

# The Chemical Core of Chemistry I

A Conceptual Approach\*

*Joachim Schummer*

**Abstract:** Given the rich diversity of research fields usually ascribed to chemistry in a broad sense, the present paper tries to dig out characteristic parts of chemistry that can be conceptually distinguished from interdisciplinary, applied, and specialized subfields of chemistry, and that may be called chemistry in a very narrow sense, or 'the chemical core of chemistry'. Unlike historical, ontological, and 'anti-reductive' approaches, I use a conceptual approach together with some methodological implications that allow to develop step by step a kind of cognitive architecture for chemistry, which basically entails: (1) systematic chemical knowledge on the experimental level; (2) clarification of chemical species; (3) chemical classification systems; (4) theoretical foundation through the chemical theory of structural formulas. In a succeeding paper the results will be checked for resisting physicalistic reduction.

**Keywords:** *chemical properties, logical structure of chemical knowledge, chemical substances, chemical classification, theory of structural formulas.*

## Introduction

Unlike particular sciences, philosophy is concerned with very general aspects of our world, and as such always running the risk of producing either over-generalizations or one-sided generalizations. The comparatively small group of present philosophers of science are faced with the strange situation that their reflections on science should in a way comprise all the activities of the past and present community of scientists, which is roughly 10,000 times bigger. The situation is even much worse in the philosophy of chemistry. Not only do chemists form the largest group among scientists, while philosophers of chemistry form the smallest and youngest group among philosophers of science. The wealth of research fields ascribed to chemistry is extraordinarily diverse too.<sup>1</sup> To talk of 'chemistry' as a somehow united field seems to obscure the plurality of historical traditions, methods, and scientific aims of this

field,<sup>2</sup> as well as the varieties of interdisciplinary projects chemists are and have been working on.

Nonetheless, whenever we talk about chemistry in general, compare it with or distinguish it from other sciences, and, in particular, whenever we teach chemistry, then we presuppose some narrow notion of chemistry, that neither comprises the entire list of research activities ascribed to chemistry, nor does it simply refer to a historically contingent situation. It need not be a clear-cut definition, but at least we presuppose some general ideas about what is more essential or peculiar to chemistry. Here philosophy of chemistry can and should help clarify or elaborate such a notion with philosophical means, keeping in mind the mentioned risks. Moreover, a narrow notion of chemistry with regard to its peculiarities may help understand the specific chemical side of all kinds of interdisciplinary research, in which a lot of chemists are involved. It may further help avoid blind reductive attitudes, based on excessive or one-sided generalizations, both with regard to biology and physics. And finally it may draw the attention to philosophical problems of chemistry *proper*, beyond the interfield problems of reduction and how mutually stimulating interdisciplinary research is possible.

The present paper tries to elaborate such a notion of chemistry *in a very narrow sense*, so that I prefer to speak of the ‘chemical core of chemistry’. Unlike historical approaches, I will not refer to the history of chemistry definitions, chemical ideas, theories, research schools, disciplines, discourses, etc.<sup>3</sup> Unlike ontological approaches, I will not refer to a certain ‘chemical level’ within a presupposed ontological hierarchy.<sup>4</sup> And unlike ‘anti-reductive’ approaches, I will at first not refer to what is (epistemologically) irreducible to microphysics;<sup>5</sup> however the results of the present part will be subject to such an ‘irreducibility test’ in a succeeding part. Instead, the present approach is basically a conceptual one, by which I try to develop the cognitive structure of an autonomous science step by step: from systematical experimental investigations (1), over the clarification of basic ontological species (2), and the formation of classification systems (3), to a theoretical foundation with highly systematizing, predictive, and explanatory capacities (4). In order to give substance to that methodological skeleton, a few basic conceptual decisions are necessary. The most important decision, which considerably narrows down the notion of chemical core, is that *chemical* properties – in contrast to other material properties – form the starting point (Sect. 1.2). All the following steps basically depend on the peculiar relational structure of chemical properties, which determines the logical structure of chemical knowledge on all cognitive levels.

Of course, the conceptual approach will not reproduce the historically grown distribution of weight given to current research activities by chemists. In particular, chemists working at the frontier of quantum chemistry will miss most of their work in what I am going to call the ‘chemical core of theo-

retical chemistry'. The reason is simply that quantum chemistry is placed in the interdisciplinary area between chemistry and physics and is, like other interdisciplinary fields, not directly concerned with the 'chemical core', as the succeeding part will prove in more detail. Hence, when I am not going to credit the many achievements of quantum chemistry in this paper, that does not mean that I do not acknowledge them, as a commentator of a former version once suggested; it is just because these are outside of the present paper's systematical focus. Moreover, the technical term 'chemical core' does not imply any evaluation on whatever value basis, let alone a devaluation of research outside of the 'chemical core'. It simply denotes parts of chemistry that can be conceptually well distinguished from interdisciplinary, applied, and specialized subfields of chemistry, and at the same time resist physicalistic reduction, as the succeeding part will show.

## 1. Basic decisions: material aspects, chemical properties, and the structure of chemical knowledge

### 1.1 Material aspects

Chemistry, like many other natural sciences and non-scientific activities, is first of all concerned with empirical objects. But its specific focus is on *material aspects* of these objects. What that means, is getting clearer if we first regard what it not means. For material aspects are only part of a rich diversity of aspects that we apply to look at empirical objects.

To take an unproblematic instance of an empirical object, regard a coin. Within a certain economical society, it has a fixed exchange value, which essentially qualifies that empirical object for being a coin. The exchange value is an interesting functional property, that is neither a material property nor does it basically depend on the coin's material. If the coin is antique, our focus may be rather on the past of this little object, and we might wonder about its long history from an antique empire towards our hands. A closer look may reveal interesting signs on both sides of the coin, written in a foreign language, or a portrait of an emperor, or an emblem, that we like to interpret. If you are a numismatist, you might collect coins according to certain motifs, and you might have certain aesthetical preferences for one or the other motif and its specific artistic representation. However, if you are interested only in the material of that coin, then you abstract from, *i.e.* you ignore, all these economical, historical, semiotical, aesthetical, *etc.* properties of empirical objects.

Furthermore, our material focus analogously abstracts from the given size and form of empirical objects. Any representative sample would be sufficient to carry out material investigations. From the numismatic point of view, sampling a coin would be an ignorant destruction of what is essential to that object, whereas from the material point of view, the object remains unchanged. What counts is completely independent of geometrical form, structure, size, number of parts, weight, mass, space-time coordinates or location and movement in space, as long as the object is appropriate for material investigations. It is by no means pure chance that the mentioned list of properties corresponds to what philosophers have called 'primary qualities' since the 17th century. The material focus ignores 'primary qualities', because what matters from the material point of view is just what is invariant to changes of 'primary qualities'. Such as mechanistic philosophers have ignored material properties, because what mattered from their point of view is just what is invariant to changes of material properties. Thus, these two perspectives on empirical objects are in a sense orthogonal.

The mechanistic search for invariants with regard to differences in material properties was, in a certain sense, the search for the 'essence of matter', something most general that remains constant in the course of all specific material changes. Sciences of materials, on the other hand, have no ambition for such metaphysical (over)generalizations. In contrast, they seek for a subtly sophisticated system of material concepts, in order to describe *the diversity of material phenomena* as precisely and unambiguously as possible. A set of material concepts is a *system of classification*, if every concept allows at least a binary discrimination of material phenomena and if all concepts are logically independent of each other. Such a classification is not (and cannot be) deductively inferred from the 'dematerialized essence of matter'. Instead, it is (and must be) developed from some primitive material concepts step by step through concept differentiation and introduction, and through empirical check of its actual discriminating power.

Chemistry is the most general science of materials, in the sense that it provides the most general system of concepts. Unlike for instance mineralogy, metallurgy, pharmacy, and the wealth of applied subdisciplines of chemistry (such as chemistry of polymers, or of ceramic, magnetic, electronic, photonic, *etc.* materials), the concepts of (general) chemistry are applicable to and discriminating with regard to all empirical objects.

In the next section we will have a closer look at current material properties among which chemical properties play an eminent role to found a sophisticated material classification.

## 1.2 A typology of scientific material properties

The nature of scientific material properties gets clear only if we widen our everyday understanding. Philosophically speaking, we must give up phenomenalism, the epistemology of everyday life that does not question our ordinary contexts of experience. For sciences of materials, with chemistry at the center, have been, from the earliest stages on, experimental science in the original meaning of studying the behavior of objects in various and controlled artificial contexts. *A material property is reproducible behavior within certain reproducible contextual conditions.* It is important to note that material properties are attributed not to isolated objects but to objects *and* contexts. Since everything looks red under red light, we have to specify the color both of the object under investigation and of the light, in order to make qualified color statements. Since everything is solid at a certain temperature and pressure,<sup>6</sup> solidness always implies specification of thermodynamic conditions. Sometimes it is more the context that matters. To speak of a toxic substance, does not mean that the substance itself but the context, a biological organism, falls sick or dies, if it gets in contact with the substance. Precise material predicates require precise and systematic details of the contexts of investigation, making contexts themselves a central subject matter of sciences of materials.

The context is also the central aspect according to which material properties are distinguished, each type being characterized by focussing on a certain contextual factor:<sup>7</sup>

1. mechanical forces: *mechanical properties* (like elasticity, viscosity);
2. thermodynamic conditions (temperature, hydrostatic pressure): *thermodynamic properties* (like specific heat capacity, melting point);
3. electromagnetic fields: *electromagnetic properties* (like specific magnetic susceptibility, specific electric conductivity, optical absorption coefficient);
4. other chemical substances: *chemical properties* (like the capacity for oxidation or the solubility in a certain liquid);
5. biological organisms: *biological or biochemical properties* (like LD<sub>50</sub>, antibiotic or anaesthetic effect);
6. ecological systems: *ecological properties* (like ozone depletion potential [ODP], green house effect factor).

Since every experimental context can be described in terms of each factor, it is necessary to introduce standard or neutral conditions for each (*e.g.*, restricting mechanical forces to gravitation and sometimes stirring; minimizing electromagnetic fields; working with standard pressure and temperature, inert container materials, abiotic and closed systems). If we vary only one factor and keep the others standardized, then we are investigating the correspon-

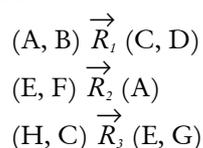
dent type of material property. If two or more factors of interest are combined, one can create new types of ‘mixed’ material properties (photochemical, thermo-electrical, thermo-electro-chemical *etc.*).

Properties of type (1)-(3) are called *physical material properties*, covering a great part of physical chemistry.<sup>8</sup> Physical properties are kept distinct from chemical properties by ‘excluding the chemical factor’, *i.e.* by working with inert container materials and atmospheres. But we speak of *chemical properties*, if and only if the ‘chemical factor’ is considered to be relevant for the behavior.<sup>9</sup>

The typology of material properties allows us to make the first, and the most important, conceptual distinction regarding our main question what we should consider the chemical core of chemistry. To be sure, all types of investigations are performed by chemists, including physico-chemists, bio-chemists, ecological chemists, and so on. All these studies have their own rights and are indispensable contributions to our overall knowledge about the material world. However, if we want to make a *conceptual decision*, what kind of investigation is central to chemistry, there seems to be no doubt that the *investigations of chemical properties form the chemical core of experimental chemistry*.

### 1.3 Chemical properties and the logical structure of chemical knowledge

While nonreactivities are also important chemical properties,<sup>10</sup> chemists are particularly interested in chemical properties that include chemical reaction behavior. The concept of *chemical reaction* can be roughly defined as a change of *chemical identity* according to the classification of chemical substances.<sup>11</sup> From the logical point of view chemical reaction properties show extraordinary features. They are asymmetrical *dynamical relations* with two classes of *relata*: initial chemical substances before the change and different chemical substances afterwards. Thus, chemical experiments yield complex dynamic relations connecting several chemical substances with each other. Take for instance:



(A, B, ... H: chemical substances;  $\xrightarrow{R_1}$  -  $\xrightarrow{R_3}$ : chemical relations including certain conditions according to the contextual factors above).

We get more insight into the logical structure of chemical knowledge, when we systematically combine all chemical relations into a complex net-

work that connects all chemical substances with each other in many direct or indirect ways (Fig. 1).

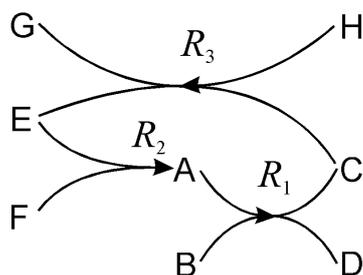


Figure 1. Schematic sketch of the chemical network.

It is first of all the literary tradition of linearly recording chemical knowledge in books and papers, which has been obscuring the *nonlinear network structure* that naturally comes out by connecting chemical facts systematically. Chemistry at the core is a science of peculiar relations. Instead of studying isolated objects to be measured, compared and put into a classificatory scheme, dynamic relations between objects constitute the basic set of chemical knowledge, and, at the same time, provide the grounds for the classifications of the objects themselves, as we will see later (Sect. 3). The chemical network includes all current empirical information about transformations of chemical substances, and new relations and *relata* supplement it every day. Starting from any chemical substance X we can follow the arrays in both directions providing two different types of chemical information: First we know how to produce X from certain different chemical substances, and second we know how and what kind of different chemical substances can be produced from X.

Concerning our conceptual search for the chemical core of chemistry, we have found another important result. If the investigation of chemical properties is the core field of experimental chemistry, as we have suggested in the last section, then we must conclude that the logical structure of systematical chemical knowledge is a network structure. In other words, *the chemical network, with chemical substances as the nodes and chemical relations as the connections, forms the chemical core of experimental chemistry.*

In the succeeding paper we will see that reductive accounts of physicalism mainly fail, because they follow an ‘isolated-object-approach’ that does not provide the relational kind of knowledge specific to chemistry.

## 2. Intermezzo: the problem of pure substances

In the preceding section we have assumed that chemical substances, the nodes of the chemical network, are entities that everybody, not only chemists, has at least some familiarity with. The usual and frequent talk of ‘chemical substances’ or ‘pure substances’ (both considered as synonymous in the present paper) gives the impression, that also the *concept* of ‘pure substance’ has a simple and well-defined meaning. However, it will turn out in this section that the concept is actually very intricate, despite its central importance. Let us first regard, if there are some *empirical* properties to make a general distinction between pure substances and others.

### 2.1. Empirical approaches

Take a couple of *unknown* homogeneous samples and try to find out which one is pure and which one is mixed by investigating empirical properties only. Simple physical material properties do not allow us to make a decision. Whatever the resulting values of physical properties such as melting point, refraction index, viscosity, *etc.* are, we do not get the slightest idea from such values, whether the sample is pure or not. For all types of properties the values of pure substances vary in the same range as the values of mixtures. Furthermore, for many electromagnetic and mechanical properties it is even true that the values of different modifications and crystal forms of the same chemical substance vary in the same range as the values of mixtures. Thus, these properties do not even tell us, whether two samples belong to the same chemical substance or not.

The same even goes for sophisticated electromagnetic properties provided by spectroscopic means. Spectroscopy, today’s main method to *prove* pureness of *known* substances, does not give a definite answer, whether an *unknown* sample is a pure substance or a not. Only if we have gathered sufficient background information about the sample (*e.g.* from its history) to make predictions about how the tentative pure substance and its possible impurities may spectroscopically look like, are we able to make some decision. But that is not what we are looking for, namely a general method to distinguish between completely *unknown* samples, which may also include unknown substances.

The chemical approach is by no means better. Pure and impure samples do not chemically differ in a way that allows us to draw general conclusions. We cannot claim, for instance, that pure substances differ in qualitative terms whereas homogeneous mixtures differ only in quantitative terms. Sometimes two distinct chemical substances, *e.g.* two stereoisomers, even react to form the same products under certain conditions. And two mixtures of the same

substances but of different proportions may yield quite different products by chemical reaction, as the famous law of multiple proportions already stated.

Thus far, our attempts to make out the distinction between chemical substances and homogeneous mixtures by means of material properties are rather unsatisfactory. Given the importance of that distinction to any chemical practice, concept building, and theorizing, such a negative result is surprising. Before we turn to the question how chemical substances are produced by purification, we will first have a brief look at theoretical approaches.

## 2.2 Theoretical approaches

Introductory textbooks of chemistry address our problem, if at all, by using a molecular approach. A sample of a pure chemical substance, it is said in the introduction, consists of molecules of a single sort, whereas homogeneous mixtures consist of molecules of different sorts. Pure water, today even a standard example of analytical philosophers,<sup>12</sup> consists only of  $\text{H}_2\text{O}$  molecules. Underlining the conceptual precision of that approach, we are told that 18 g water consists of  $6.22045 \cdot 10^{23}$   $\text{H}_2\text{O}$  molecules.

If the textbook has some affinity to atomic physics, the statement is very soon modified by admitting that the  $\text{H}_2\text{O}$  molecules are, in a strict sense, not of the same sort. Instead they are built up of different isotopes of hydrogen and oxygen in various combinations. But such a difference is said to be unimportant to general chemical problems. A few chapters later, when the first chemical problems are faced, students are confused by reading that the same 'pure' water now consists not only of  $\text{H}_2\text{O}$  molecules, but also of  $\text{H}_3\text{O}^+$ -ions and  $\text{OH}^-$ -ions, and that the concentration of all three components vary with temperature and pressure. In such a context of basic acid-base theory, we learn that the concentration of  $\text{H}_3\text{O}^+$ -ions, small as it is compared with the  $\text{H}_2\text{O}$  concentration, substantially governs the chemical properties of water. Another chapter that deals with electrical properties of the same pure water tells us a dynamical story of various forms of aggregation and disaggregation giving the impression that the  $\text{H}_2\text{O}$  molecules have nearly disappeared. The simple picture of countable building blocks of a single sort has been gradually changed into a chaotic mass that is said to be governed by various forms of changing interactions, such as covalent bonding, hydrogen bonding, van der Waals interactions, or electrostatic and spin-spin interaction, depending on the level of description. If the student reaches the level of quantum mechanics, she learns that different entities of a single sort are no longer distinguishable and should be dealt with as a single whole entity with mysterious dimensions. She is told that the building block picture must, *strictu sensu*, give way to a holistic approach, and that only mathematical complexity forces one to

use sometimes several simplistic quasi-building block approximations that are adapted to one or the other problem.

The conclusion we can draw from this little story is that the molecular approach of pure substances as consisting of building blocks of a single sort, does not work as a general criterion, not even in the light of classical chemistry, *e.g.* acid-base theory or electrochemistry. *There is no simple one-to-one relation between chemical substances and quasi-molecular species.* If the theoretical criterion were taken seriously, we would rule out a lot, if not all, of what we usually consider to be pure substances. Moreover, there is not a single molecular approach but a variety of different theoretical approaches adapted (and restricted) to one or the other problem. For many problems of today's chemistry, *e.g.* of complex chemistry, solid state chemistry, chemistry of transition states, supramolecular chemistry, *etc.*, the view of static building blocks that correspond to pure substances is inappropriate. While their experimental and conceptual starting points are still pure substances, their theoretical concepts refer to quasi-molecular species that do not correspond to pure substances any longer.

Thus, we are faced with the strange situation that the concept of pure substances, though central for all chemistry, seems to evade both empirical and theoretical determinations.

### 2.3. Making pure substances: operational criterion

Pure substances are produced by purification methods. Purification is an artificial intervention into the samples we have first at hand. That means, pureness is not a natural category, pure substances are no natural kinds in the simple sense that the material world as such would be divided up into pure substances. Nor is the material world divided up by simply applying our mental concepts, by putting our classificatory grid onto the world. In chemistry, there seems to be no more need to involve into the old philosophical struggle between nominalism and (natural kind) realism. Instead, our material world is such that it can be *experimentally* divided up into pure substances by certain means – that is the chemical way to solve the philosophical puzzle, from the philosophical point of view it is a kind of operational solution. A material sample is called pure, if any further efforts of purification, according to operations and properties that we address soon, have no more recognizable results. Since we know that our means to recognize the effects of purification have limited resolution, the concept of pureness may be extrapolated into the ideal realm of unlimited resolution. The ideal pure substance is hypothesized as passing an ideal purification test with ideally unlimited resolution.

The basic means of purification are, at least by the final step, various thermal processes all including one or the other *phase transition*, such as

distillation, crystallization, sublimation, *etc.*<sup>13</sup> Though there are today a wealth of other separation methods, the classical thermal processing has still some methodological priority. Chromatography, today's main purification processing in chemistry laboratories, does not work as a general and independent pureness test. For we cannot claim pureness of chromatographic fractions of an *unknown* sample, because we do not know if we have actually chosen a pair of effective chromatographic phases appropriate to the mixture to be separated. That is to say, that chromatography works in practice only if we have some background information about what we are going to separate. Thus it does not provide a general criterion of pureness.

*A material is usually regarded as a pure substance, if the temperature or any other material property remains constant in the course of further phase transition.* But such a sophisticated property does not work as blind operational criterion, since it turns out that, for instance, azeotropic and enantiomer mixtures, though showing constant transition temperatures, can nonetheless be separated further under modified conditions. Mixtures like those are handled either by modifying thermodynamic conditions or by performing other appropriate separation steps before. But finally, at least whenever we want to be sure and do not have enough background information, the sample has to pass the thermal purification test.

Against the background of the discussed failures of empirical and theoretical approaches to define pure substances, the operational criterion we found deserves special attention. If it is true that phase transition is a necessary part of the decisive processing, then we can draw some conclusions about the concept of pure substances. First, we understand now why substances such as polymers, that do not undergo definite phase transitions, are ruled out by the concept. Second, a pure substance is something that persists during phase transition; that is why modifications and crystal forms, though differing in physical properties, are considered to be the *same* chemical substance. Third, phase transition involves adding or taking away certain amounts of energy, *i.e.* a pure substance is something that persists through variation of thermodynamic conditions within certain limits. From the theoretical point of view it becomes clear now, that only molecular entities of a certain energetic stability can be related to pure substances, if at all. Fourth, a single molecular entity can only indirectly *represent* a pure substance, but is by no means itself a pure substance. For phase transitions are theoretical conceived as a change of interactions between molecular entities of a phase and not as being a property of a single molecular entity.

#### 2.4 The chemical need of substances as basic chemical species

The operationally based concept of pure substance turns out to be rather intricate. And what is more, it seems to be at odds with theoretical accounts.

There is no wonder that many theoretical and analytical chemists today feel uncomfortable with the restrictive concept of pure substances as defining the basic chemical species. From their point of view the concept is old-fashioned and bears some arbitrariness of selecting only those material entities that can be purified and put into bottles. They prefer to define chemical species through one or the other quasi-molecular approach supported by spectroscopic evidence. From such a viewpoint quasi-molecular species are identified independently of their environment. As a consequence, the distinction between pure substances and homogeneous mixtures does not make sense any longer for them. In fact, the distinction is, as we have pointed out, not a natural category. It is imposed onto the material world by our experimental approach. To be sure, the distinction comes out naturally, so to speak, if we apply our purification processing. But we could do else, or, at least, we need not attach too much conceptual importance to it. Berthollet's original approach, which gave up the distinction between pure substances and solutions, would actually be more appropriate and consistent with modern theoretical chemistry.

The reason why we still adhere to that distinction finally leads us back to what I have called the chemical core of chemistry. For the distinction is essential to the chemical approach of characterizing and classifying materials – and only to that approach as we will see now.

From the classificatory point of view, homogeneous mixtures could be perfectly distinguished and classified by an appropriate set of quantitative physical properties. With the help of high-resolution instruments we are able to determine the slightest difference in concentration according to any physical property. Once we have built up an appropriate database, identification of any homogeneous material is routine work. Thus, there seems to be no *classificatory* need at all to refer to such idealizations as perfectly pure and distinct chemical substances. We need not make the detour of describing homogeneous materials in terms of concentrations of pure substances, something that may come out after violent purification. We could go directly into *medias res*, towards the various quasi-molecular species that tumble around both in homogeneous mixtures and pure substances, proving again that the distinction is artificial, arbitrary and useless.

To be sure, the distinction between pure substances and homogeneous mixtures is artificial in the sense that it depends on our experimental interventions. But it is neither useless nor arbitrary, as long as we do not have an alternative. By then the distinction is the only one we have that allows us to single out from the continuous variety of homogeneous materials perfectly distinct substances. And that is exactly what we need for our chemical properties, the complex relations that we have pointed out above. *We need distinct substances as definite starting and end points of chemical reactions, as relata of the chemical relations.* And we need them even more for connecting the rela-

tions together in order to build up the chemical network, the logical structure of chemical knowledge. The chemical approach of material characterization and classification depends on distinct entities at the same ontological level that is to be classified. It is essentially a requirement of concept precision and systematization that forces chemists towards ontology of distinct entities. There is no way, at least we do not yet know any, to define a systematic set of chemical properties within the continuous realm of homogeneous mixtures, without referring to pure substances as distinct entities. Of course, we could describe any chemical reaction as a change of homogeneous mixtures, *i.e.* as a change of values of any set of quantitative properties. But we would not be able to put such descriptions into a systematical context. All we could do is collecting infinitely many facts, each of them highly precise but without any systematic connection with each other.

In sum, *the concept of pure substances, artificial as it may appear, perfectly meets the chemical requirement – and only that – of building a systematic structure of chemical knowledge.*

## 2.5 Can we do the same with quasi-molecular species?

A last objection against the pure substance approach remains. Our argument for pure substances was in essence a logical argument that refers to the logical structure of chemical knowledge: we need distinct *relata* for the chemical relations in order to build precise concepts and to connect our experimental results in a systematical way to form a network structure. Couldn't we build up an analogous structure in terms of quasi-molecular species instead of pure substances? Modern spectroscopy enables us to detect quasi-molecular species at very low concentrations persisting only for picoseconds. With the help of sophisticated theory we are able to interpret the data in terms of existence and change of such structurally specified entities.<sup>14</sup> Thus, what comes out has the same logical structure that we found for chemical properties of pure substances, *i.e.* relations of the type, say,  $A+B \rightarrow C+D$ . The only difference is that the letters A, B, C, D now refer to quasi-molecular species instead of pure substances. As we have seen in Sect 2.2, there is no one-to-one relation between pure substance and quasi-molecular species. Instead, we can sometimes detect a wealth of quasi-molecular species within a sample of a pure substance, and we even know that they are changing under varying conditions. Moreover, many quasi-molecular species, such as van der Waals complexes, appear only in homogeneous mixtures, so that we are unable to put them in purified form into bottles. Thus, the quasi-molecule approach appears to be a revolutionary shift in chemistry, bringing about a much greater diversity of chemical species and changes than the coarse stuff approach. Once we drop the distinction between pure substances and homogeneous mixtures and take quasi-molecular species as the basic species, chemistry be-

comes considerably richer. There is no doubt that many chemists have been working on that since decades, giving us more and more insight into the complexity of even the simplest pure substance. So what are the reasons to adhere to pure substance as the basic chemical species any longer?

There is a practical reason of great importance. Our ontology of chemical substances ensures per definition that certain kinds of material processing do not change the identity of substances. Chemical substances are *per definition* invariant to purification processing, change of thermodynamic state, sampling, *etc.* What seems to be, at first glance, an arbitrary and coarse approach, turns out to be a necessary requirement for systematical chemical research. First, it is purification that enables us to start a chemical reaction with a definite number of chemical species, *i.e.* to have controlled and comprehensible chemical conditions. And secondly, it is the invariance to purification processing that allows us to isolate definite chemical species as reaction products afterwards. As a consequence, the resulting chemical species can then, each of their own, be the starting point of further chemical investigations under new controlled conditions, *i.e.* with other combinations of chemical species. Thus, it is only because our chemical species *per definition* retain their identity during purification, that we are able to connect single facts of chemical relations with each other to build a systematic network structure of chemical knowledge.

Do we have an equivalent within our conceived ontology of quasi-molecular species? Every processing, every change of thermodynamic and other conditions, in particular, every attempt of isolation, may change the quasi-molecular species. It is the burden of precision and the lack of operationally well-adapted coarseness that leaves no room for any processing which these species are definitely invariant to. There may be singular cases in which quasi-molecular species do not change a lot in the course of stuff purification. But we know from spectroscopy that many do. The geometrical structure of quasi-molecular entities, *i.e.* distances and angles, change with changing temperature, let alone phase transitions. If structure is the essential property of a quasi-molecular entity, then we are ontologically committed to consider all that as changes of quasi-molecular identity.<sup>15</sup> Even if we could pick up a single quasi-molecular entity from a certain quasi-molecular surrounding and put it into another one: why not considering that a change of identity too? At least we know that the solvent surrounding has considerable impact on the structure of a quasi-molecular solute. Some quantum chemists even think that is only the impact of the surrounding what constitutes a molecular structure at all (Amann 1993).

Given the lack of processing that do not change quasi-molecular species, *systematical* chemical research gets into serious trouble. Once we have started investigating the changes of quasi-molecular species of a *real* system, *i.e.* not of a simplified computer simulation, the system is getting more and more

complex without return. We cannot restore controlled conditions, since any intervention is itself a change of species. In particular, we cannot systematically connect quasi-molecular species from different systems, so that the outcome of one reaction would be the starting point of another, as we can do with chemical substances. Such connections may be, in singular cases, possible, if we spectroscopically recognize quasi-molecular species known from other systems. But we cannot do that systematically; *i.e.* the relations between quasi-molecular species cannot be connected in a way to yield a comprehensive network structure of chemical knowledge. Thus, the costs of a richer ontology are a fragmentation of chemical knowledge. There is no doubt that a quasi-molecular ontology is extremely helpful for a better understanding of material systems of their own including chemical changes.<sup>16</sup> However such an ontology is inappropriate for the logical structure of chemical knowledge. That is the main reason, why a *chemistry of quasi-molecular species cannot replace but only support the chemistry of pure substances*. Hence, we have good reasons, to consider pure substances, despite the mentioned problems, as the basic chemical species.

## 2.6 Conclusion

The concept of pure substances has turned out to be rather complicated, despite its fundamental role in chemistry. While spectroscopic detection of *known* instances is routine work, pure substances do not reveal any simple physical or chemical characteristics to distinguish them from ‘impures’. Nor do we have any theoretical account that tells us something about the general nature of pure substances. Instead it is only what we call ‘purification processing’, that allow us both to make pure substances and to give a general definition in operational terms.

We have pointed out, that the distinction between pure substances and homogeneous mixtures is in some sense artificial and may be regarded as arbitrary and old-fashioned from a certain point of view. If it has a justification, then it is only within the framework of the ‘chemical core of chemistry’, because it perfectly meets requirements of the logical structure of chemical knowledge. For it selects distinct substances from among a continuous diversity, something that we need as *relata* of chemical relations and, finally, as nodes of the chemical network. Since building an analogous network of quasi-molecular species is impossible – and will probably remain so forever –, quasi-molecular species are no general alternatives in chemistry. In sum, the choice for pure substances as the basic chemical species is grounded on no other good reasons than chemical core reasons.

### 3. Chemical classification

A long-standing tradition of philosophy of science, mainly fascinated by the mathematical elegance of Newtonian mechanics and its succeeding theories, has been neglecting the fact, that classification is one of the fundamental means and aims of sciences.<sup>17</sup> First of all it provides us with a system of notions that allows us to talk about our empirical objects unambiguously and subtly differentiated. Secondly, it opens our eyes to the diversity of phenomena, preventing us from one-sided and blind over-simplifications. Third, it allows us to make predictions: if an object is identified to belong to a certain class, then we are able to predict all other properties belonging to that class. And finally, if our classification is systematical, we may assume the existence of objects from obvious gaps in the classification – and what is more in chemistry's classification, we even get instructions to make these objects as exemplars of new chemical substances or substance classes.

#### 3.1 From pure substances to chemical substance classes

Once a decision is made for pure substances as the basic chemical species, chemical classification can go on to divide up pure substance – and only that – into substance classes according to chemical similarities (like alcohols, carboxylic acids, esters, aldehydes, amines, diazonium salts, dibenzothiophenes, and the rest) despite great differences in physical properties. Substance classification is not strictly hierarchical in chemistry; *e.g.*, a substance may belong to the group of acids as well as to the group of aromatics, though neither of them is a subset of the other. However, it is always chemical similarity and dissimilarity, what the classification is about. We do not use equivalence classes of any quantitative physical material property as the guiding principle of substance classification. It does not mean much, for instance, if two samples show the same or nearly the same refraction index or melting point, if they are chemically quite different. However, chemically similar behavior moves them closer together, despite any quantitative difference in refraction indices, melting points, *etc.*

One might object that today's chemists detect a substance as belonging to a certain substance class, say to alcohols, by spectroscopic means, *i.e.* by sophisticated electromagnetic properties. A certain characteristic part of the IR- or NMR-spectrum unambiguously reveals an alcohol, so that chemical properties seem to be no longer important. It is certainly true that chemical properties are getting less important in chemical *detection practice*, but that does not mean that chemical properties are no longer what the classification is about. A difference is to be made between properties used for classificatory detection and properties with regard to which the classification is build up. A strong correlation between certain chemical properties of alcohols and certain

characteristics of spectra has once enabled us to select the latter as auxiliary means for detecting alcohols. But once the correlation turns out to be bad, it is not the chemical classification but the auxiliary detection method that will be modified to detect again alcohols, *i.e.* substances that chemically behave like alcohols.

It is important to note that the *concept of similarity* would not have any precise and non-arbitrary meaning in the continuous realm of homogeneous mixtures. Moreover, 'similarity' would be vague and arbitrary either, if it were related to quantitative closeness and difference in value of any physical material property. Instead the concept of similarity used for substance classification refers to the chemical behavior of chemical substances. What makes the concept of chemical similarity precise is that it refers to relations with *distinct relata*.

While the logic of relations is still an underdeveloped field of today's philosophical logic, chemistry has historically found its own ways to solve its classificatory problems. From the logical point of view, chemical substance classification reveals a certain kind of circularity, since it appears that the concepts of 'chemical similarity' and 'substance class' mutually define each other. Let us regard the following schematic definitions:

- (1) Two substances belong to the same substance class, if they are chemically similar.
- (2) Two substances are chemically similar, if each of them react under the same conditions to form product substances of a common substance class.

Note that this is not a simple *circulus vitiosus* of mutual concept definition, since we do *not* define:

- (2') Two substances are chemically similar, if they belong to the same substance class.

Instead one substance class is defined with reference to another substance class, *i.e.* we do not have a circle but a chain, or, to be more correct, a network, since substance classes refer to each other by many direct and indirect ways, including also mutual reference. Nonetheless, the substance class concepts depend on each other in a way that they remain completely loose without foundation.

How can we solve the logical puzzle? Two ways are conceivable. If we have a comprehensive set of chemical relations between substances, we could start a structural analysis of the resulting network. Because that approach requires some technical details, I turn to the second solution, which is simpler and more historical. Since every substance class is defined with reference to another class, all we need is an anchor point. We could start with some intuitive substance grouping, according to similarities of non-chemical properties.

The history of chemistry is particularly rich in such starting points. For instance, once acids were grouped according to their sour taste, the corresponding bases and salts could be classified according to their chemical relations with acids, and so on. A similar case could be made for metals and their peculiarly gleaming appearance. However, there are also more precise chemical starting points. Given the identity of chemical substances with themselves, we can derive from definition (2):

- (2a) Two different substances are chemically similar, if each of them react under the same conditions to form the same substance(s).

If we chose elementary substances as reference points and take elemental decomposition as a basic chemical relation, then we have:

- (2b) Two different substances are chemically similar, if they can be decomposed to the same elementary substance(s).

Definition (2b) still works today as a basic classificatory account according to elementary composition. The lucky fact that all substances can be chemically related to a comprehensive and systematical (though open) set of elementary substances as their final decomposition products provides the most important and basic account of chemical classification. Thus, we have a systematic set of definite anchor points from which all classification can be developed further on towards subtly differentiated substance classes.

I cannot go into details here, how cross-classifications and subclassification are achieved by the chemical approach, and how that approach has been developed and transformed towards a highly sophisticated and theoretically supported classification apparatus. The only two points I want to make here are that, first, pure substances are *necessarily* the basic objects of the chemical classification. That proves once more the classificatory importance of the distinction between pure substances and homogeneous mixtures to chemistry. And, second, the chemical classification, as the name suggests, is essentially based upon similarities of *chemical properties*, and, consequently, it retains their peculiar nature of a relational network. The importance of the second point gets clear, when we recall what classification in general means for a natural science (s.a.): namely, to provide a subtly sophisticated set of concepts enabling us to grasp the diversity of entities and phenomena at all; and to provide for predictions of properties and of new entities, including the ways to find or make them. As a consequence, whatever the chemical classification provides us in these regards is necessarily centered around chemical properties, other properties like physical, biological, and ecological properties being merely secondary aspects.

### 3.2 Predictive power of substance classification

If a substance is found to belong to the substance class of, say, carboxylic acids, that means little for its physical, biological and ecological properties. We would hardly find any nonchemical characteristic that makes formic acid and stearic acid belonging together in a particular manner.<sup>18</sup> However, from the chemical point of view, carboxylic acids are related to a wealth of other substance classes, each relation being characteristic to carboxylic acids. Thus, *on the level of substance classes, we meet the same type of relational network that we have pointed out for pure substances* (Fig. 1). For instance, carboxylic acids are related to esters and alcohols, because the formers are the products of reactions with the latter. Corresponding relations are to amines and amides; various metal salts or metals and carboxylates; halogens and halogen carboxylic acids; reducing reagents and alcohols, aldehydes, alkyls, or aryls; and hundreds more. The network structure also suggests that carboxylic acids being in turn chemically related to other substance classes from which they can be made. Thus, detecting a carboxylic acid enables us to make a wealth of predictions about its chemical properties. And what is more, the chemical classification even provides predictions of what kind of new substances can be made from it, including instructions how to make them. That characteristic, which basically depends on the relational structure of chemical properties, distinguishes it from every other type of classification we have in science.

Prediction on epistemologically reliable grounds is one of the main goals of science. As we have seen, the chemical classification is already a very strong approach for predicting chemical properties, whereas it is rather weak with regard to other material properties.

All modern natural sciences, except one, have started their enterprise with classification for the mentioned reasons, each of them from a specific perspective. The exceptional science physics tried to do without classification and started with mathematical relations between quantitative properties. Both types of approaches enabled predictions of properties, that each of the sciences was looking for. That means, the laws of physics and the (predictive) classifications of other natural sciences have exactly the same methodological status. All are theories on a basic level, but in the full meaning that philosophy of science has given to that term, *i.e.* they systematically order, predict, and, consequently, explain a certain realm of phenomena.<sup>19</sup>

### 3.3 Searching for ‘foundation’

Mere classification of phenomena and entities according to empirical properties was, despite its theoretical capacity, not the last word of classificatory sciences. Biologists, for instance, have founded their taxonomic approaches

on evolution theory. In a similar way, geology has got its foundation at least partly by tectonics. In both cases 'foundation' means, roughly speaking, relating the synchronic classificatory concepts to concepts of a diachronic theory, so that the classificatory approach is additionally justified by, *i.e.* 'founded on', a different theoretical approach. We need not go into details here, because in chemistry we obviously have no diachronic theory on which the chemical classification could be founded. What else do we have for founding the chemical classification? Is there any theory that may justify our classificatory approach, *i.e.* the networks of chemical substances and substance classes? Do we have anything else that enables us to make the wealth of predictions of chemical properties and new substances, something that provides us with instructions to make new substances, so that it actually deserves the term 'foundation'?

In a succeeding paper we will see in detail that quantum mechanics, the celebrated candidate for universal foundation, does not help at all here. Not only do we miss a reformulation, let alone a 'foundation', of the concept of pure substance. More importantly, quantum mechanics provides no classificatory concepts at all, and it cannot grasp chemical relations. Only if we ignore the essential difference between chemical and physical properties (Sect. 1.2), then we tend to overlook quantum mechanics' obvious silence with regard to chemical properties. But try to ask quantum mechanics, for instance, whether or how acetic acid can react to form an ester, and you know what I mean by silence. The crudest chemical classification on empirical grounds is incomparably more predictive with regard to chemical properties than quantum mechanics. However, quantum mechanics is in turn the most powerful account we have to predict physical, in particular, electromagnetic properties.

#### 4. Chemical theory: the language of structural formulas

What we are now looking for, is a theoretical approach that systematically represents the chemical relations, *i.e.* the networks of chemical substances and substance classes. Moreover, we call for a theory that systematizes, predicts, and explains chemical relations even more powerfully than our empirical classification does. Such an account is, as we will see, the *sign system of structural formulas*, classical chemical structure theory, that has been developed in organic chemistry since the 1860s until today. In a certain sense, this account is a molecular approach, for it represents chemical relations between substances and substance classes on the level of structural formulas that seem to represent molecules.

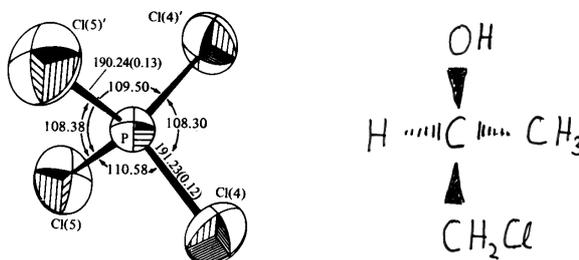
Before we regard this approach in more detail, we should first note the simple fact that neither a single molecular entity nor a mere set of such entities is a theory. Furthermore, if spectroscopic measurement or quantum chemical calculation allows us to give a molecular representation in geometrical terms, such a representation is not a theory but the *outcome* of a theory (and empirical data). We cannot derive predictions from the mere representation, unless we refer again to a theory. Thus, a molecular representation as such (a graph, a stick-and-ball model, a geometrical data set, *etc.*) is a *sign* that has to be interpreted within the framework of a certain theory in order to derive the information we are looking for. Take, for instance, a comprehensive geometrical data set about the nuclei of a molecule. The geometrical features *as such* are meaningless regarding any predictions of material properties. However, they store certain information for certain theoretical interpretations. Only if we take the data as the framework of nuclei to solve the Schrödinger equation of the corresponding electronic system (or an appropriate semi-empirical model), then we get information about energy levels that allow us to make predictions of electromagnetic properties, say. In sum, *neither a set of molecules nor a set of molecular representations is as such a theory. Instead, the theory is what provides the rules to interpret (and develop) molecular representations.*

#### 4.1 Structural formulas versus pictographic representations

It has been pointed out before (Ourisson 1986, Hoffmann/Laszlo 1991) that chemists use a lot of different molecular representations. Such a difference in molecular representation indicates that there are different interpreting theories at work, each having its own interpretation rules for certain predictive and explanative purposes. Without going into details here, I would like to point out only the main difference between structural formulas, on the one hand, and precise geometrical representations that include exact details about interatomic distances and angles, on the other. Figure 2a presents an instance of the latter: a pictographic representation of the geometrical data set of a nuclei system of a molecule as it comes out from spectroscopic measurement, x-ray diffraction, or quantum chemical calculation. Figure 2b presents a structural formula as it is used in chemical communications; to emphasize the difference, I have chosen a lax drawing that, nonetheless, does not lack any information a structural formula usually stores.

What strikes first is that, unlike Figure 2a, the structural formula of Figure 2b can hardly be interpreted in the mentioned way. Though we get some information about the constitution and symmetry, exact geometrical data of the nuclei are missing to derive quantum chemical predictions of electromagnetic properties. Our quantum chemist gets information only about *what kind* of molecule is to be calculated, but the calculation has to be started nearly from

the beginning. Thus, structural formulas are obviously not the appropriate form of molecular representation for quantum chemical interpretation and predictions of electromagnetic properties.



**Figure 2.** Two molecular representations: (a) a pictographic representation of  $\text{PCl}_4$ , including exact information about interatomic angles, distances, and thermal ellipsoids of atomic motion, (b) a structural formula of  $\text{CH}_2\text{Cl}-\text{CH}(\text{CH}_2)\text{OH}$  with information about configuration and functional groups.

However, structural formulas are actually the kind of representation that chemists mainly use, as an arbitrary look at general chemistry textbooks or journals reveals. If the use of structural formulas has some rationality, they must obviously be appropriate to a different kind of theory for a different kind of predictions.

My thesis is that structural formulas are the appropriate representation to make predictions of *chemical properties*, and that it is the only way we have until today by which such predictions can be derived systematically.

As we have pointed out above, a mere set of molecular representation as such is not a theory. What then is the theory that provides the rules for interpreting structural formulas in terms of chemical properties? In other words, how can a simple formula store information about the wealth of such complex dynamical relations that we have found to be the logical structure of chemical properties? For instance, how can a formula of acetic acid, say, tell us that the substance can react to form esters, amids, alcohols, *etc.*, that it can be made from acetone, and so on? The strange puzzle we are faced with is that the formula of a single substance must somehow carry information about many, if not all other substances, *i.e.* about the whole chemical network!

Despite their apparent similarity and common designation as ‘molecular representation’, structural formulas obviously work completely different than the geometrical graphs of Figure 2b do. The latter represent molecules (and the corresponding substances) each of their own like the signs of a pictographic language do. Structural formulas, on the other hand, are elements of a systematical language; they represent substances in certain relations with

each other, *i.e.* substances within the chemical network. It would be a cardinal error to mix up these two types of languages,<sup>20</sup> though we have some translation rules, as we will see below (Sect. 4.5).

## 4.2 Mapping the chemical network

Let us turn now to the interpretation (and generation) rules for structural formulas, which I can necessarily present only as a general sketch since the whole field is much too complex.<sup>21</sup> At first glance, structural formulas are, like the pictographic representations, built up from atomic ‘building blocks’ as the basic units. However, what counts are not the atomic units, but so-called *functional groups*, *i.e.* certain groupings of atomic units. As the name suggests, these groupings represent each of their own a certain functionality, *i.e.* a chemical reactivity with substances of certain different substance classes to form substances of certain other substance classes. Once a chemist has recognized a functional group as part of a structural formula, she is able to make predictions about the reactivity of the corresponding substance. How does she know what functional group represents what reactivity? She must have learned it before, such as we learn the rules of a language. Since similarity in chemical reactivity defines, as we have pointed out in Sect. 3.1, a substance class, it comes out that every functional group represents a substance class. As a consequence, a sufficiently developed system of functional groups exactly maps our network of substance classes (Sect. 3.1/2).

A substance can belong to several substance classes, so that its structural formula contains several functional groups. In addition, other structural parts of the formula that (still) have no functional meaning in terms of chemical functionality – such as the carbon skeleton of classical organic chemistry – makes the system of structural formulas sufficiently rich, so that every pure substance is unambiguously represented by a single structural formula. As a consequence, the system of structural formula is also able to exactly map the network of chemical substances (Sect. 1.3). Since all structural formulas are then correspondingly connected with each other to form a network, our puzzle is going to be cleared up now. Unlike pictographic representations, every structural formula represents a substance in its manifold relations to other substances; *i.e.* it represents its place within the chemical network. Thus, our chemist is able to interpret a given structural formula both in terms of how the corresponding substance can react to form other substances and how it can itself be produced from other substances.

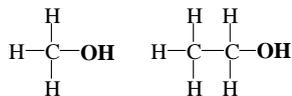
## 4.3 Theoretical surplus capacity of the sign language

Thus far, our theoretical system of structural formulas only reproduces the empirical classification of the chemical network, and as such it has the same

systematical, predictive, and explanative capacity as the latter. What about theoretical surplus capacity?

First we should admit that historically the empirical classification has been developed from rudimentary forms toward its present state only along with and under supervision of the theoretical framework of structural formulas. The chemical sign language has such an enormous systematizing power, that we can hardly imagine distinguishing chemical substances without referring to structural formulas. A nearly endless list of chemical properties of one chemical substance can be easily grasped by a single formula. For a functional group does not only represent a single property, but all chemical properties characteristic to a certain substance class, including possibly still unknown properties. Thus, structural formulas (as already the predecessors of type theory and radical theory) played an important role in guiding and sometimes correcting the empirical classification of substances.

Furthermore, structural formulas owe their semiotical precision to a systematical coarseness. What seems to be, at first glance, a paradox, is actually a rational strategy of every systematical language, namely excluding gray areas both with regard to syntactic and semantic features.<sup>22</sup> This case is very well illustrated by the structural representation of chemical similarities. *Chemical similarity between two substances is analyzed in terms of sameness and difference in structural parts of the two corresponding formulas.* For instance, unlike pictographic representations, the structural formulas both of methanol and ethanol have exactly the *same* (not only a similar!) functional OH-group and are *different* with regard to the rest (Fig. 3).<sup>23</sup>



**Figure 3.** The structural formulas of methanol (left) and ethanol (right) both have the *same* functional OH-group and are different regarding to the rest

Once we have transferred the logic of similarities of wholes into a (binary) logic of sameness and differences of parts, the field of classical logic as well as some mathematical theories (such as topology and group theory, as we will see below) becomes applicable to the chemical sign language, making it a tool with new systematizing and predictive capacities of its own. Only if we consider functional groups in different formulas as the *same*, then a system of *general rules* can be developed to interpret and, in particular, to generate new structural formulas, *i.e.* to make predictions of new chemical substances. In other words, only if we leave pictographic representations in favor of a systematical sign language, then we are able to develop a full-fledged theory.

#### 4.4 Thinking in structural formulas and working with substances

Chemists think in terms of structural formulas, their structural modifications, connections, and rearrangements, while they work with substances in the laboratory and perform chemical reactions. Apparently, structural formulas have nothing of the material properties of substances and *vice versa*. And, of course, the same goes for the two kinds of operations. If there were not a strong formal correlation between these two kinds of operations, chemists would be running the risk of falling into a kind of schizophrenia.

The correlation can be ideally conceived as a kind of translation manual: every structural formula unambiguously corresponds to a certain pure substance and *vice versa*; and every modification of a structural formula corresponds to a certain chemical change of a substance, and *vice versa*. Given both a set of structural formulas and a set of rules for allowed structural modification, we are able to generate new structural formulas by applying the rules to the former ones. The outcome of such a structural modification is, per definition, a representation of a possible substance and as such subject to further structural modification according to the rules. If we translate that into the language of substances and chemical properties, it comes out that we have *predicted new substances, including its chemical properties and the chemical way to produce them*. That is exactly, how millions of new substances have been predicted and produced during the last hundred years (Schummer 1997b/c), proving that the *chemical sign language is actually one of the most powerful predictive theories of science at all!*

Furthermore, structural formulas are subject to types of reasoning which have no direct correlation in the realm of substances, but allows us to make predictions thereof. A good example is the application of *topology* and *mathematical group theory* in such areas where possible structural formulas are limited by strict valence rules. For a given quantitative elementary composition (an ‘empirical formula’) we can calculate the number of different structural formulas, *i.e.* isomeric structures, with regard both to differences in (topological) configuration and in (geometrical) constitution. Since different isomeric structural formulas correspond to different substances, we have an indirect prediction of new substances by mathematical approaches, which are not directly applicable to the language of substances.

‘Thinking in structural formulas’ has brought about a great many rules and laws that could not have been developed on the level of substances, though a correspondence relationship was established afterwards. Most importantly, the concept of *reaction mechanism* through the introduction of intermediate formulas, that need not correspond to pure substances, but represents intermediate steps of structural modifications with its own kinds of hypothesized functional groups. Such intermediate formulas and their ‘functionalities’ can be hypothetically derived from the outcome of system-

atical reaction analysis by combinatorics. In addition, the chemical sign language has been enriched by dynamical elements, like arrows, 'switching bonds', 'jumping electrons' *etc.*, that may look strange from the pictographic or quantum mechanical point of view. But together with intermediate formulas they perfectly and uniquely fulfil the need for a more sophisticated theory to make more precise and more specified chemical predictions by taking into account systematic variations of contextual conditions. Since chemical reactions basically depend on contextual conditions (such as temperature, concentration, solvent *etc.*) with regard to both velocity and the kind of products, the study of reaction mechanisms allows us to systematically diversify chemical relations on a theoretical level.

Moreover, reaction mechanisms such as 'addition', 'elimination', 'substitution' do not only sound like algebraic operation, they also show a strong resemblance to algebraic groups, though an axiomatization has not yet been accomplished. And finally structural formulas as well as reaction mechanism open a field of their own for new kinds of analogical reasoning and similarity concepts of great heuristic value.<sup>24</sup>

#### 4.5 Translation between structural formulas and pictographic representations, and spectroscopic study of reaction mechanisms

Mixing up the difference between a structural formula and a pictographic molecular representation is a cardinal error, because the two signs belong to semiotically different types of languages and store different kinds of information for different types of theories. None of the two can substitute the other. If we ignore the difference and, say, produce only pictographic representations, then the language would in the long run lose its chemical information.

Despite that semiotical and epistemological difference, there is a strong tendency among chemists to relate both signs to a common referent, a molecule: the two signs may have different connotations, they may emphasize different aspects of reality, but they refer to the same real entity, a molecule. This view has its merits, particularly because of its strong heuristic power of mutual stimulation and supplementing. However, one should not forget that the concept of molecule is itself theory dependant, and that it is neither as unambiguous nor as universal as many believe. Simple substances like water, metals, and salts obviously evade molecular approaches. To be sure, there are many unproblematic substances. But what about the intermediate realm: the protic solvents, the neither-pure-covalent-nor-pure-ionic solids, the van der Waals complexes, and so on? To deal with these substances by a molecular approach means emphasizing certain aspects and neglecting others. And we often ignore intermolecular interactions or Einstein-Podolsky-Rosen correlations (another theory!) for no other good reasons than that they do not

help but prevent us from solving our current problems. Hence, the claim for a common referent of molecular representations called ‘molecule’, something that is more ‘real’ and more independent of a theoretical perspective is difficult to hold, if we go into details. But it perfectly serves as a heuristic working hypothesis, in order to integrate chemical knowledge. Without it, we would not have developed methods to translate structural formulas into pictographic representations and *vice versa*. The relative success of such translations, that we address now, gives at least some evidence of a common referential basis of the underlying theories of both types of representations.

The first kind of translation, from structural formulas to pictographic representations, was already mentioned (Sect. 4.1). Structural formulas provide us with information about molecular constitution but give no details about interatomic distances and exact angles. If we put the constitutional information into quantum chemical accounts, we may calculate these geometrical details and draw pictographic representations.

The other direction, translating pictographic representations into structural formulas, is more intricate though perhaps more important. Our best experimental methods for structural analysis, such as various forms of spectroscopy, including NMR and x-ray diffraction, as well as quantum chemical calculations, provide us – at least finally after successful data interpretation – with pictographic structure representations, *i.e.* with a lot of geometrical details. But these methods do not tell us, what we may consider functional groups. Translating pictographic representations into structural formulas is at first a geometrical simplification; *i.e.* we reduce details of atomic distances and angles to standard distances and angles according to standard notations for structural formulas. But the most important step is *recognizing* functional groups according to what we have learned before as part of the chemical sign language. We have to make out structural parts for which we have already established rules to interpret them as representations of certain kinds of chemical reactivity. If we do not find anything that we have learned before, then our translation fails. (The deeper reason behind that translation limit is that we have no independent theoretical approach of the concept of functional groups; or, as we will put it in the succeeding paper, there is not even a reductive approach of functional groups, let alone a successful reduction.)

However, spectroscopic methods can be used to establish new functional groups, if the methods are used to systematically investigate chemical relations between substance classes. If we find out similarities in atomic arrangement between pictographic representations of several substances, and if systematic chemical investigation reveals regular changes of that kind of arrangement into another kind of atomic arrangement along with chemical reactions under certain conditions, then we can define, in terms of structural formulas, two new functional groups related by a new rule.

Moreover, laser spectroscopic methods are today an indispensable means for investigating intermediate states of chemical changes and, thus, going into details of reaction mechanisms. The outcomes of spectroscopic analysis are pictographic representations of intermediate quasi-molecular states. Hence, they must first be translated into the language of structural formulas and assigned certain dynamical ‘functionalities’, in order to become an intermediate step of a *reaction mechanism* in terms of structural formulas.

The study of reaction mechanisms particularly tends to mix up the difference between pictographic representations and structural formulas. But here again we meet the same distinctions and the same consequences that we have pointed out for static representations. A reaction mechanism can be regarded either as a kind of ‘video presentation’ of a singular molecular event, or as a systematical rule for modifying structural formulas. The ‘video presentation’ seems to be ‘closer to reality’ than our chemical sign language. But the apparent ‘closeness to reality’ is again at the expense of theoretically. *Only if we formulate reaction mechanisms as general rules in terms of structural formulas, then we get the predictive capacity that all theoretical efforts aim at.*

While chemistry in general and structural formulas in particular are usually neglected by philosophers (of science), there is at least one philosophical classic who has pointed out the theoretical capacity of structural formulas, comparable only to mathematical formalism in physics. In the introduction of his famous *Philosophie der symbolischen Formen* Ernst Cassirer emphatically describes the systematizing capacities of structural formulas:

Am klarsten tritt vielleicht diese Tendenz [die den Zeichen ursprüngliche Kraft der Verknüpfung und Vereinheitlichung, J.S.] in der Funktion der *wissenschaftlichen* Zeichensysteme heraus. Die abstrakte chemische ‘Formel’ etwa, die als Bezeichnung eines bestimmten Stoffes gebraucht wird, enthält nichts mehr von dem, was die direkte Beobachtung und die sinnliche Wahrnehmung uns an diesem Stoffe kennen lehrt; – aber statt dessen stellt sie den besonderen Körper in einen außerordentlich reichen und fein gegliederten Beziehungskomplex ein, [...] sie faßt ihn als einen Inbegriff möglicher ‘Reaktionen’, möglicher kausaler Zusammenhänge, die durch allgemeine Regeln bestimmt werden. Die Gesamtheit dieser gesetzlichen Beziehungen ist es, die in der chemischen Konstitutionsformel mit dem Ausdruck des Einzelnen verschmilzt, und durch die nun dieser Ausdruck ein durchaus neues charakteristisches Gepräge erhält.<sup>25</sup>

And finally, in his concluding chapter on ‘the foundations of scientific knowledge’, he comes back to structural formulas emphasizing their predictive capacities:

Ganz allgemein besteht der wissenschaftliche Wert einer [Struktur-]Formel nicht nur darin, daß sie gegebene empirische Tatbestände *zusammenfaßt*, son-

dern daß sie neue Tatbestände gewissermaßen *hervorlockt*. Sie stellt Probleme von Zusammenhängen, von Verknüpfungen und Reihenbildungen auf, die der unmittelbaren Beobachtung vorausseilen. So wird sie zu einem der hervorragendsten Mittel dessen, was Leibniz die 'Logik der Entdeckung', die *logica inventiois* genannt hat.<sup>26</sup>

## 5. Conclusion

It was the aim of the previous sections to dig out characteristic parts from the rich diversity of research fields, all of which are usually called chemistry, by conceptual means and step by step:

Starting from the common sense idea that chemistry in general is concerned with material aspects of our world (1.1), we have carefully distinguished chemical properties from other material properties (1.2). Against the background of that distinction we made the basic decision to take the systematic investigation of chemical properties as the chemical core of experimental chemistry. Further analysis of the logical features of chemical properties revealed that the logical structure of systematical chemical knowledge is a peculiar network structure (1.3).

A lengthy excursion was necessary to discuss the reasons why pure substances are considered as the basic chemical species, despite serious problems of defining them by empirical (2.1) and theoretical (2.2) means. Pure substances, though artificially produced and definable only in operational terms (2.3), perfectly fulfil the chemical requirement of distinct substances serving as the nodes of the chemical network (2.4). And they do that in a way that has no alternative on the level of quasi-molecular species (2.5).

Not only are pure substances the basic chemical species; they also form the nodes of the chemical network which is already a basic kind of chemical classification. In Sect. 3 we have addressed higher order classification and regarded why and how precise concepts of substance classes are achieved only with regard to chemical similarity (3.1). The resulting classification has turned out to be again a network structure, with substance classes as nodes and chemical class relations as connections; it has enormous systematizing and predictive power with regard to chemical properties (3.2). We have then pointed out that classifications the like are already theories on a basic level and that further 'foundation' means justifying the classification by another theory with more systematizing and predictive power (3.4).

It was the aim of Sect. 4 to demonstrate that the chemical sign language of structural formulas uniquely fulfils that need of 'foundation' of the chemical classification. Thus, from the point of view of our conceptual approach we are forced to consider it as the chemical core of theoretical chemistry –

despite the fact that today's theoretical chemists are mainly concerned with quite different accounts. Since the current neglect of the theoreticity of the sign language may be due to a misinterpretation of structural formulas, we have first pointed out the fundamental difference from pictographic representations of molecular structure (4.1). A closer look at the sign language, *i.e.* its rules of interpreting and developing structural formulas, has revealed that it not only reproduces the network structures of substances and substance classes with all their systematizing and predictive capacities (4.2). The sign language also establishes a new theoretical level with new systematizing and predictive capacities (4.3-4). In a final section we have discussed the possibilities and limits of translation between the two types of molecular representations with special regard to the investigation of reaction mechanisms by spectroscopic means as a way to enhance the chemical sign language (4.5).

The whole complex – from the systematical investigation of chemical properties, over the classificatory networks of chemical substances and substance classes, to the chemical sign language – is considered the 'chemical core of chemistry'. Since the conceptual approach as well as its results might appear unusual, some final remarks shall be added to avoid possible misunderstandings. First, the above presentation was necessarily only a sketch that requires more detailed analysis and supplementing of neglected aspects such as, for instance, the role of the periodic system of elements for chemical classification. Second, it should once again be emphasized that our results have no normative implications on whatever value basis concerning the value of chemical research inside or outside of the core. Instead, 'the chemical core' is simply that part of chemistry that can be conceptually well-distinguished from interdisciplinary, applied, and specialized fields of chemistry, all of which have their own rights, of course. Third, because our approach was basically a conceptual one, one might say that we have been following traditional philosophical lines of searching after 'the essence' of chemistry. There is in fact some similarity. However, we have tried to avoid dogmatic essentialism by starting with some basic and reasonable *decisions* concerning the content and, in particular, the logical structure of basic chemical knowledge. These decisions may be subject to further discussion. But once we accept them – together with the methodological thesis that science seeks for theories, *i.e.* cognitive structures with highly systematizing and predictive capacities on epistemologically reliable grounds – the rest comes out nearly consequently. Finally, a second important meaning of the term 'chemical core of chemistry' was only briefly mentioned here and will be the subject of a succeeding paper (Schummer 1999); namely that the chemical core of chemistry is what is stubbornly resisting all attempts of physicalistic reduction.

## Notes

- \* A first draft of this paper was read and delivered at the 'International Workshop: Chemistry and Reduction', 15-17th June 1995 in Konstanz (under the title 'Problems of Physicalistic Reductions: The Chemical Core of Chemistry'). A revised version was submitted to the journal *Synthese* for a special issue on the philosophy of chemistry. Unfortunately, the anonymous referees of *Synthese* have been unable to make a decision for nearly 2 years, so that I was forced to withdraw the submission after repeatedly unsuccessful reclamation. During the years I have got many opportunities to present and discuss parts of the paper, to work on related pieces, and to read new publications on related issues, so that I finally felt the need of completely rewriting and extending the manuscript, of which the present paper is only the first part.
- <sup>1</sup> Cf. the division of chemistry into 80 sections by Chemical Abstracts.
  - <sup>2</sup> Cf. Schummer 1997b for a typology of aims and methods in preparative chemistry, and Schummer 1998 for a distinction of various approaches in physical chemistry.
  - <sup>3</sup> E.g. Duncan 1981, Nye 1989, 1993, Ruthenberg 1993, Beretta 1993, Hiebert 1995.
  - <sup>4</sup> More influential in this regard than Comte's 'positivistic' hierarchy of sciences was Engels hierarchy of *Bewegungsformen* as part of his dialectical materialism. This materialistic approach of specifying chemistry, enhanced with epistemological and methodological aspects, was continued by many philosophers of Eastern Europe; e.g. in works of Kedrow, Laitko, Richter, Simon, and many others listed in the bibliography Schummer 1996c.
  - <sup>5</sup> E.g. Primas 1981, 1985; Del Re 1987, who also takes into account the historical and ontological approach; van Brakel 1997, and many other publications to be cited in the succeeding part.
  - <sup>6</sup> I ignore the quantum curiosity of boson condensation.
  - <sup>7</sup> Cf. Schummer 1997a, p. 311, and Note 10 for reasons why a typology of material properties according to behavior runs into trouble.
  - <sup>8</sup> For a detailed study of the area of physical chemistry and its peculiar fields of material investigations cf. Schummer 1998.
  - <sup>9</sup> Note that thermo-dissociation and photo-dissociation are no chemical properties according to our typology, but thermodynamic and electromagnetic properties, respectively.
  - <sup>10</sup> If two substances are put together in a reaction vessel and nothing happens, then we get precise (and worthwhile) *chemical* information about two substances. Thus, nonreactivity is not a nonproperty, as a typology of material properties according to kinds of behavior would suggest, but a chemical property. That case is similar to the introduction of the number zero into arithmetic: zero is not a nonnumber but a 'number-of-no-quantity'.
  - <sup>11</sup> For more details and precision of the concept of chemical reaction cf. Schummer 1997a, p. 320; the concept of chemical or pure substances will be dealt with at length in Sect. 2.
  - <sup>12</sup> Most famously, of course, Putnam's semantic reduction 'water is H<sub>2</sub>O'. The original version of the present paper contained a separate section against this kind of semantic reduction and its related Quinean and Kripkean variants. A chemically well-founded criticism of Putnam's naive essentialism is presented in van Brakel 1986, cf. also van Brakel 1991, 1997.

- <sup>13</sup> There is some literature on this issue, most importantly the ‘classic’ of Timmermans (1963). For philosophical discussions and further references *cf.* van Brakel 1986, and Schummer 1996a, pp. 175-180.
- <sup>14</sup> *Cf.* the article of Joseph Earley in the present HYLE issue.
- <sup>15</sup> For more detailed arguments against an essentialism of structures *cf.* van Brakel 1991, 1997.
- <sup>16</sup> We will see later (Sect. 4.5), that spectroscopic information about quasi-molecular species as transition states is today a main source for a sophisticated understanding of chemical reactions. But that works only, if we have first translated our coarse but operationally well adapted ontology of pure substance one-by-one into a system of structural formulas, and secondly transferred the spectroscopic information into the language of structural formulas.
- <sup>17</sup> A rare exception among the forefathers of modern philosophy of science is J.F.W. Herschel.
- <sup>18</sup> Exceptions of great practical importance are the already mentioned characteristic of IR- and NMR-spectra.
- <sup>19</sup> *Cf. e.g.* Stegmüller 1970. Since we have in the present context only (non-statistical) scientific law based types of predictions, we can follow the general agreement among philosophers of science about the structural identity of prediction and explanation; *cf.* Stegmüller 1969.
- <sup>20</sup> For a general semiotical analysis of structural formulas against the background of Peirce’s semiotic *cf.* Schummer 1996b.
- <sup>21</sup> For a more detailed but still restricted presentation *cf.* Schummer 1996a, sect. 6.4. What I call ‘interpretation and generation rules for structural formulas’ is basically the semiotical side of what chemists call ‘reactions’ or ‘reaction mechanisms’ of molecules, of which Chemical Abstract Service has registered now over 2.5 Mio. The ‘ontological speech’ of chemists tends to overlook the theoreticity of these ‘reactions’, for they describe not only singular molecular events or relations between individual substances, but mostly generalized chemical relations between (open) substance classes. Thus, *the chemical sign language currently integrates millions of theoretically founded different laws!*
- <sup>22</sup> Here I basically mean what Goodman (1968, chapt. IV) has called syntactical and semantic density with reference to the languages of art and science; *cf.* Schummer 1995, pp. 165ff. for details why structural formulas do not meet Goodman’s criteria.
- <sup>23</sup> That does not, of course, exclude the possibility of subdividing and reorganizing the system of functional groups. *E.g.* we can make subdivisions according to the closeness of the OH-group to other functional groups (such as by  $\alpha$ -hydroxy-ketones and  $\beta$ -hydroxy-ketones), or consider the OH-grouping as part of another functional group (such as by carboxy-groups. Our point is a more general and logical one, that a functional group, on whatever state of the art, is considered the same in different formulas.
- <sup>24</sup> *Cf.* Schummer 1996a, sect. 6.5.3 for more details and references.
- <sup>25</sup> “This tendency [the signs’ original power of connection and unification] probably most clearly comes out in the function of *scientific* sign languages. For instance, an abstract chemical ‘formula’ used to designate a certain substance, contains nothing of what direct observation and sense perception tell us about this substance; – instead it places the particular substance into an extraordinarily rich and subtly

structured complex of relations, [...] it [the formula] grasps it [the substance] as the embodiment of possible 'reactions', of possible causal relationships determined by general rules. The totality of these regular relationships merges with the expression of the individual into the chemical constitution formula, so that this expression now receives quite a new characteristic." (Cassirer 1923, vol. 1, p. 44f.).

- <sup>26</sup> "In general, the scientific value of a [structural] formula is not only that it *unites* given empirical facts, but that it *lures out*, so to speak, new facts. It puts forward problems about relations, connections, and formation of order, which precede immediate observation. Thus, it becomes one of the most outstanding means of what Leibniz has called the 'logic of invention', *logica inventionis*." (Cassirer 1923, vol. 3, p. 513).

## References

- Amann, A.: 1993, 'The Gestalt Problem in Quantum Theory: Generation of Molecular Shape by the Environment', *Synthese*, **97**, 125-156.
- Beretta, M.: 1993, *The Enlightenment of Matter. The Definition of Chemistry from Agricola to Lavoisier*, Watson Publishing International, Canton/MA.
- Cassirer E.: 1923, *Philosophie der symbolischen Formen*, 3 vols., Oxford (2nd edn. Oxford 1956)
- Del Re, G.: 1987, 'The Historical Perspective and the Specificity of Chemistry', *Epistemologia*, **10**, 231-240.
- Duncan, A. M.: 1981, 'Styles of Language and Modes of Chemical Thought', *Ambix*, **28**, 83-107.
- Goodman, N.: 1968, *Languages of Art. An Approach to a Theory of Symbols*, Bobbs-Merril, Indianapolis.
- Hiebert, E. N.: 1996, 'Discipline Identification in Chemistry and Physics', *Science in Context*, **9**, 93-119.
- Hoffmann, R.; Laszlo, P.: 1991, 'Darstellungen in der Chemie – die Sprache der Chemiker', *Angewandte Chemie*, **103**, 1-16. (*Angew. Chem. Int. Ed. Engl.*, **30** (1991) 1-16).
- Luisi, P.-L.; Thomas, R.M.: 1990, 'The Pictographic Molecular Paradigm', *Naturwissenschaften*, **77**, 67-74.
- Nye, M.J.: 1989, 'Chemical Explanation and Physical Dynamics: Two Research Schools at the First Solvay Conferences 1922-1928', *Annals of Science*, **46**, 461-480.
- Nye, M.J.: 1993, *From Chemical Philosophy to Theoretical Chemistry. Dynamics of Matter and Dynamics of Disciplines 1800-1950*, Univ. of California Press, Berkeley.
- Ourisson, G.: 1986, 'Le langage universel de la chimie: les idéogrammes. Ambiguïtés et laxismes', *L'actualité chimique*, 41-46.
- Primas, H.: 1981, *Chemistry, Quantum Mechanics and Reductionism*, Springer, Berlin-Heidelberg-New York.
- Primas, H.: 1985, 'Kann Chemie auf Physik reduziert werden?', *Chemie in unserer Zeit*, **19**, 109-119, 160-166.
- Ruthenberg, K.: 1993, *Was ist Chemie? Die Selbstbestimmung einer Disziplin im historischen Wandel*, Univ. of Göttingen (unpubl. Magister thesis).

- Schummer, J.: 1995, 'Ist die Chemie eine schöne Kunst? Ein Beitrag zum Verhältnis von Kunst und Wissenschaft', *Zeitschrift für Ästhetik und allgemeine Kunstwissenschaft*, **40**, 145-178.
- Schummer, J.: 1996a, *Realismus und Chemie. Philosophische Untersuchungen der Wissenschaft von den Stoffen*, Königshausen & Neumann, Würzburg.
- Schummer, J.: 1996b, 'Zur Semiotik der chemischen Zeichensprache: Die Repräsentation dynamischer Verhältnisse mit statischen Mitteln', in: P. Janich, N. Psarros (eds.): *Die Sprache der Chemie*, Königshausen & Neumann, Würzburg, pp. 113-126.
- Schummer, J.: 1996c, 'Bibliographie chemiephilosophischer Literatur der DDR', *Hyle*, **2**, 3-11.
- Schummer, J.: 1997a, 'Towards a Philosophy of Chemistry', *Journal for General Philosophy of Science*, **28**, 307-336.
- Schummer, J.: 1997b, 'Scientometric Studies on Chemistry I: The Exponential Growth Chemical Substances, 1800-1995', *Scientometrics*, **39**, 107-123.
- Schummer, J.: 1997c, 'Scientometric Studies on Chemistry II: Aims and Methods of Producing new Chemical Substances', *Scientometrics*, **39**, 125-40.
- Schummer, J.: 1998, 'Physical Chemistry – neither Fish nor Fowl?', in: P. Janich, N. Psarros (eds.): *The Autonomy of Chemistry*, Königshausen & Neumann, Würzburg (forthcoming).
- Schummer, J.: 1999, 'The Chemical Core of Chemistry II: Resisting Physicalistic Reduction' (forthcoming).
- Stegmüller, W.: 1969, *Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie, Band I: Wissenschaftliche Erklärung und Begründung*, Springer, Berlin-Heidelberg-New York.
- Stegmüller, W.: 1970, *Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie, Band II: Theorie und Erfahrung, 1. Halbband*, Springer, Berlin-Heidelberg-New York.
- Timmermans, J.: 1963, *The Concept of Species in Chemistry*, 2nd edn., New York.
- van Brakel, J.: 1986, 'The Chemistry of Substances and the Philosophy of Mass Terms', *Synthese*, **69**, 291-324.
- van Brakel, J.: 1991, 'Chemistry', in: H. Burkhardt, B. Smith (ed.): *Handbook of Metaphysics and Ontology*, vol. I, Philosophia, München, pp. 146-147.
- van Brakel, J.: 1997, 'Chemistry as the Science of the Transformation of Substances', *Synthese*, **111**, 253-282.

Joachim Schummer:

Institute of Philosophy, University of Karlsruhe, D-76128 Karlsruhe,  
Germany; Joachim.Schummer@geist-soz.uni-karlsruhe.de