The isotherms for $T<T_c$ are qualitatively different from isotherms at temperatures higher than T_c (see Rueger 2000a, 485).

We can thus reason as follows: the systemic property of being liquid or gaseous is emergent with respect to the microscopic properties of the components (the molecules) because certain small, purely quantitative, changes in the properties of the components are responsible for qualitative changes in the systemic property. A small change in the kinetic energy of the molecules (which corresponds to a change in the temperature of the gas) that causes it to pass through the critical value T_c results in a radical change in the dependence of pressure on volume. This change corresponds to the transition from a regime without phase transition to a regime with phase transition.

A consequence of my hypothesis is that a system that does not exhibit phase transitions, for example an ideal gas, has only resultant global properties: they change only quantitatively when the properties of the components undergo a small quantitative change, such as a change in kinetic energy. In contrast, a system — such as a van der Waals gas — that undergoes phase transitions has emergent global properties since these properties can change qualitatively when there is a small quantitative change in the properties of the components.

Rueger (2000a) presents systems with phase transitions as examples of "diachronic emergence." According to Rueger, a new global property, for example that of being in a gaseous state, emerges over time from another global property, that of being in a liquid state, when the pressure drops and the system is below the critical temperature. However, we can also use concepts of mathematical similarity and difference to characterize the concept of emergence of interest here. Synchronic emergence characterizes global properties with respect to the properties of the components that determine them synchronously. Unlike Rueger, I do not consider the qualitative change itself during the phase transition as a case of emergence (in the course of time); rather, I take it as a *criterion* showing that the global property that undergoes this qualitative change is emergent. One can justify this use of qualitative change as a criterion of synchronic emergence in the following way: since a quantitative change in the properties of the components determines (by a non-linear law of composition) a qualitative change in a global property, there must be a qualitative difference between the properties of the components and the global property. If the properties C (of the components) are only quantitatively different from the properties G (global), then properties C_1 and C_2, which are only quantitatively different, would determine properties G_1 and G_2, which also would be only quantitatively different. In the case of substances with a phase transition, the antecedent of the latter conditional can be true and the consequent false (in my example, if the difference between C_1 and C_2 corresponds to the difference between a state below T_c and a state above T_c); therefore, the G properties are not only quantitatively different from the C properties.

In this example, the global properties examined are dynamic properties: they are dispositional properties (powers) of the system that determine its evolution. This does not prevent them from being intrinsic properties that the system possesses at a given moment. The properties that are emergent according to this criterion, in particular the property of being disposed to a qualitatively new dynamic behaviour, are "new" compared with the properties of the components because the emergent properties are subject to different laws: they follow an evolution qualitatively different from the evolution of the isolated components by virtue of their properties and the laws that apply to them. It is the difference in the laws to which the properties are subject that proves that they are indeed different properties.[50] Accordingly, the emergent property imposes a specific evolution on the system. At each moment when the system possesses the structural conditions required for emergence (i.e., when its parts have the appropriate properties and relationships), the law of

50 In Kistler (2002a), I show that the identity of a property is determined by the set of laws in which it is involved.

composition has the nomological consequence that the system possesses the emergent property. This nomological determination is synchronic — even if the properties involved are predominantly characterized diachronically — by virtue of how their possession causes the system to evolve.

The property of a van der Waals gas of being disposed to undergo a phase transition is an emergent property: it obeys the fundamental criterion of novelty in relation to the properties of the components, a dynamic novelty accompanied by the novelty of the laws that it obeys. However, it does not obey the condition of emergence imposed by classical emergentists of being irreducible to the properties of the parts and their relationships or of being inexplicable or unpredictable. According to the concept of reduction developed above, the existence of a law of composition that determines (nomologically) that the system possesses G if its parts possess $P_{11} \ldots P_{nm}$ suffices for the reducibility of G. Once the law is known, the possession of G can be deduced, and consequently explained (and predicted), in the sense of the deductive-nomological model.[51]

The possibility of defining precisely what is meant by a qualitatively new property of a system, compared with the properties of its parts, allows us to add the condition of novelty to our necessary condition of emergence.

51 Rueger (2000a, 2000b) distinguishes between weak and strong emergence. A property is strongly emergent if it gives the system causal powers that the parts of the system do not possess. A property of a complex system is weakly emergent if, first, it gives the system a new dynamic behaviour and, second, it is irreducible. From the point of view of the nomological conception of the identity of properties, it seems to be contradictory to associate the novelty of the dynamic behaviour of a system with a weakly emergent property but to deny that this property gives new causal powers to the system. Only a different causal power can make the system evolve in a different way. However, Rueger's thesis, according to which even weakly emergent properties are irreducible, is only apparently incompatible with my thesis according to which emergent properties are reducible in principle. The appearance results from the use of a different concept of reduction. Rueger uses the concept of reduction common in science and analyzed by Nickles (1973). This concept characterizes a relationship between two successive theories describing phenomena at the same level, for example Newtonian mechanics and relativistic mechanics. Wimsatt calls it "successional reduction" as opposed to "explanatory reduction" (1976b, 677). In Chapter 1, I introduced the distinction between the "intralevel" successive reduction between two theories describing phenomena of the same level and the "micro-macro" explanatory reduction. An equation of relativistic mechanics, for example the equation that determines the addition of velocities, is said to "reduce" to the analogous equation of Newtonian mechanics, in the limit where the relative velocity is negligible compared with the speed of light. This sense of the term "reduce" has nothing to do with the sense that allows us to affirm the reducibility of emergent properties, which concerns the relationship between properties of different mereological levels.

(C5) A global property G of the complex object s is (strongly) emergent if and only if (it is weakly emergent and if) (1) the fact that s has G is determined by the fact that s has parts $p_1 \ldots p_n$ that have properties $P_{11} \ldots P_{nm}$, as well as by a non-linear law of composition, which is not a logical consequence of laws governing the properties $P_{11} \ldots P_{nm}$ of the parts in isolation, and if (2) the property G is qualitatively different from properties P in the sense of topological equivalence or other mathematical criteria.

The qualitative novelty of emergent properties gives rise to *stability*: the emergent property depends, of course, on the properties of the parts of the system, but it nonetheless possesses a certain autonomy that manifests itself as independence from certain microscopic changes in the parts. The fact that a system is in a solid or liquid phase is a *robust* emergent property.[52] An emergent property is robust in the sense that it has a certain (relative) independence from the properties of the components, such that a large number of small changes in the properties of the parts of the system does not induce a qualitative change in the overall property. The temperature of a gas, for example, is robust in this sense, with respect to many changes in the velocity and position of the particles that compose it: insofar as these changes — which occur continuously in the gas — do not affect the *average* kinetic energy over all particles and over time, the temperature is not affected at all and can therefore be said to be robust.[53]

52 Rueger (2000a) gives the opposite meaning to the concept of robustness, in which robust properties are non-emergent, or resultant, properties: a small change in the properties of components is reflected in a small change in the properties of the whole. For Rueger, emergent properties are therefore not robust in the sense that the non-linear nature of their dependence on component properties means that certain small changes in component properties are accompanied by (i.e., determine) radical changes in overall properties.

53 We can hypothesize that properties that are "robust" in this sense obey causal laws that are "insensitive" in the sense of Lewis (1986, 184) or "stable" in the sense of Woodward (2010). The robustness of emergent properties corresponds to the insensitivity of functional properties with respect to their different physical realizations, according to Putnam's machine functionalism. Indeed, Putnam points out that, insofar as mental states are structural functional states, they are invariant in terms of small changes in their physical realization. He draws an analogy to explanation. In his terms, "a good explanation is invariant under small perturbations of the assumptions" (1975b, 301). The good explanation refers to robust functional properties, and the assumptions refer to the physical realization. See Menzies and List (2010).

To bolster the concept of robustness as the invariance of a systemic property with respect to small variations in the properties of the components of the system, we can use the concept of *attractor* from the theory of dynamic systems. In the phase space used to represent the evolution of a dynamic system, a "fixed point" is a set of values of the variables determining the state of the system, for which the system remains stable. A fixed point is an attractor if it has a neighbourhood — called its basin — such that, if the system enters this neighbourhood, its trajectory converges toward the fixed point.

Using the notion of an attractor, we can hypothesize that each robust emergent property corresponds to an attractor, with respect to the trajectory of the system in its phase space. In this sense, the existence of an attractor can serve as a rigorous formal criterion for the existence of a robust property. Newman (1996, 2001) offers an analysis of the emergence of a property of a system in terms of its dynamics. In particular, he considers mental properties to be emergent properties of the brain, insofar as they correspond to the existence of basins of "strange" attractors in the phase diagram of the non-linear dynamic system of the brain (2001, 190). However, the condition imposed by Newman seems to be too strong: it is not necessary for the attractor to be of the "strange" type. The reason for imposing this condition is that a strange attractor causes non-periodic trajectories in phase space, which means that the system never returns to the same state twice and that two arbitrarily close points diverge exponentially in the course of the system's evolution (1996, 254–55). Newman argues that emergence requires the system to be in the basin of a strange attractor because this makes it impossible to predict the states of the system far enough into the future. But the impossibility of predicting the future states of the system seems to be neither necessary nor sufficient for emergence.

1. It is not necessary: the appearance of emergent properties, such as the ordered crystallized state during the solidification of a liquid, corresponds to a point attractor; it is perfectly predictable.

2. It is not sufficient: when the system is in the basin of a strange attractor, it does not robustly possess any systemic property. The state is sensitive to small variations in the initial conditions, whereas emergent properties are not.

The same objections can be made to a recent version of the epistemic conception of emergence, according to which a macroscopic property *P* of a dynamic system is emergent if the possession of *P* can be derived only from a microscopic description of the system and external conditions by making use of simulation (Bechtel and Richardson 1992; Bedau 1997; Holland 1998; Huneman 2008). This condition is both too weak and too strong.

1. The condition is not sufficient for emergence. A system can be complex in the sense that its state can be predicted only by simulation yet have no emergent properties. This is the case of chaotic systems that have only strange attractors, such as the double pendulum.

2. Nor is it a necessary condition for emergence. When a system has a robust emergent property, in particular a property that corresponds to the existence of an attractor, in the phase diagram of the system, that has the topology of a point or cycle, its appearance can be predicted without any simulation, simply from the presence of the state of the system in the basin of the attractor.

8. The Limits of Explaining Emergent Properties

We have seen that we can take up the vocabulary of classical emergentism only by imposing new meanings on the terms "emergent" and "resultant": none of these types of properties is determined by causal laws, and none can be discovered a priori. However, the non-causal (rather than causal) and ontological (rather than epistemic) interpretation of the concepts "emergent" and "resultant" is not the only, and perhaps not even the most important, difference between the classical concept of emergence and the concept developed here. The thesis whose refutation is generally considered to be responsible for the defeat of emergentism is that of the mysterious character of emergent properties; indeed, according to classical emergentists, the essential justification for the assertion that a property is emergent is that it is impossible to *explain* its origin.

Since Locke, the association of the perceptive experience of secondary qualities with certain stimuli has been considered as a paradigm of the mysterious character of emergent properties. Locke expresses the idea that this is a "brute nomological fact," inaccessible to any explanation, by saying that it is the impenetrable will of God, which appears to us as arbitrary, that associates a given stimulus with the smell of violets rather than with the pain caused by a steel knife cutting our flesh. He asks us to suppose

> that a Violet . . . by the impulse of such insensible particles of matter . . . causes the *Ideas* of the blue Colour, and sweet Scent of that Flower to be produced in our Minds. It being no more impossible, to conceive, that God should annex such *Ideas* to such Motions, with which they have no similitude; than that he should annex the *Idea* of Pain to the motion of a piece of Steel dividing our Flesh, with which that *Idea* hath no resemblance. (1689, Book II, Chapter VIII, para. 13)

Without using the notion of God, Alexander essentially expresses the same idea. In his words, "the existence of emergent qualities . . . is something to be noted, as some would say, under the compulsion of brute empirical fact, or, as I should prefer to say in less harsh terms, to be accepted with the 'natural piety' of the investigator. It admits no explanation" (1920, 46–47). Similarly, Broad (1925) and Lloyd Morgan (1926) insist that it is impossible to *predict* the emergent properties of a complex object on the mere basis of knowledge of its components and their properties, whereas it is possible to predict its resulting properties.

We can accept these theses of the impossibility of predicting and explaining emergent properties only if we interpret them in a new way. It is correct that we can neither predict nor explain the properties of a complex object if we know only the properties that its component parts possess in isolation from each other. However, contemporary science gives us no positive reason to think that — among chemical, biological, or psychological properties — there are emergent properties in the sense of Alexander, Broad, and Lloyd Morgan, impossible even in principle to predict and explain. The discovery of a reductive explanation, such as that of chemical bonding, of the mechanism of oxygen transport in blood or of the fixing of long-term memory refutes the idea that emergent properties cannot be explained.

Nevertheless, we can admit that the existence of emergent properties remains in a sense mysterious even if it is the subject of a reductive scientific explanation based on laws of composition. The sense of mystery stems from the fact that these laws are truths of fact rather than truths of reason. Since we *discover* the laws, in other words, because they cannot be known a priori, there is a sense in which we cannot understand why the laws have the forms that they do. When it comes to fundamental laws, we have to content ourselves with observing them without being able to hope to explain them. Ontologically, the laws of nature are necessary. However, they are epistemically a posteriori. This has the consequence that we can *conceive of* them being different from what they are. We can express this by saying that laws are "epistemically contingent."

I suggest that the persistent intuition of mystery surrounding emergent properties can be explained by the fact that their existence cannot be derived in a purely a priori way. For example, the emergence of a regular crystal from atoms that have no perceptible shape, the emergence of the shape of a flower from biological molecules, and above all the emergence of the qualitative experience of colours from the physiological mechanisms in our body continue to amaze us even when we possess a reductive explanation of these phenomena. The existence of these properties and phenomena remains partly mysterious in the same sense and for the same reason that the fundamental laws of nature remain mysterious.

Although I have just acknowledged that we can accept, in a sense, that the existence of emergent properties is partly beyond our comprehension, there is one big difference between this and classical emergentism: here emergent properties are not *absolutely* mysterious; the "natural piety," to use Samuel Alexander's (1920) expression, with which we are forced to accept their existence no longer needs to be unlimited as might have seemed to be inevitable before the advent of quantum mechanics, molecular biology, and cognitive neurophysiology. These sciences show how emergent properties can be reduced and explained. The only mystery that remains is the a posteriori nature of the laws of composition used in the reductive explanation. These laws, or at least the fundamental laws from which in turn they can be derived, always remain "brute nomological facts": even when the reductive explanation of a particular law of composition is discovered, that explanation is based ultimately on axiomatic laws that in turn cannot be explained. This observation is independent of whether the relevant law is more or less deeply integrated into a theory.[54] As a theory covering a certain domain of phenomena develops, the laws governing these phenomena change their status: what originally were "brute" empirical laws, hypothesized on the basis of induction from observation, become deducible laws within the framework of a theory. Thus, to discover an experimental law describing the behaviour of a chemical substance, at first it might be "absolutely necessary to study samples of *that particular compound*" (Broad 1925, 64). But the construction of a theory of the chemical behaviour of all compounds of a certain type can later make it possible to deduce (and in this sense explain and predict) this particular law from more general principles or laws and ultimately from the most general axioms and principles of the theory. Nevertheless, it always remains an empirical law whose explanation must end with principles and axioms that themselves are neither explicable nor predictable. For example, the law of composition that imposes the form (*) on the Hamiltonian of the molecular ion H_2^+ allows us to explain, up to a certain point, the properties of the molecule and, above all, the fact that it exists in a stable manner, thanks to the existence of a minimum of the overall energy at a certain distance R_0 between the protons. However, this explanation does not have the transparency of an a priori explanation. It necessarily starts from premises that contain an ineliminable empirical element: the particular form that the terms of the Hamiltonian take because of the interaction among the different components of the system. To distill the truth from Alexander's statement quoted above, we would have to say that emergent properties exist "under the compulsion of brute empirical law," where Alexander said "under the compulsion of brute empirical fact" (1920, 46), and instead of saying that their existence "admits no explanation" (47) it is more correct to say that it "admits no a priori explanation" or that it "admits only explanations that rest on ultimately inexplicable premises."

The concept of a "microbased" macroscopic property (henceforth MB property) developed by Kim might seem to provide a way of invalidating this result by conceiving of macroscopic properties — emergent or otherwise — as purely logical functions of underlying microscopic properties. Kim starts

54 We can put this idea parallel to the conception of explanation as resulting from the unification of the system of scientific knowledge: according to Kitcher, science explains by teaching us "*how to reduce the number of types of facts we have to accept as ultimate (or brute)*" (1989, 432).

from the concept of a "structural" property introduced by Armstrong.[55] A structural property of *s* is defined by the fact that the parts of *s* have certain properties and stand in certain relationships. Kim defines the concept of an MB property as follows: "*P* is a *micro-based property* just in case *P* is the property of being completely decomposable into nonoverlapping proper parts a_1, . . ., a_n, such that $P_1(a_1)$, $P_2(a_2)$, . . ., $P_n(a_n)$, and $R(a_1, \ldots, a_n)$" (1998, 84), where P_1 . . . P_n represent properties of the parts a_1 . . . a_n and *R* relations among these parts. This concept makes it possible to attribute to an object the property of having parts that in turn have mutual properties and relations. In this sense, it points the way to the development of the concept of a causally efficient macroproperty. However, the conditions that Kim imposes on MB properties are too weak to guarantee that the properties thus conceived are real or, in other words, causally efficient.[56] Most MB properties are not real properties in this sense (see Kistler 2005c, 149–50). Take a mereological sum whose elements do not interact with each other. The mereological whole composed of the electrons of billiard ball *A* and the nuclei of billiard ball *B* has neither the causal powers of a billiard ball nor any other causal powers. The existence of an MB property, conceived of in Kim's manner, is a *logical* consequence of the existence of the "parts" (of the mereological whole to which the MB properties are attributed), whereas the existence of a whole endowed with its own causal powers depends on the existence of appropriate interactions among the parts. Kim's definition does not impose any constraint on the relations *R* among the parts, and it does not require in particular that they are physical interactions. The subject who sees a yellow spot at a certain point in her visual field has an MB thanks to her neurons and their activation states. But even if these neurons and their activities are in their "normal" spatiotemporal relations (i.e., in conditions appropriate for producing the mental state of perceiving a yellow spot at that location), if they are prevented from interacting, then the perceptual experience disappears.

Another way of showing that the conditions that Kim imposes on MB properties do not guarantee that every MB is causally efficient is this: a given

---

55 "A property, *S*, is structural if and only if proper parts of particulars having *S* have some property or properties, *T* . . . not identical with *S*, and this state of affairs is, in part at least, constitutive of *S*" (Armstrong 1978, 2: 69).

56 "Real" properties are distinct from purely nominal properties that can be attributed to a system on a purely logical basis, based on the properties of its parts.

whole has a different MB for each possible decomposition. But many different decompositions do not give rise to an equivalent number of causally efficacious properties. Conversely, an object that has only one natural decomposition can have different causal powers. This is because of the interactions among the different properties of the parts. A hydrogen molecule H_2 whose most natural decomposition is into two H atoms nevertheless has several causal powers, such as its magnetic moment and a fundamental frequency of oscillation.

Unlike an MB property, an emergent property of an object is determined by a law of composition that is not purely logical and requires physical interactions among the parts. An emergent property is not identical to a structural or MB property because it is typically "multi-realizable" in two senses. First, different types of systems can possess it: the dynamic property of being disposed to undergo a phase transition can relate to the phase transition in the magnetization of an iron crystal or to the phase transition in water during freezing.[57] Second, there are many changes in the properties of the parts and their relationships that do not result in any change in the overall property of being disposed to undergo a phase transition. Many constellations of parts determine the same global property. This insensitivity of the overall property to variations in the determining properties of the parts characterizes robust systemic properties (see the previous section).

9. Avoiding Panpsychism

It seems to be difficult to deny that mental properties — as well as other properties usually taken to be non-physical, such as biological or chemical properties — are systemic. In other words, these properties appear only at a specific level: only a compound object of sufficient complexity can possess them but not its parts. Only a body more complex than an atom can be solid; only an organism can be adapted to its environment (in the sense of having fitness); only a cognitive system can represent its environment; only a human

57 Kim (1992a) shows that we can save the identity thesis in a situation of multi-realizability of this type by relativizing the global property to the different types of systems that might possess it. There would not be "the" phase transition but a different phase transition property for each type of system. However, if we consider that the identity of a property is determined by the laws in which it is involved, then we can justify the intuition that it is a single property common to different types of systems.

being can learn a systematic and creative language. To accept this idea is to adopt a doctrine that Girill (1976) calls "Democritean" in opposition to the "Empedoclean" doctrine. According to the latter, a micro-explanation of the possession of the macroproperty *P* by a system *s* must start from the fact that, among the components of *s*, some part already possesses, or some parts already possess, *P* (Klee 1984, 50). The Democritean conception, universally recognized today, admits that it is possible to explain the possession of a macroscopic property *P* of a system *s* in a way that does not presuppose that there are parts of *s* that possess *P*. For example, quantum mechanics explains the stability of molecules and the solidity of solid bodies by the stability of the bonds among atoms without attributing solidity to those atoms.

When we do not know the laws of composition that would explain a certain emergent property by a microreduction, we might be tempted to conclude that the Empedoclean doctrine is true. As far as mental properties are concerned, this is tantamount to accepting panpsychism. William James develops the argument in favour of panpsychism in this way: "*If evolution is to work smoothly, consciousness in some shape must have been present at the very origin of things.* Accordingly, we find that the more clear-sighted evolutionary philosophers are beginning to posit it there. Each atom of the nebula, they suppose, must have had an aboriginal atom of consciousness linked with it" (1890, 149).

Thomas Nagel (1979) more recently took up this argument for panpsychism, taking as premises the existence of mental properties and the impossibility of the only two conceivable ways of explaining their presence: they cannot be logically deduced from physical properties, and we have no Democritean explanation either. Therefore, we must conclude that panpsychism is true. More precisely, Nagel argues as follows.

1. Human beings are complex systems composed entirely of matter.

2. Mental properties are not logically implied by physical properties.

3. Human beings have mental properties.

4. There are no emergent properties. In other words, all properties of a complex system that are not relations

> between this system and something else derive from the properties of their constituents and their mode of combination.

From these premises, Nagel concludes that panpsychism is true: "The basic physical constituents of the universe have mental properties" (181). The validity of his argument depends on the interpretation of the word *derive* in premise (4), which states that all real properties of a system can be derived from the properties of its components and how they are combined. If the word *derive* means "logically deduce," then the argument is valid.

It is well known that what is a *modus ponens* for one is a *modus tollens* for another. If the conclusion of the argument (i.e., the truth of panpsychism) is taken to be unacceptable, then one of the premises must be false. I take premises (1) and (3) to be undeniable. Thus, either (2) or (4) is false, or both are false. In Chapter 2, I critically analyzed the doctrine of "cosmic hermeneutics," according to which it is possible to deduce a priori all of the properties of the universe (including psychological properties but with the notable exclusion of qualia) from the knowledge of physical properties alone. My refutation of cosmic hermeneutics comes down to justifying premise (2) of Nagel's argument.

Now this refutation of cosmic hermeneutics also shows that (4) is false if we understand the word *derive* to mean "deduce from logical principles alone." Conversely, if the word means "deduce, possibly by means of laws of composition," then (4) simply corresponds to the physicalist requirement that I imposed at the beginning of this chapter, among other conditions, on "weak" emergence. In this case, all premises are true, but the argument is no longer valid: if mental properties are derived nomologically (4), without being derivable logically (2), then humans can be composed solely of matter (1), while having mental properties (3), without the panpsychist conclusion being true (i.e., without humans' microscopic components having mental properties).

10. Response to a Version of Kripke's Argument against the Identity Theory

According to the conception developed in this book, mental properties are global properties that emerge from the properties of the parts of the body of their possessor and from the interactions among those parts. Global properties, and mental properties in particular, are determined by laws of composition. Since these laws are necessary, it is also necessary that any individual who possesses the physical configuration underlying a mental state possesses the mental state determined by the relevant law of composition.

Kripke (1972) developed a famous argument against the thesis of the identity of a phenomenal mental property, such as pain, with an underlying physical property. Kripke began with the thesis that identity statements expressed with rigid designators are necessary and then argued that the relation between phenomenal mental properties and underlying neurophysiological properties is really, not just apparently, contingent. Now we can construct an argument analogous to Kripke's that seems to refute my conception of the relation between physical and phenomenal properties.

According to my approach, pain is an emergent property of persons and animals, determined nomologically by the interaction of certain parts of their organisms according to their properties.

> (*) Pain (*S*) = the global property *G* of the organism, determined by the law of composition *L*, from the physical property *P* (an MB macroproperty of the organism).

To simplify the analysis, let us assume that pain is not multi-realizable. In other words, let us assume that only *P* gives rise to *S* by virtue of the law *L*.

Here is an argument analogous to Kripke's. (1) If the identity (*) is true, then it is necessarily true because "pain" and "*P*" are rigid designators. (2) (*) appears to be contingent, but (3) this apparent contingency cannot be explained by a confusion between metaphysical modality and epistemic modality. According to Kripke, such a confusion explains in particular that

> (**) heat = kinetic energy

seems to be contingent. Indeed, we confuse this identity of the properties themselves (which is necessarily true) with the contingent identity of

> (***) what appears to be hot to us = kinetic energy,

where "what appears to be hot to us" is a non-rigid designator that can designate different properties in different possible worlds.

For this explanation to work, (**) and (***) must be modally different. In the case of pain, the analogue of (***) is

> (*') what appears to us as pain = the property *G* that *P* determines as a function of *L*.

According to Kripke, there is no difference between (*) and (*') since pain is *essentially* what appears to us as pain.

The apparent contingency of (*) therefore cannot be explained by a confusion analogous to the confusion between "heat" and "what appears to us as heat." Therefore, (*) is really contingent; (*) is not necessary; therefore, (*) is not true.

The conception of mental properties developed in this book allows us to reply to this argument as follows: both (*) and (*') are necessarily true. But (*) and (*') do not appear to us as contingent: what appears to us as contingent is the fact that it is *P* and the law *L*, or (in Kripke's original presentation) the activation of *C* fibres, which determine *G* and therefore pain in us. This appearance is justified if *G* is multi-realizable. *G*, and therefore pain, are not identical to the microbased property *P* that determines them. If pain is multi-realizable, then different properties P_1, P_2, et cetera can determine nomologically the same global property *G*. If *G* is multi-realizable, then it is contingent that *G* emerges in the human species from P_1 and not from P_2. But even if *G* is not multi-realizable, it is not identical to the microbased macroproperty *P*, which determines it through the interaction of the microscopic components of the organism. If *G* is not multi-realizable, in other words if *G*, in terms of the laws of nature, can emerge only from *P*, then *P* gives rise to *G* in a necessary way. The appearance of contingency is then simply the appearance of the contingency of the laws of nature (see section 4.7).

11. Emergence, Reduction, and Supervenience

Reducibility is typically understood to be the opposite of emergence. What is not emergent is often called "reducible," whereas what is irreducible is often called "emergent."[58] This equivalence is plausible only if we give the concept of reduction a very narrow meaning. According to this "simple notion of reduc-

58 This conception is implicit in Wimsatt, who distinguishes between the situation in which "the phenomena of each [level] is [sic] explained by and reduced to those of the level below" and the

tion," reducing the property of a whole consists of explaining it by "studying the parts in isolation" (Holland 1998, 14) without taking into account their interactions. I have not found it appropriate to retain this meaning, which underlies expressions such as "this conception is reductive" in the sense of "it is simplistic" or "it constitutes an oversimplification." However, as soon as we construe reduction in terms of interactions among the parts of systems, emergence becomes compatible with reduction, whereby "*an emergent property is — roughly — a system property which is dependent upon the mode of organization of the system's parts*. This is compatible with reductionism" (Wimsatt 1996, S373).[59]

Whereas reduction is often (wrongly) taken to be incompatible with emergence because the constraints that it imposes on explanation seem to be too strong, supervenience is often considered a necessary but not sufficient condition for emergence. Davidson (1970) suggested analyzing the relationship between mind and body in terms of supervenience. Supervenience is a form of correlation between mental and physical properties that appears to be sufficiently weak to be compatible with the autonomy and irreducibility of psychology. At the same time, supervenience seems to justify the physicalist thesis that the mental depends on the physical, and that the physical determines the mental, whereas the mental does not determine the physical.

However, as Kim (1990, 1993a, 1997a, 189) has shown, the concept of supervenience is independent in fact of dependence and determination.[60]

situation in which a phenomenon "might not have an explanation and thus be emergent" (1976a, 252, 253).

59 On the distinction between the widespread concept of emergence according to which emergent properties are ipso facto irreducible, and the concept of reduction adopted here according to which emergence and reducibility are compatible, see also Wimsatt (1986) and Kistler (2007).

60 Charles (1992), Horgan (1993), and McLaughlin (1995) have also emphasized the weakness of the supervenience relation, which asserts only the systematic covariation of the properties in the two sets. Charles explains that supervenience entails neither explanatory priority nor the fact that the supervenience base is the ontological basis of the supervenient properties. He points out, for example, that in a deterministic world where *S* always has effect *T*, and where *T* can only be caused by *S*, all properties that supervene on *S* also supervene on *T*. This shows that supervenience does not allow one to identify "the appropriate basis for the occurrence of a given mental property" (Charles 1992, 275). As McLaughlin points out, even strong supervenience "does not imply *explanatory* connections between supervenient and subvenient properties" and that, "*if* reduction is an explanatory relation, then the SS_m [strong supervenience, as defined with the modal operator of necessity, in McLaughlin 1995, 25] of *A*-properties on *B*-properties with metaphysical necessity fails to suffice for reduction" (1995, 48). Humphreys draws the conclusion that "supervenience does not

Supervenience does not guarantee that mental properties depend on physical properties or that physical properties determine mental properties. On the contrary, it is compatible with parallelism or occasionalism: in these doctrines, mental and physical properties are determined by an independent cause, and depend only on it, namely the will of God.

This shows two things.

1. The existence of a universal correlation between mental properties and physical properties, and even a necessary correlation as in strong supervenience, contains no indication of the origin or explanation of this correlation (see Horgan 1984; Kim 1990, 26–27, 1998, 9–15; Horgan 1993, 577 ff.).

2. The thesis that mental properties supervene on physical properties does not guarantee a materialist position. As Horgan puts it, "the mere supervenience of higher-order properties and facts on physical properties and facts cannot be enough to confer materialistic respectability" (1993, 565). What would make the explanation of the relationship complete, while showing its materialist character, would be the demonstration of a relationship stronger than supervenience (Horgan calls it "superdupervenience"), which would "constitute a kind of ontic determination which . . . confers materialistic respectability on higher-order properties and facts" (566).

The concept of emergence in terms of laws of composition, themselves grounded in laws of interaction, aims at bridging that gap. It aims to identify the nature of the determination of the properties *P* of an object *s* by the properties of the components of *s* and their relations by virtue of non-causal laws of composition. This allows us to understand why the properties *P* arise from the properties of the components, whereas supervenience is usually seen as

provide any understanding of *ontological* relationships holding between levels. For that emergence is required" (1997b, S341).

a fundamental relationship that cannot be explained.[61] Emergence implies supervenience but goes far beyond it by providing an explanation of its origin.[62] Moreover, the thesis that mental properties are emergent is incompatible with parallelist and occasionalist doctrines.

This line of reasoning allows us to accept Van Cleve's thesis that emergence is "a species of supervenience" (1990, 222).[63] But my conception reverses the order of explanation. Van Cleve proposes, in the tradition of the contemporary philosophy of mind, to start from the relation of supervenience as a genus and then to define emergence as one of its species. He proposes using the type of modality as the specific difference: whereas the definition of supervenience contains a general operator of necessity, he proposes to conceive of emergence as the species of supervenience in which the necessity is nomological and not logical. "If *P* is a property of *w*, then *P* is emergent if *P* supervenes with nomological necessity, but *not* with logical necessity, on the properties of the parts of *w*" (222).

Van Cleve's proposal shares with mine the main idea of conceiving of emergence in terms of laws of nature, but his way of specifying this idea poses two problems (O'Connor 1994, 96–97). First, if we interpret Van Cleve's distinction ontologically (and not epistemically), it presupposes the widespread but controversial doctrine that the laws of nature themselves are contingent. On the ontological level, the only way to conceive of a difference between these two types of supervenience is in terms of modal strength. Nomological necessity is supposed to be weaker than logical necessity in the sense that it only

61 In Kim (1998), supervenience provides the starting point for philosophical reflection on the relationship between mental properties and underlying physical properties.

62 We could express the relationship between these concepts by inverting Blackburn's formula according to which "supervenience is physically fixed emergence" (1993, 233). According to the conception developed in this book, emergence is a form of supervenience that obeys physical constraints: emergent properties are determined by laws that do not necessarily remain "nomological danglers." In other words, these laws have the potential to become theorems, derivable from the laws of interaction that relate to the properties of the emergent basis. In this sense, all emergent properties obey physical constraints. However, the relation of emergence is stronger than the relation of supervenience; it implies it, whereas supervenience is compatible with irreducible bridge laws and even with brute extensional correlation in the absence of any law.

63 Van Cleve (1990, 224) cites Webster's unabridged dictionary of English as evidence of the current usage in 1960, according to which supervenience was considered synonymous with emergence. The entry reads thus: "Supervene 2. *Philos.* To occur otherwise than as an additive resultant; to occur in a manner not antecedently predictable; to accrue in the manner of what is evolutionally emergent."

constrains possible worlds that share the laws of the actual world. However, the latter — the worlds that share our actual laws — are a strict subset of all possible worlds (and the associated necessity is weaker than the logical necessity) only if the laws are contingent. But there are reasons to think, on the contrary, that laws are necessary because they determine the identities of properties (see Shoemaker 1980, 1998; Kistler 2002a, 2005a). However, if laws are necessary, then there is no difference between logical and nomological necessity in terms of modal force — in other words, in terms of the set of worlds within their scope. Nevertheless, it could be argued that there is an epistemic difference between nomological necessity and logical necessity even if they are ontologically equivalent: logical necessity is accessible a priori, whereas nomological necessity is known only a posteriori. However, this epistemic difference cannot be used to develop a conception in which the difference between emergent properties and resultant properties is *ontological*; the distinction between a priori and a posteriori is an epistemic difference that does not give the two forms of arising distinguished by Van Cleve a different modal force at the metaphysical level.

Second, if we disregard this first problem and accept the premise that nomological necessity is weaker than logical necessity, we are left with the problem that Van Cleve's criterion makes all systemic properties of complex objects emergent. As we have seen, no property of a system can be deduced a priori from the properties of its components. On the contrary, they are all determined by laws that can be known only a posteriori. Therefore, the necessity with which a compound possesses them, given its components, is always nomological and not logical.

I have made the determination of emergent properties of complex objects from the properties of their parts and their interactions a necessary condition for emergence. Such a determination leads to the emergent properties arising from the properties of the parts.[64] I have argued that the discovery of the law of composition that gives rise to an emergent property goes beyond supervenience in the sense that it provides a metaphysical explanation for it. Accordingly, one might be surprised by the thesis of Humphreys that there are systems with emergent properties that violate the principle of mereological

64 My conception of cognitive properties is incompatible with Bernal Velasquez's (2012) thesis that phenomenal consciousness can have causal powers of its own only if it *does not supervene on* physical properties.

supervenience. True, Humphreys (1996, 66; 1997a, 15–16) takes the only uncontroversial example of emergent properties in this sense to be that of the properties of entangled quantum systems. However, his conception of the "fusion" of instances of properties is supposed to apply equally to cases of emergence outside quantum properties. The result of the fusion of two instances of level *i* properties, P^i_m and P^i_n, possessed respectively by the objects x^i_r and x^i_s, at time t_1, is the instance of a new property at level *i+1*, represented by $[P_{im}{*}P_{in}]$, possessed at time t_1', after the fusion, by the object $x^i_r{+}x^i_s$ resulting from the fusion of the objects x^i_r and x^i_s. One can present this formally:

> $[P^i_m(x^i_r)(t_1){*}P^i_n(x^i_s)(t_1)] = [P^i_m{*}P^i_n][(x^i_r)+(x^i_s)](t_1')$ (Humphreys 1996, 60; 1997a, 9)

The time difference between t_1 and t_1' represents the fact that fusion takes time: the emergent property $[P^i_m{*}P^i_n]$ is instantiated at the instant t_1', later than the instant t_1 when the base properties P^i_m and P^i_n are instantiated by the parts of the complex object. This time lag is crucial for the conception of emergence proposed by Humphreys because it depends on the assumption that base properties *disappear* during fusion. At the instant t_1', when the fusion is completed, "the original property instances $P^i_m(x^i_r)(t_1)$, $P^i_n(x^i_s)(t_1)$ no longer exist as separate entities and they do not have all of their *i*-level causal powers available for use at the (*i+1*)st level. Some of them, so to speak, have been 'used up' in forming the fused property instance" (Humphreys 1997a, 10). The time lag between the instances of the base properties and the instance of the fused property justifies both the thesis of the novelty of the causal powers of the fused property (which in turn justifies the thesis of its emergent character) and the thesis of its non-supervenience. The instance of the fused property at t_1' has its own causal powers that are not identical to the causal powers possessed by the instances of the base properties, for the simple reason that, at time t_1', those instances of the base properties no longer exist. Given the time lag between the base property instances and the fused property, and the synchronic concept of mereological supervenience of the properties of a whole, at some instant, on the properties of its parts at the same instant, it is trivial that the fused property does not supervene on the base properties that gave rise to it.[65] Assuming that there are no other instances of the same base properties at the same place either, "trivially, there is nothing at t_1' at the i-level upon which $[P^i_m * P^i_n][(x^i_r)+(x^i_s)](t_1')$ can supervene" (Humphreys 1997a, 11).

By neglecting the distinction between the synchronic relationship of emergence and the diachronic change of properties, the conception of Humphreys trivializes emergence. In his view, the mere fact that a complex object undergoes a change that affects both its macroscopic properties and the microscopic properties of its parts is a sufficient reason to consider those macroscopic properties as emergent. According to this criterion, all macroscopic properties are emergent. Let us take a simple spatial movement. The properties of parts x^i_r and x^i_s of being at spatial positions P^i_m and P^i_n no longer have an instance at t_1'. (x^i_r and x^i_s no longer have them, and it is possible that no other object has them at t_1'.) Therefore, the property of the system being at the spatial position $[P^i_m * P^i_n]$ an instant later, at t_1', cannot supervene on any instances of P^i_m and P^i_n at t_1'.

It is true that typically (and probably always), for any emergent property possessed by an object *s*, there is an instant at which *s* begins to possess it. It is legitimate to ask for the causal process that led to this first instantiation, and it is often possible to discover it. But this question does not concern the synchronous relationship between the instance of the complex property at time t_1' and the properties possessed by the parts at that time, t_1'. To find out whether the global property *G* is emergent, we need to ask (among other things) whether the parts possess it at t_1' and whether the parts possess properties that determine it according to a law of composition. To find out whether *G* supervenes on properties *F* belonging to the parts of objects that are *G*, we need to ask whether necessarily, for any object *x* that is *G* at t_1', there exist

65 Humphreys (1997a) introduces his conception of emergence in terms of the fusion of instances of properties, in the context of the debate on the supervenience argument put forward by Kim (1998). I will return to this in Chapter 5. According to Kim, the only way in which one mental property (instance) can cause another is by causing its supervenience base. Mental causation therefore presupposes downward causation. Kim tries to refute the possibility of downward causation. The conception of Humphreys allows ($i+1$)-level (e.g., mental) properties to escape this argument and thus allows them to have causal powers. A property P_1 can cause another P_2 directly rather than by causing its supervenience base: in fact, P_2 has no synchronic supervenience base. But ($i+1$)-level properties can also have effects at level *i*: since P_1 does not have any synchronic supervenience base at level *i* either, its power to influence level *i* is not called into question by any such base property, whose efficacy would exclude that of P_1. See Humphreys (1997a, 14).

properties F such that the parts of x possess F at t_1' and such that any object that possesses parts with these properties F at any time t possesses G at t.

In the case of the stable state of the hydrogen ion H_2^+ that I considered earlier in this chapter, we can ask (which I explicitly have not done above) for the causal process of the fusion, over time, of two originally isolated atoms brought together until their electronic orbits overlap sufficiently for the energy levels determined by the system to have a minimum. If we consider the molecular electronic orbit that the system possesses at t_1', at the end of this process, it is trivial that it does not supervene on the electronic orbits that the isolated atoms possessed at some instant t_1 *preceding* the overlapping of the orbits: mereological supervenience — as it is imposed on emergence within the framework of physicalism — is a synchronic relation between the properties of an object and those of its parts, at the same instant. However, given that a change has taken place, after the formation of the molecule, the atomic parts of the molecule no longer possess the orbits of isolated atoms.

However, this is not enough to show that there are not other properties of these atomic parts on which the properties of the molecule supervene. If my analysis of the synchronous determination of energy states by the Hamiltonian is correct, then such properties of the molecule's parts exist: the electric charges of protons and electrons and the existence of a region of overlap of the electron orbits.

Similarly, in order to prove the novelty of a systemic property G at t_1', it is not sufficient to show that certain properties F of the parts no longer exist at t_1' and that the causal powers of G (at t_1') cannot be identical to those of F (at t_1) by consequence. Instead, it must be shown that there are no properties of the parts whose instances are synchronous with the instance of G and that have all of the causal powers of G. It follows from this analysis that Humphreys makes the mistake of concluding from the premise that there are F-properties of parts whose instances are not synchronous with the instance of G at t_1' that there are no F-properties of parts whose instances are synchronous with the instance of G at t_1'. As a result, he arrives at an extremely weak criterion of emergence according to which all systemic properties resulting from a change in the supervenience base count as emergent.

The thesis that properties at higher-level $i+1$ have no simultaneous supervenience base appears to be even more dubious in the case of mental and neurophysiological properties. At the instant t when a subject possesses mental property M, each of her neurons, synapses, and molecules possesses well-defined properties *P*. Here is a reason to contest that the *P* properties that give rise to *M* have been "used up" during the fusion that gave rise to *M*:[66] chemical and neuronal properties determine many global properties of the brain and the subject, some of which are aggregative or structural (i.e., have no causal powers of their own) and some of which are emergent (i.e., have causal powers of their own). However, if the fusion necessary for emergence made the supervenience base disappear, then the set *P* of neuronal properties could not give rise to any systemic property, not even an aggregative property. Let us suppose, *per impossibile*, that *P* gives rise to *M* as well as to an aggregative property M_1. M_1 depends for its existence on the properties *P* in its supervenience base. However, the existence of *M* requires the disappearance of *P*. It is therefore impossible for the base *P* to give rise to any aggregative property.

12. Conclusion

In light of the great scientific successes of the twentieth century, some have judged that the concept of emergence was doomed to become obsolete. According to the emergentist tradition — culminating with the work of C.D. Broad — emergence characterizes properties and nomological regularities whose existence cannot absolutely and definitively be the subject of a reductive explanation, even though these properties and regularities are the consequences of "trans-ordinal" laws that make the properties and laws of a given level depend on the properties and laws of lower levels. For emergentists, these trans-ordinal laws are absolutely inexplicable. The discovery of reductive explanations of certain chemical properties, by quantum mechanics, and of certain properties of heredity, by molecular biology, now lends credence to the conviction that there are no such inexplicable laws: the advent of the

66 This argument is due to Wong: "If all basal instances are exhausted in fusion, then structural properties and functions which depend on these will also be destroyed" (2006, 357). Wong does not solve the problem posed by Kim that I will discuss in Chapter 5: the efficacy of ($i+1$)-level properties is called into question by the underlying i-level properties. Instead of escaping overdetermination by maintaining, as Humphreys does, that the i-level properties no longer exist at the moment when the ($i+1$) properties come into existence, Wong says two contradictory things: on the one hand, there is no overdetermination "because the basal and supervenient properties are not distinct" (357–58); on the other hand, emergent properties have new causal powers: "If basal properties *don't* possess the causal powers of emergents, then they *can't* cause the same effects; *so they can't compete as overdeterminers*" (360). It is incoherent to say both that emergent properties are not distinct from their base properties and that they have different causal powers.

reductive explanation of all the properties and laws of the macrophysical and non-physical levels of reality, starting from the microphysical level, seems to be only a matter of time.

We do not need to decide the empirical question of whether there are properties and laws that will definitively resist reductive explanation. In any case, it seems to be prudent not to base any philosophical thesis on the existence of such "emergent" properties and laws in Broad's sense. However, the concept of emergence does not lose its usefulness if we assume that they do not exist, for it can be used to account for the existence of "levels of reality." Some structured systems possess properties that are "systemic" in the sense that none of their components can possess them and are "qualitatively new." These properties are the subject of nomological regularities: that is, regularities that are not accidental but due to laws that do not exist at the level of the components. Often a particular science is devoted to the properties and laws of a given level. Chemistry, for example, studies properties and laws that are emergent in relation to those studied by physics. Emergence characterizes the relationship of determination of a higher level, such as chemistry, with respect to lower levels, such as physics. The qualitative novelty of the properties and laws of a given level is independent of the discovery of their reduction. The fact that we explain the appearance of such a property or law does not make it any less qualitatively different. The qualitative difference justifies the idea that a set of properties, linked together by laws, constitutes a distinct level of reality.

The hope of accounting for the relationship between physical and mental properties in terms of supervenience has not been realized: supervenience has turned out to be too weak to justify the belief that the underlying base properties *determine* the supervenient properties and that the latter *depend* on the former. Supervenience expresses a form of systematic and nomological correlation, but it imposes no constraint on the origin of this correlation, which makes it compatible with dualist doctrines. Emergence fills the gap left open by supervenience. It characterizes the relationship of determination between properties and laws at adjacent, but qualitatively different, levels of reality. Within the framework of physicalism, we presuppose that each level is objectively determined by lower levels and ultimately by the physical level. The discovery of the nomological form of this determination can give rise to reductive explanations (as we saw in Chapters 1 and 3), the subject of empirical discovery (as we saw in Chapter 2). Among the properties thus determined, emergence characterizes those that are qualitatively different. For example, the solidity, transparency, and redness of a ruby emerge from the properties of the atoms that make up the crystal: only a structured crystal can possess these systemic properties, qualitatively different from the properties of its components. Similarly, only a cognitive system can form representations of its environment and learn to behave appropriately in it. No single neuron — or set of neurons or even the brain cut off from the rest of the body — can represent, learn to behave, or possess properties qualitatively equivalent to such properties of cognitive systems. In this sense, representation and learning are emergent from physiological and in particular neuronal properties.

The most difficult task is to find a rigorous criterion for qualitative novelty. I have put forward the hypothesis that the relevant concept of novelty can be characterized in terms of topological equivalence and other mathematical criteria. The structure of the representations of sensory qualities, such as colours, is topologically different from that of their physical stimuli. The mathematical rigour of such mathematical criteria makes them promising. However, we are far from having specified general criteria applicable to all properties and laws that intuitively are qualitatively new. The task of elaborating such criteria is in part scientific. Philosophy can contribute to this inquiry by giving emergence a place in the conceptual landscape of the problem of the relationship between levels of reality in general and the relationship between body and mind in particular.

What remains to be done is to defend the coherence of this conception of emergence against a major objection: in order to be real, an emergent property must have causal powers of its own. However, it is doubtful that emergent properties can have such powers given that the physicalist conception of the world seems to give the monopoly of causal efficacy to physical properties underlying them. The next chapter is devoted to examining this objection.