Sergio Chibbaro · Lamberto Rondoni
Angelo Vulpiani

Reductionism, Emergence and Levels of Reality

The Importance of Being Borderline

Springer

Reductionism, Emergence and Levels of Reality

Sergio Chibbaro · Lamberto Rondoni
Angelo Vulpiani

Reductionism, Emergence and Levels of Reality

The Importance of Being Borderline

Springer

Sergio Chibbaro
University Pierre et Marie Curie
Institut Jean Le Rond d'Alembert
Paris
France

Angelo Vulpiani
"La Sapienza" Dipartimento di Fisica
University of Rome
Roma
Italy

Lamberto Rondoni
Dipartimento di Scienze Matematiche
Polytechnic University of Turin
Torino
Italy

ISBN 978-3-319-06360-7 ISBN 978-3-319-06361-4 (eBook)
DOI 10.1007/978-3-319-06361-4
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014938209

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To the glorious memory of Ludwig Boltzmann who struggled against the stream of his time

Foreword

Newton's third law does not apply to the interaction between philosophers ('them') and physicists ('us'). It has usually been asymmetrical, with 'us' influencing 'them', without 'them' acting on 'us'. In a way this is natural, because the raw material that philosophers study are the discoveries and theories of science and the interactions between scientists, while the primary preoccupation of physicists is not the study of philosophy or philosophers. I do not deny that there have been eminent scientists (Einstein, Poincaré, Bohr…) who have pondered on the philosophical significance of the scientific picture of the world, and much of what they said has been immediately appreciated by practicing scientists. But their wise intellectual interventions have usually been outside the philosophical mainstream.

This book by Sergio Chibbaro, Lamberto Rondoni and Angelo Vulpiani (CRV) is an exception. Although their day job is the practice of theoretical physics, they have something genuinely new to say about the physicist's picture of the world, that should be of interest to philosophers. Their focus is on what has long been studied by philosophers as 'the problem of reduction'. This concerns the relations between different levels of description of physical phenomena.

Optics is a good example. Light can be described in terms of the rays of geometrical optics, as interfering waves, as electromagnetic fields, or as the photons of quantum field theory. These are levels of increasing generality; each encompasses phenomena described at the previous levels and also includes new phenomena that were unexplainable earlier. But the concepts and mathematical expressions of these levels are very different, and moving between them is almost always challenging. It is far from straightforward to derive the formula relating the object and image of a simple lens by starting from the field operators of quantum optics supposedly the deepest of our current pictures of light.

This illustrates a wider difficulty. One can have a general theory—or even, as some envisage, the theory of everything—which stubbornly resists attempts to employ it to explain phenomena that were well understood at more elementary levels: at the more sophisticated level, they are emergent. This was cleverly caricatured in Ian McEwen's novel *Solar*: a string theorist, caught by his wife in a compromising situation with another woman, tries to reassure her: "Darling, I can explain everything". It often happens that theories claiming great explanatory

reach are in fact powerless to explain many particular phenomena. A political analogy comes to mind: the ideologist who loves all humanity but behaves badly to every individual person he encounters.

The resolution of these difficulties starts from the observation that the theories of physics are mathematical, and relations between them involve limits as some parameter vanishes: wave optics 'reduces to' geometrical optics when the wavelength is negligibly small, quantum physics 'reduces' to classical physics when Planck's constant can be neglected, etc. Therefore understanding relations between levels must involve the study of limits, that is, mathematical asymptotics. And the central reason why 'reduces to' is so problematic is the fact that the limits involved are usually singular. These singular limits should be not regarded as a nuisance, and certainly not as deficiencies of the more general theories. On the contrary, they should be embraced with enthusiasm, because they are responsible for fundamental phenomena inhabiting the borderlands between theories—phenomena at the forefront of physics research, such as critical phenomena in statistical mechanics, fluid turbulence and the universal statistics of the energy levels of highly excited quantum systems.

CRV fully appreciate these ideas; hence their subtitle *The Importance of Being Borderline*. And they explore them in depth and in detail. There are some philosophers who have grasped the significance of singularities in mathematical asymptotics for the understanding of theory reduction—Robert Batterman and Alisa Bokulich come to mind—but they remain a minority. The value of CRV's account is that it is the first full-length and wide-ranging exposition of this point of view by physicists who are sensitive to the concerns of philosophers.

Bristol, UK, December 2013 Michael Berry

Preface

Considerate la vostra semenza:
fatti non foste a viver come bruti,
ma per seguir virtute e canoscenza

(Consider your origin;
you were not born to live like brutes,
but to follow virtue and knowledge)

Inferno, Canto XXVI Dante Alighieri

The attempt to explain the sensible world in terms of a few unifying principles has been a constant of scientific thought since the times of the Presocratic philosophers, who believed to have identified several fundamental elements of which the world had to be made: air for Anaximenes; water for Tales; and air, water, earth and fire for Empedocles. Subsequently, Pythagoras tried to interpret everything sorting out of relationships among integer numbers.

In the fifth century B.C. Democritus boldly hypothesised that reality, which appears so varied and changeable, is nothing but a collection of indivisible and eternal parts, or *atoms*. In one of the few fragments which has survived to our times, the fundamental hypothesis of Democritus can be quite clearly summarised:

Reason: *Sweet exists by convention, bitter by convention, heat by convention, cold by convention, color by convention; but atoms and the void exist in truth.*

Over the centuries, despite the ups and downs, the doctrine of Democritus, has been a landmark of science, and in physics there has been a great deal of progress thanks to the clear distinction that has been made between subjective perception and objective reality. Currently, for example, it is understood how apparently subjective properties such as temperature and colour can be unambiguously understood in terms of physical objective quantities, such as molecular motion and the vibrations of the electromagnetic field.

The grand theory of Democritus was visionary but faced major challenges, because the total objectivity that seems so attractive in the development of a unified conceptual framework poses problems that are difficult to solve. Much as one can be firmly convinced that the ultimate essence is represented by atoms, certain questions cannot be avoided: only through our senses, and with the mediation of appropriate technological tools that are conceptual extensions of our senses, can one hope to reveal the atomic structure of matter.

This was already clear to Democritus; here, as in his dialogue, the Senses respond to Reason:

Wretched mind, after receiving from us your knowledge, do you try to overthrow us? The overthrow will be your downfall.

The founder of atomism formulated in a very clear fashion the problem that will be discussed in this book: the issue of reduction in physics. The idea, in its simplest form, is that *the whole is nothing more than the sum of the parts*; that is, the behaviour of things is directly determined by the properties of their elementary constituents. The properties observed at the level of the whole are related to the properties at a lower level of observation, which hence seem to be more fundamental. For example, the motion of a fluid can be related to that of its molecules, or, climbing much higher in the hierarchy of levels, consciousness can be related to the behaviour of neurons. Reductionism is then the activity concerned with the relation between different theories that attempt to describe different levels of reality, or different levels of observation, and, through a qualitative leap consisting of many simplifications, ultimately to relate the different sciences: psychology, physiology, biology, chemistry and physics.

Reductionism, or more generally the relationship connecting the different sciences, is perhaps one of the few issues of the scientific culture which has been also considered in literature. How can one forget the bewildered Mr. Palomar, immortalised by Italo Calvino, who muses over the sea, cheese, the flight of birds and the meadows:

The lawn is a collection of grasses—this is how the problem must be formulated that includes a subcollection of cultivated grasses and a subcollection of spontaneous grasses… The two subcollections, in their turn, include various species, each of which is a subcollection, or rather it is a collection that includes the subcollection of its own members, which are members also of the lawn and the sub-collection of those alien to the lawn… is "the lawn" what we see or do we see one grass plus one grass plus one grass…? What we call "seeing the lawn" is only an effect of our coarse and slapdash senses; a collection exists only because it is formed of discrete elements. There is no point in counting them, the number does not matter; what matters is grasping in one glance the individual little plants, one by one, in their individualities and differences (Calvino 1983).

It is clear that Mr. Palomar, and not only he, is continually oscillating between the two opposite points of view that see the lawn either as the sum of so many blades of grass or as something more than just all the individual seedlings. The reality certainly shows a complicated (or complex?) structure of relations, and the dream of a unified interpretation of all phenomena in terms of several simple laws, from which all can be deduced, has attracted and continues to attract almost anyone who has genuine philosophical and scientific interests.

Reductionism seems also to be one of the few topics still capable of stirring peremptory and robust discussions among dedicated scholars. As example, we propose three contributions with significant titles, respectively by Atkins, Midgley and Edelman: *The Power of Science Limitedness; Reductive Megalomania* and *Individual and Soul Memory: Against Silly Reductionism*, included in an

interesting volume that collects the opinions of authors having different points of view (Cornwell 1995).

Pondering the difficulties faced in the attempt to show that the behaviour of real systems can be deduced from simple laws, the 'hard-core' reductionists, who have their stronghold in high-energy physics, typically acknowledge obvious technical problems, but insist that reduction is in principle possible and conceptually correct. This radical reductionism is sometimes shared by scientists working in other branches of science. The well-known chemical-physicist Atkins claims that a true scientist has to be reductionist and that an anti-reductionist point of view is necessarily obscurantist: *Theism (and the implicit rejection of reductionism) is a system of knowledge based on ignorance, and that twin of ignorance, fear.*

Atkins champions a certain vulgate which tends to equate the radical reductionist with the 'hard-core', true lovers of the mathematical and hard sciences, such as physics, who do not indulge in extra-scientific considerations, like those concerning morality and religion. In contrast, anti-reductionists would be incurable romantic people, who complain about the cold rationality of 'official science'. Nevertheless, the very same source (Cornwell 1995) collects papers by scientists of impeccable reputation, like Dyson, Chaitin and Edelman, whose positions are clearly non-reductionist.

In fact, the state of affairs is rather complex. For instance, the Italian philosopher Severino (1997), contrary to folkloristic vulgarisations, believes that reductionism is one of the pillars of religion:

> "reductionism" is not the enfant terrible of present-day scientific, physicalistic culture: the desire to connect something to its origin—especially the desire to connect the world to its divine origin—is reductionism. Theology is the fundamental form of reductionism. Indeed, theology reduces the essence of the world to God, in the same fashion that in science, one day, reductionists will reduce all human reality to movements of elementary particles.

Feigenbaum, one of the "fathers" of chaos theory, similarly maintains in an interview reported in (Horgan 1997), that: *many physicists like the idea of final theories because they use it to replace God.* Finally the words of Laughlin and Pines (2000) are particularly poignant: *The belief on the part of many that the renormalisability of the universe is a constraint on an underlying Theory of Everything rather than an emergent property is nothing but an unfalsifiable article of faith.*

If the most radical reductionist point of view was correct, the relationship between the different scientific disciplines would be of strictly inclusive type: chemistry contained in physics, biology in chemistry and so on. Eventually, only one science, indeed just a single theory, would survive, since all the others would eventually be embedded in the Theory of Everything. Is the current coexistence of different sciences a mere historical parenthesis which will end when the Theory of Everything has been worked out?

At the same time, it seems that at every moment in history, some have doubted the possibility of a unified description of nature, and not only among scientists of little formalised disciplines, i.e. with a mathematical apparatus less advanced than

that of physics, such as biology. For instance, in 1856, the young Maxwell addressed the Cambridge Apostles, stating: *perhaps the "book", as it has been called, of nature is regularly paged; if so, no doubt the introductory parts will explain those that follow, and the methods taught in the first chapters will be taken for granted and used as illustrations in the more advanced parts of the course; but if it is not a 'book' at all, but a magazine, nothing is more foolish than to suppose that one part can throw light on another* (Campbell and Garnett 1882).

Laughlin and Pines (2000) suggest that to refute the radical reductionist approach leads one to propose new categories for those properties which do not appear likely to be reduced to the mere sum of their constituents. The most appropriate category is then that of emergence.

The first purpose of this book is to analyse some aspects of theory reduction in physics and to stimulate some reflection on questions that spontaneously arise in this field:

1. Is there any evidence of actual reduction of theories? Are the examples that are found in books on the philosophy of science too simplistic or not completely correct?
   For example, we shall discuss the reduction of thermodynamics to statistical mechanics, which is considered by many a paradigm case of reductionism. This relation is for others a counterexample. Indeed, the passage from the microscopic level to the macroscopic one is not a simple translation between two different languages, as can be appreciated observing, for instance, that the erratic behaviour of the elementary parts (the molecules) leads to total order.
2. If the microscopic level is essential to determine the macroscopic one, as Weinberg (1987) says when he writes *we understand perfectly well that hydrodynamics and thermodynamics are what they are because of the principles of microscopic physics*, why have we understood hydrodynamics quite well for centuries?
   More specifically, if microscopic laws are truly relevant, how can we explain the possibility of reproducing hydrodynamic behaviour by means of cellular automata that violate several fundamental properties of the microscopic dynamics (e.g. they occur in discrete states and do not follow strictly deterministic rules)?
3. The most extreme reductionists argue that by abandoning research on the Theory of Everything, unification of the sciences would never be achieved and just two alternatives would remain: a devastating fragmentation of applied knowledge and a weak science, like a sort of mysticism based upon sickening mantra along the lines of: *everything is more than the sum of its parts.* Is this really so?
4. What has been endangered by the search for (the) ultimate truth? Has the dream of reductionist reason created any monsters? Is *big science* one such monster? We have built accelerators tens of kilometres in size, to study phenomena that occur on scales of the order of 10^{-20} sec. while we do not understand many issues concerning the macroscopic level. In an effort to study smaller and smaller scale phenomena, from those of atomic physics to those of nuclear

physics and elementary particles, physics, the leading science for a long period, has marginalised many important issues, some of which have been recently rediscovered and repopularised.

5. What is the point of embedding science *Y* within science *X*, if predictions concerning science *Y* cannot be made starting from science *X*?

Even some of the most extreme reductionists admit that there are practical, maybe insurmountable, difficulties in performing theory reductions. For instance, after having praised the merits of the Standard Model, Weinberg (1995) states: *It seems that quantum chromodynamics is mathematically self-consistent, but it describes an impoverished universe in which there are only nuclear particles—there are no atoms, there are no people*. Nevertheless, the reductionist approach in its radical forms remains very influential in physics, and Weinberg himself concludes the same paper proposing his dream of an ultimate truth: *Perhaps our best hope for a final explanation is to discover a set of final laws of nature and show that this is the only logically consistent rich theory, rich enough for example to allow for the existence of ourselves*.

The theme of reductionism has been, and still is, the source of many discussions in areas other than physics. In recent times, it has been analysed especially in relation to one of its corollaries, *emergent* properties, which is a traditional theme in philosophy as well as in the biological sciences and those related to them (such as neurology, philosophy of the mind and psychology). As in the philosophy of science, the problem of reductionism typically arises between two completely formalised theories. The motivation of this approach is technically flawless: if the two theories are not even formalised, the problem is just too vague even to be posed. But in doing so, the danger is that the issue is rendered void or uninteresting. However, completely formalised theories are very few, even in physics, and the areas of most advanced research cannot be axiomatised, as mathematics has been,[1] because of the insufficient lack of understanding, which is unavoidable in the early stages of research in any given field.

Given all this, it would seem that no reasonable person could deny that some notion of reductionism has contributed to major advances in many areas of science. However, as pointed out by Mayr (2004), in doing so one should take care not to confuse reductionism with analysis, which is the method of breaking down, of dissecting a complicated (or complex) system into its parts. This approach, which we call 'methodological reductionism', is almost inevitable, but by no means does it imply that starting from the study of individual components of a given object, i.e. starting from a useful description of those parts, can the behaviour of the system, seen as a whole, be justified.

Perhaps it is no coincidence that the reductions of theories discussed by philosophers, with just a few exceptions, are interesting from a historical point of

[1] In reality, even mathematics is not completely formalised, and that is not always considered a negative fact. For example, Thom joked about the fact that the term 'rigour' reminds him of the 'rigor mortis'.

view, allow a rigorous logical analysis but are not very stimulating for scientists active in research, and are difficult to apply to the situations of current scientific interest. Among the few exceptions we find those considered in two interesting books, Batterman (2002) and Bokulich (2008), and in the collection of articles (Humphreys and Bedau 2006). A statement by the editors of this last source is particularly noteworthy: *we believe that progress in understanding emergence will be helped by a familiarity with work in areas outside psychology and the philosophy of mind. By understanding how emergent phenomena occur and are represented in physics and artificial life, for example…*

Einstein said that to understand how theoretical physicists work it is not necessary to care too much about what they say (especially on official occasions) but rather to look at what they do. In a similar way, to get an idea of how reduction of theories really works, it seems better to go beyond the statements of the various scientists, distinguished as they may be, and beyond the general formulations of epistemologists. Indeed, even Einstein was not immune from inconsistent thoughts, for example, he claimed that: *The supreme task of the physicist is to arrive at those universal elementary laws from which the cosmos can be built up by pure deduction.* But his major scientific achievements, beginning with those published in the *annus mirabilis*, did not at all follow a reductionist approach. Only after reaching the age of 40 did he embark on a project of unification, which was however doomed to failure. The same path was followed by Hilbert, who after having made so many major contributions to the most diverse sectors of mathematics and physics, entered into a fruitless programme of the complete formalisation of mathematics.

It is worth mentioning that declarations of principle sometimes arise from legitimate reasons in defence of academic interests. It is no secret that the passionate defence of reductionism pursued by Weinberg was largely motivated by lobbying in the U.S. Congress in support of the SSC project, the giant Superconducting Supercollider intended to study the physics of elementary particles. Similarly, Anderson's manifesto, *More is Different* (Anderson 1972), constitutes an attempt of condensed matter physicists, which we appreciate, to fight at a cultural level the overwhelming power of elementary particle physicists. After the papers by Weinberg and Anderson, with their opposite views on the role of the Big Science at the beginning of the 1990s, reductionism has been widely reconsidered and debated by physicists, see e.g. Schweber (1993); Anderson (1995).

The second purpose of our book is to look at a few cases taken from contemporary physics and to clearly present their relevance to the debate on reductionism, hence emergence.

Indeed, we do not find it particularly interesting to argue in favour or against the reductionist point of view, which in its most extreme formulation may even be a ghost, since virtually no one has really supported it and even fewer have practised it. We find it much more interesting to study in detail some specific cases without hesitating to enter into technical matters. The point that seems important to us is to understand the reasons for which a very detailed understanding of many complex phenomena is possible without having to resort to any Theory of Everything: adopting this perspective, we would like to understand the real connection between

the different levels of description of reality. In the same vein, we also argue that emergent phenomena are by no means inexplicable. Indeed, they can be understood through the analysis of the connections between different levels of description or theories.

Therefore, one main focus of this book is to investigate the fact that the structure of various sciences (or more often theories) $X_1, X_2, \ldots, X_n, \ldots$ is not hierarchical and inclusive—which would mean that X_1 is contained in X_2, X_2 contained in X_3 and so on—but that, on the contrary, there is only a partial overlap between 'neighboring' sciences. By contrast, 'distant' sciences are almost completely disconnected: X_1 partially overlaps with X_2, but has almost no intersection with X_3; X_2 partially overlaps with X_3 but has almost no connection with X_4 and so on.

The really crucial point that we want to make here is that the interesting things occur at the borders, where two different levels of description meet: this often leads to novel concepts that cannot be categorised into what epistemologists call *bridge laws*, but rather can be considered as emergent properties. Indeed, the transition from one theory to another does not happen in a regular or simple way, in general, but through a procedure technically known as a *singular limit*.

This type of structure, which perhaps leaves the fundamentalists of ultimate truth unsatisfied, is in our opinion the motivation for which science exists and speaks sensibly about the natural phenomena occurring at a given level of observation (e.g. through the laws of hydrodynamics) without having to refer to a detailed comprehension of the underlying levels of observation (e.g. of the microscopic dynamics).

The central aspect resides in the fact that the transition from a more fundamental theory to a more phenomenological one almost never consists in a mere process of approximation, but rather in the emergence of new properties and concepts which are not present in the more fundamental level of description.

In our view the relevance of these facts is not always well understood and appreciated. Berry (1994) and Primas (1998) are among the few who have strongly and explicitly acknowledged the need for a singular limit in some of the most interesting reductions of theories. This does not amount to a mere technicality but is the crucial point from which concepts that cannot be contained in the more detailed theory emerge.

This book cannot and does not aim to be a text on the philosophy or history of science. Our main intention is to discuss, with specific examples, the importance of the singular limits and the emergence of new concepts in the process of theory reduction. Therefore, this book is primarily dedicated to all researchers in physics and in neighbouring sciences, chemistry and mathematics in particular, who want to reflect critically on their work. We hope, however, that the book is of interest to those philosophers of science who like to consider the most recent results obtained in the natural sciences and their consequences for more general studies.

Our selection of arguments follows a natural criterion: it is dictated by our personal research interests and competence, as developed in our recent scientific activity.

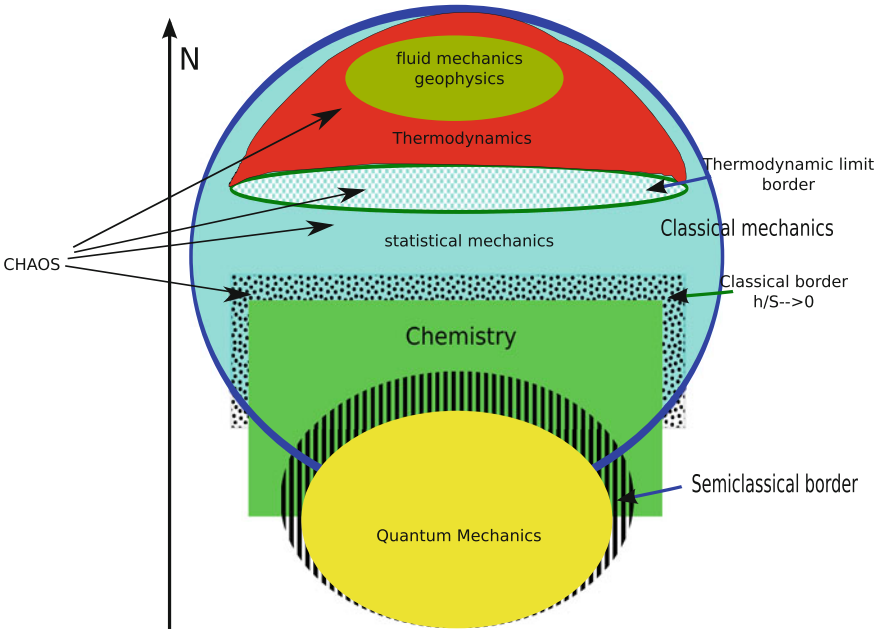

Fig. 1 Inter-theoretic relations treated in this book. The levels of reality correspond roughly to increasing the number of elements involved (*N* in the picture)

The structure of the book is as follows: we begin with a Dialogue that we would like to think of as being in the style of Galileo, which informally introduces the technical issues discussed in the rest of the book.

Chapter 2 introduces the problem of reductionism from a historical and philosophic view and does not claim to be either complete or original. Furthermore, the possible original implications for the field of epistemology of our analysis of physical theories are beyond the scope of the present book. Nevertheless it has been included so that the book is self-contained, and to the benefit of readers lacking a solid background in the philosophy of science. For the same reasons, we have kept a general description outlining the precise framework of our thesis.

Each of the central chapters (Chaps. 3–6) treats a specific subject: *Statistical Mechanics and Universality*; *Irreversibility and macroscopic behaviour*; *Chaos*; and *Quantum Mechanics and Chemistry*. The main purpose of these chapters is to demonstrate the importance of singular limits in all attempts at inter-theoretic reductions. These chapters also stress that these reductions do not conform to the standard textbook views of philosophy of science, based on the identification of microscopic theory, macroscopic theory and bridge laws. A sketch of this structure is shown in Fig. 1. These self-contained chapters can be read independently of one another.

Chapter 7 includes a partial analysis of more recent developments. In particular, we present some random thoughts about the unity of science in a non-reductionist point of view, as well as discussing some propositions of unifying theories concerning, e.g. fractals, dissipative structures, computational algorithms, which have become popular over the last few decades. In our opinion, those theories are often more or less concealed or disguised descriptions of different forms of reductionism, which deserve to have a critical eye cast over them at the very least, because of their popularity.

Perhaps the reader wonders about the roles of the characters Simplicio, Sagredo and Salviati, in the dialogue. Salviati expresses opinions that the authors share and, as in Galileo's dialogue, Sagredo is a non-impartial referee. Simplicio, at variance with the character of the Galilean dialogue, is not totally naive, he represents points of view that are commonly advocated by a number of contemporary scientists. Since active researchers are not all interested in the foundations of their disciplines, it is not unusual to find inconsistencies in their views.

We conclude these introductory notes with some clarifications.

In the first place, we observe that our purpose is that of making a simple point concerning theory reduction in physics, by examining several examples of current interest. We do not attempt to exhaustively address all foundational issues related to all branches of physics, including those concerning the various interpretations of quantum mechanics. We could have made the same point discussing other examples, such as those of statistical/entropic forces, of Casimir forces, of cold atoms, of the coherence or decoherence phenomena in modern macroscopic experiments, of Verlinde's theory of gravity as fluctuation induced, etc. Our selection is motivated by our familiarity with the chosen examples.

Throughout the book we quote authors who are well known in our fields of research, but who may not be known to a wide readership. Nevertheless, we preferred not to add bibliographic notes for them, for sake of simplicity and of readability.

Finally, historical references are made, including anecdotical ones. Surely, they do not meet the exacting standards of professional historians. However, we make no pretence of producing a historical opus. These references have always been used as conceptual examples. In any case we have detailed our bibliographic references.

The authors thank D. Aubin, M. V. Berry, M. Falcioni, S. Goldstein, G. Jona-Lasinio, C. Malaterre, M. Rossi, E. Sanchez-Palencia, H. A. Stone and N. Zanghí for the careful reading and precious suggestions. A. V. thanks R. Burioni, M. Casartelli and D. Cassi with whom, some years ago, he had many fruitful discussions about reductionism and emergence. Thanks are in order also to O. G. Jepps, who has competently and extensively revised our English. We acknowledge the support of the Kavli Institute for Theoretical Physics China (KITPC), Institute of Theoretical Physics (ITP), Chinese Academy of Sciences (CAS) in Beijing. Lastly, we are grateful to M. V. Berry for his friendly encouragement and the many remarks.

References

Anderson, P.W.: More is different. Science. **177**(4047), 393 (1972)
Anderson, P.: Historical overview of twentieth century physics. In: Brown, L., Pais, A., and Pippard, B. (eds.) Twentieth Century Physics, p. 2017. American Institute of Physics press, New York (1995)
Batterman, R.W.: The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence. Oxford University Press, New York (2002)
Berry, M.V.: Asymptotics, singularities and the reduction of theories. In: Logic, Methodology and Philosophy of Science IX, p. 597. (1994)
Bokulich, A.: Reexamining the Quantum-Classical Relation. Cambridge University Press, Cambridge (2008)
Calvino, I.: Palomar, Einaudi (1983)
Campbell, L., Garnett, W.: The life of James Clerk Maxwell. MacMillan and Co., London (1882)
Cornwell, J. (ed.): Nature's Imagination: The Frontiers of Scientific Vision. Oxford University Press, USA (1995)
Horgan, J.: The End of Science. Broadway Books, New York (1997)
Humphreys, P.E., Bedau, M.A.: Emergence: Contemporary Readings in Philosophy and Science. MIT Press, Cambridge (2006)
Laughlin, R.B., Pines, D.: The Theory of Everything. P.N.A.S., vol. 97, p. 28 (2000)
Mayr, E.: What Makes Biology Unique? Considerations on the Autonomy of a Scientific Discipline. Cambridge University Press, Cambridge (2004)
Primas, H.: Emergence in exact natural science. Acta Polytech. Scand. 91, 83 (1998)
Schweber, S.S.: Physics, community and the crisis in physical theory. Physics Today. 46, 34 (1993)
Severino, E.: Si può capire l'universo? In: Coyone, G.V., Giorello, G., and Sindoni, E. (eds.) La Favola dell'Universo, p. 62. Piemme, Casal Monferrato (1997)
Weinberg, S.: Newtonianism, reductionism and the art of congressional testimony. Nature. **330**(6147), 433 (1987)
Weinberg, S.: Reductionism Redux. The New York Review of Books (1995)

Contents

Chapter 1
A Galilean Dialogue on the Levels of Reality

A number of flawed individuals can often add up to a brilliant social unit.

There is, in short, no great idea that stupidity could not put to its own uses; it can move in all directions, and put on all the guises of truth. The truth, by comparison, has only one appearance and only one path, and is always at a disadvantage.
The man without qualities, Robert Musil

SALVIATI: Nearly four centuries after our first memorable dialogue, I think it is high time to meet again and debate the new problems which have arisen over such a long time. If I remember rightly, at the end of the fourth day last time, we promised that one day we would resume our discussion.

Therefore, I suggest that we analyse thoroughly the arguments of the proponents of reductionist philosophy—according to which Nature at all its levels can be understood just from a knowledge of its ultimate elementary constituents—as well as the arguments of its opponents.

SIMPLICIO: It is certainly a good idea to bring our dialogue up to date, but this time I'm sure there'll be no major disagreements or misunderstandings between us.

The Masters of the past have shown the way ahead with great clarity; hoisted on to their shoulders, we can almost see the final solution.

Galileo, Newton and Maxwell effectively solved all mechanical and electrodynamics problems. Superb theories such as relativity, quantum mechanics and quantum electrodynamics, give us perfect insight into the molecular and atomic structure of matter. Soon the few remaining mysteries will be definitively clarified and we can already say that there is nothing more to be understood, in physics at least. Furthermore, we must acknowledge the formidable successes obtained in molecular biology by Watson and Crick, who discovered DNA, and we are enlightened by Monod's work.

SALVIATI: Do you refer to the thesis put forward by the cosmologist S. Hawking, now famous for his beautiful books who, accepting the Lucasian chair (the same of

S. Chibbaro et al., *Reductionism, Emergence and Levels of Reality*,
DOI: 10.1007/978-3-319-06361-4_1, © Springer International Publishing Switzerland 2014

Newton and Dirac) a few decades ago, gave an inaugural lecture in which he argued that the end of theoretical physics was rapidly approaching?
SIMPLICIO: Exactly! Maybe Hawking was overly optimistic but he was substantially right. Is this not the culmination of the program that the great masters such as Galileo, Newton, Maxwell, Boltzmann and Einstein started, to explain Nature in terms of simple and universal principles?
SAGREDO: Honestly, I'm a little surprised by your enthusiasm for the "modern times". Wasn't your admiration reserved for the philosopher of Stagira, in the past? With all due respect, a scientist such as Hawking, who holds the Chair of the great Newton, should not narrow his horizons and neglect challenging topics, just for the sake of his own research interests.
SIMPLICIO: I am eager to know the arguments that call into question the philosophy of so many eminent scientists.
SAGREDO: Well, we can find many examples to show that things are not as simple as you claim. Allow me to mention one that I know well, because I have pondered it over for a long time. Perhaps it sounds unbelievable, but today we understand the atomic nucleus better than a cubic centimetre of turbulent atmosphere.
SIMPLICIO: I feel like you're teasing me, setting one of those traps that you often set for me during our first dialogue. What could possibly be so conceptually hard to understand in a fluid?
SAGREDO: As a matter of fact, the dynamics of fluids is extremely complex. If any of us were brave and skilled enough to deal with the tremendous mathematical difficulties of fluid dynamics, he could earn a tidy sum. Are you not aware of the American billionaire[1] who will pay one million dollars for the mathematical proof of existence and uniqueness of the solutions of the Navier-Stokes equations?
SALVIATI: If I remember well, these equations are Newton's laws for a continuous medium. For gases, they may also be obtained starting from the microscopic level of description, that is from a very large collection of interacting molecules, whose evolution is governed by the classical laws of dynamics.
SIMPLICIO: Well said! Of course there will be technical details to master, such as the theorem whose proof would be worth so much money, but in principle everything can be dealt with.

Actually, fluids are a perfect example, even though despised by some, in support of reductionism: the deepest level of description (that based on atomism and mechanical laws) determines the macroscopic level of description (fluid dynamics). Weinberg (1987) has summed up this point superbly:

> we understand perfectly well that hydrodynamics and thermodynamics are what they are because of the principles of microscopic physics.

[1] L. T. Clay, businessman from Boston, in 1998 founded the Clay Mathematical Institute (CMI). In 2000, to celebrate the 100th anniversary of the famous Paris conference in which D. Hilbert formulated 23 important problems in mathematics (almost all resolved today), the CMI announced seven awards (a million dollars each) for seven different issues, one of which is the proof of existence, uniqueness and regularity of the solution of Navier-Stokes equations in three dimensions. See Notices of the AMS, May 2000, p. 877.

There can be no doubt: in principle (if not in practice, owing to technical difficulties), everything is manageable.

SADREGO: Much as I am willing to listen, I cannot share your enthusiasm. Let me ask you a question: how could Euler and Bernoulli understand hydrodynamics so well in the eighteenth century that their results still hold? They had no clear understanding at all of the microscopic dynamics, since at their time there was no way of knowing how molecules interact.

SIMPLICIO: I am no historian of physics, and unfortunately I have not read the original works of Euler and Bernoulli. I'm willing to admit that not all details of the microscopic deterministic dynamics are important for the derivation of the Navier-Stokes equations. Nevertheless, I believe that premises such as the atomic hypothesis and Newton's laws are indispensable.

SAGREDO: You are only partly right. What is true is that many details of the dynamics are not essential, but this is not all: one may even abandon the atomic hypothesis.

SIMPLICIO: I am amazed; in the first dialogue, you were totally aligned with Galileo, one of the founders of modern atomism, and now you question the existence of atoms!?

SAGREDO: But seriously! How could anyone doubt the existence of atoms, after Maxwell's and Boltzmann's work on kinetic theory, after Einstein's, Smoluchowski's and Perrin's studies on Brownian motion? I merely referred to the findings of Frisch et al. (1986), who proved that the most complex hydrodynamic features can be reproduced by means of cellular automata.

SALVIATI: If I remember correctly, this artificial system does not fulfill some of the fundamental properties of the microscopic dynamics. For example, only microscopic discrete states are allowed: the "molecules" in this system move on a lattice and their velocities assume only a finite number of values. Moreover, rather than strictly deterministic rules, such as those of classical dynamics, the system follows probabilistic rules. I think this is a good example of what Kadanoff (1986) calls *two levels of reality*: the fluid dynamic level is practically independent of the microscopic level, from all points of view. It is also what Laughlin and Pines (2000) mean when speaking about "protectorates". Moving from the microscopic to macroscopic level of reality, it is possible to unfold several protectorates which know virtually nothing about the underlying levels. Basically, the dynamics of each of these levels are independent of the lower levels.

SIMPLICIO: I recognise that the work quoted by Sagredo is an example of great physical and mathematical virtuosity, but it does not affect in any way the focal point of reductionism. It only shows that, with regard to practical applications, hydrodynamics can be treated in conceptually diverse fashions. Yet, to me, that seems an exquisitely technical question. Fluids are certainly useful in applications, for instance in weather forecasting. But, as brilliantly stated by Weinberg (1995):

> We don't know the final laws of nature, but we know that they are not expressed in terms of cold fronts or thundestorms.

SALVIATI: You already quoted Weinberg with great emphasis before, though I think he meant something different from what you imply.

SIMPLICIO: Besides merely technical aspects, I think that the reductionist program is the only one that can save us from a "weak science", from forms of mysticism such as those expressed by the slogan "everything is greater than the sum of its parts."
SAGREDO: Are you referring to ideas such as those advocated by Lovelock (2007) with his "Gaia hypothesis", according to which our Planet is a sort of gigantic living system, where everything (the atmosphere, the ocean, glaciers, tropical forests, animals, and so on, up to human life) interacts with everything else?
SIMPLICIO: Exactly. This is the harm caused by abandoning reductionism: loss of unity in science, loss of predictive capacity. Even philosophers claim that the unity of science must be our working hypothesis Putnam and Oppenheim (1958)!
SAGREDO: Nevertheless, you cannot deny that very serious biologists expressed strong reservations against reductionism. Moreover, if the unity of science was very attractive some decades ago, today the disunity of science seems to be much more appealing, Fodor (1974) particularly when addressing issues such as life, consciousness and thought!
SALVIATI: Allow me, my friends, to make a proposal: let's limit our discussion to physics, or chemistry at most, in this discussion. This way, we shall safely avoid very delicate themes, which might not be considered properly scientific.
SAGREDO: Fine. It's better to avoid philosophical and religious discussions; we do not want to fall into invective or sensationalistic journalism. Yes, the analysis of reductionism in physics and chemistry bears relevant consequences in wider, even philosophical, contexts. However, I agree that it is appropriate to begin with natural sciences.
SIMPLICIO: Let us recall the words of Galileo:

> the book of Nature is written in the language of mathematics.

Following his teaching of *difalcare gli impedimenti*,[2] the study of Nature should aim to understand the fundamental laws that explain the apparent complexity of the World we see. This is necessary for the unity of science.
SAGREDO: Whether complexity is apparent or real, is still a highly controversial issue. In any case, with all due respect, nature is not perceived as a book by everyone.
SIMPLICIO: I am surprised by what you have to say. Please elaborate.
SAGREDO: Here is what Maxwell, who unified electricity with magnetism and optics and made great contributions to statistical mechanics, said in 1856 at a conference held at the Cambridge Apostles, which gathered the best students in Cambridge[3]:

> Perhaps the "book", as it has been called, of nature is regularly paged; if so, no doubt the introductory parts will explain those that follow, and the methods taught in the first chapters will be taken for granted and used as illustrations in the more advanced parts of the course; but if it is not a "book" at all, but a magazine, nothing is more foolish to suppose that one part can throw light on another.

[2] Eliminate hindrances.

[3] See the book of Campbell and Garnett (1882).

SIMPLICIO: To appeal to a youth conference of 1856 seems like a rhetorical trick! We agreed, I remind you, not to stoop to the level of journalism. Think of how much we have learned in the past centuries!

To make a long story short; what do you hope to prove by quoting the thoughts of famous physicists? Rather, let us not lose sight of the real problem and of the dangers of a science without fundamental principles. I have a relevant quote too: Goya admonished that:

> el sueño de la razon produce monstruos [4]

The monsters of holism and the new age are but two examples.

SAGREDO: I hope you do not believe that I am in favour of fads like the new age! Unfortunately I do not know the language of Cervantes, but a Spanish colleague of mine tells me that the word *sueño* means also dream: maybe Goya was deliberately ambiguous. It could also mean "the dream of reason produces monsters."

SIMPLICIO: Does reason produce monsters?! Could you please be less cryptic?

SAGREDO: I mean that the dream of reductionism may have produced the monster of "big science" and the consequent foolhardy race to reach ever smaller structures, passed off as more fundamental. The result of this rush are accelerators tens of kilometres long, built to study phenomena that occur on scales of 10^{-20} s. Yet, we still have no explanation of many things at the macroscopic level, on the scale of our everyday lives.

SIMPLICIO: Please, spare me the ideas of Thom, that we do not know how old walls crumble, and other banalities of this kind, which impress ladies on elegant lounges.[5] We are talking about the ultimate principles, the conceptual ones, not about merely technical difficulties.

The history of science teaches how physical problems were addressed within the framework of classical mechanics. Of course, this theory has been surpassed by the physics of the twentieth century, but its example still indicates the only way ahea perfectly.

SALVIATI: At this point, our readers would surely appreciate a small historical digression to illustrate your point, Simplicio.

SIMPLICIO: Wholeheartedly. What can be expected from a scientific theory? Obviously, that it describe reality through a mathematical model. Given the equations governing a particular phenomenon, and the corresponding initial state, one expects to predict the state of the system completely at any future time, by solving those equations with those initial conditions. Consider the motion of gravitating bodies, or the periodic behaviour of pendulums.

---

[4] The sleep of reason produces monsters.

[5] The quote that Sagredo has in mind is: *Many phenomena of common experience, in themselves trivial (often to the point that they escape attention altogether!)—for example, the cracks in an old wall, the shape of a cloud, the path of a falling leaf, or the froth on a pint of beer—are very difficult to formalise, but is it not possible that a mathematical theory launched for such homely phenomena might, in the end, be more profitable for science?* Thom (1993).

Certainly, things are not always so simple. However, nothing should in principle prevent the (at least approximate) solution of the evolution equations, by means of calculations of suitable difficulty.

SALVIATI: I suppose that you refer to Pierre-Simon marquis de Laplace (1829) and to his celebrated description of mathematical intelligence:

> We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes.

Supreme scientist Laplace, was a man of dubious moral stature and of astonishing opportunism. He succeeded in climbing up the social ladder under the monarchy, during the revolution, under Napoleon and, finally, under the Bourbon Restoration, eventually obtaining a title of nobility for himself.[6]

SIMPLICIO: Sometimes the giants of science are the pygmies of morality, but this sad aspect of human nature diminishes neither Laplace's greatness nor the merits of his mechanistic approach, which was one of the culminating achievements of the Enlightenment.

SALVIATI: Far be it from me to impugn Laplace's reputation, or the role of classical mechanics in the advancement of science. Please continue.

SIMPLICIO: I was going to say that the calculations may be far from simple. Nevertheless, could we forget the generations of astronomers who have computed with incredible patience and constancy the orbits of planets and asteroids from the fundamental equations of mechanics?

I cannot resist the temptation of telling you of the discovery of the planet Neptune. In the nineteenth century a series of observations indicated a significant deviation from the motion of Uranus' orbit foreseen by Newtonian mechanics. Adams in England and Le Verrier in France suggested that this discrepancy was not due to a deficiency in Newton's theory but to the presence of an unknown planet. Based on the laws of motion and gravitation, they calculated the location of this hypothetical planet, which was shortly after observed by Galle with a telescope.

The calculations that astronomers of the past performed with pen and paper are now carried out by computers and we can quickly predict the motion of celestial bodies and of natural satellites with high precision. We can almost affirm that Laplace's high mathematical intelligence, capable of submitting the data to analysis, exists today; it is not a skilled mathematician, but NASA's computing facilities. . .

SAGREDO: I didn't mean to provoke you by citing Thom; nevertheless I believe that these views are in stark contrast with everyday life, in which many events, unlike astronomical phenomena, do not seem to follow a predictable path. Consider, for instance, the weather, a falling leaf, or a rolling stone. How can we reconcile Laplace's fundamental assumption with the irregularity of most of our direct experience?

---

[6] It seems that he was actually ashamed of his humble origins.

SIMPLICIO: It seems to me, and it is of course just my point of view, that the problem can be solved attributing the irregularity of common evolutions to the impossibility of dealing with huge numbers of complicated equations analytically or numerically. I claim that these irregularities are merely "apparent", generated by myriads causes, each of which is, in fact, simple.

In short, one should strive to explain the complex visible world in terms of the simple invisible realm. In all problems that nobody can solve with just pen and paper, one may almost always nonetheless start from the basic principles and, using a sufficiently powerful computer, obtain the solution with the desired accuracy.

SALVIATI: This way to approach the problem, though shared in the past by distinguished physicists, seems outdated. Staying within fluid dynamics, one finds the theory for the onset of turbulence, proposed by the great Landau (1944). In that theory, the very complicated behaviour of a fluid was thought to result precisely from the superposition of many simple harmonic oscillations. However, this was proven mathematically inconsistent and quite in contrast with experiment.[7]

SIMPLICIO: This example only shows that even the greatest of scientists may be wrong.

SAGREDO: Of course everyone makes mistakes. This error is particularly significant for our discussion, however, because it seems that Landau, who was definitely a genius, was only caught in error two or three times in his entire scientific career. I mean that his was not a technical mistake, but a conceptual one, typical of physicists up until the 1960s. The difficulty was to conceive of a complex systems as something other than the sum of many simple subsystems. This attitude was in agreement with logical positivism, which had been the dominant philosophy for decades. It is important to observe that Landau, while apparently not particularly interested in epistemological issues, understood that qualitatively new laws emerge when passing from the mechanical microscopic level to the statistical macroscopic one [something that epistemology experts call "emerging properties" Nagel (1979), Humphreys and Bedau (2006)]. In their textbook on statistical mechanics, which generations of young physicists have spent many days, and sometimes nights, pondering, Landau and Lifshitz (1980) very clearly explain how new special laws appear in the presence of a very large number of particles constituting a macroscopic object. They called them called statistical laws.

The statistical laws are more than purely mechanical laws. Their specific feature is that they lose meaning in systems with a small number of degrees of freedom. Thus, although the particles obey the same mechanical laws, whether they belong to a large or small assembly, in the case of a large number of degrees of freedom, the collective behaviour follows qualitatively different rules.

SIMPLICIO: I'm sorry but I cannot follow your argument. All this discussion on fluids, beginning with unproven theorems that would make you rich, cellular automata that do not follow deterministic laws, the onset of turbulence and, now, statistical laws, thermodynamics and the very few mistakes made by Landau. All topics of great erudition, but were do they lead us? Let's return to modern physics.

---

[7] For a short discussion on the theory of turbulence see Cencini et al. (2009).

SALVIATI: If I understand him right, after having concocted this story, our friend Sagredo is now ready to move to modern physics and launch another attack.
SIMPLICIO: I hope it is not another of his rhetorical tricks. Don't you feel that the highest goal of scientific research should be to seek ever simpler principles which allow us to understand Nature in terms of a few simple laws obeyed by the ultimate constituents of matter?
SALVIATI: I guess you're referring to Weinberg's point of view. Together with Salam, Weinberg proposed a theory for the unification of the weak force (which is responsible for the decay of atomic nuclei) with electromagnetism.

> Perhaps our best hope for a final explanation is to discover a set of final laws of nature and show that this is the only logically consistent rich theory, rich enough for example to allow for the existence of ourselves.

SIMPLICIO: Indeed. After the unification of electromagnetic and optical phenomena, electromagnetism and weak interactions have also been unified. This constitutes the so-called Standard Model, whose predictions have been tested with great precision by skilled investigators like Rubbia. And soon we shall have the Theory of Everything, by which we will be able to account for all phenomena of physics and chemistry.
SAGREDO: Surely the Standard Model is a cornerstone of physics. But what exactly do you mean by: "we will be able to account for all phenomena"?
SIMPLICIO: Perhaps you remember what Dirac (1929) wrote about chemistry:

> now that quantum mechanics is well defined all the problems of chemistry are conceptually resolved. It suffices to solve the Schrödinger equation for the system of interest. Once the theory of everything has been established, we will be in the same situation, not limited to chemistry but open to all natural phenomena.

SALVIATI: I do not think that Dirac's famous statement can be taken literally.

As far as I know, the Schrödinger equation can be solved for the hydrogen atom, which has only one electron, while for helium, which has two electrons, it is impossible to find a solution. In this second case, subtle mathematical tools allow us to find good approximations of the Schrödinger equation. But for an atom such as lead, and even worse for systems with many interacting atoms, which are of interest e.g. in solid state physics, one faces practically insurmountable difficulties.

Apart from technical aspects that make it difficult to deal with the problem analytically and even numerically, the most interesting feature of many-body systems is perhaps the emergence of collective phenomena which differ substantially from the simple ones concerning the hydrogen atom. These emergent phenomena have never been treated in terms of the basic microscopic mechanical laws, but only in terms of "simplified" models.

I think that scientists continue to work in departments of chemistry, considering themselves chemists and not physicists, precisely for that reason: chemistry did not disappear and was not absorbed by physics, when physics discovered the laws governing the atomic realities.
SIMPLICIO: You, my good friend, do not understand. Obviously, the departments of chemistry will continue to exist, and chemists will continue to use techniques and

even concepts different from those of physicists. However, I cannot see how you can deny that, conceptually at least, chemistry is nothing but a difficult (albeit very difficult indeed) exercise in quantum mechanics.
SAGREDO: If I understand correctly, you recognise that chemistry is a science which differs from quantum mechanics. Indeed, even the departments of chemistry of our universities usually differ from those of physics. But you argue that, in some sense, physics subsumes chemistry.
SIMPLICIO: That's right; chemistry is a science describing a lower level than quantum mechanics. From the latter it is possible, in principle at least, to derive the laws of chemistry, but not vice versa. Basically, this is an instance of what the epistemologist Nagel (1979) calls inter-theoretic reduction.
SALVIATI: Epistemologists are particularly discerning. They investigated the issue deeply, distinguishing various forms of reductionism. I think that it's too delicate a topic for us, it is full of subtleties which may escape the inexpert. I think it is better to confine our dialogue to the context of physics, so that we may reach reliable conclusions at least on the foundations of physics. Later, at the end of our endeavour, we may try to put forward some speculations in a broader context.
SAGREDO: What is the purpose of having science Y reduced to science X, if it is impossible to derive Y from X? This impossibility, for which we have historical evidence, is indeed recognised even by champions of radical reductionism, such as Weinberg, who stated that:

> It seems that quantum chromodynamics is mathematically self-consistent, but it describes an impoverished universe in which there are only nuclear particles—there are no atoms, there are no people.

SIMPLICIO: You keep confusing the possibility of doing something in principle with that of doing it in practice.
SAGREDO: To be honest, I do not value reduction *in principle*, if nobody can carry it out in practice. Because the claim is unverifiable, to me it looks like a purely ideological stance. If I say that it is impossible to recover science X from to Y, and you say that it is possible but you cannot show me how, I have every right to be dissatisfied.
SIMPLICIO: Let's come back to the relationship between chemistry and quantum mechanics. As properly mentioned by our friend Salviati, it is impossible to solve explicitly the Schrödinger equation for a given atom, except in particularly simple cases. However, at least for atoms with few electrons and for some molecular systems, one can obtain good approximations of the solution, hence may conclude that all chemical properties can be traced back to the appropriate Schrödinger equation. For instance, approximation techniques allowed the calculation of the ionisation energy and other relevant properties for the helium atom, in remarkable agreement with experimental observations. I see no reasons to doubt that the same could be done (in principle) for even more complicated systems, such as solids.
SAGREDO: To tell you the truth, often chemistry and, more generally, condensed-matter physics, have been treated with little respect and even contempt by some physicists of renown. In response to the young R. Peirls' request for advice on the

prospect of entering the field of the structure of matter, W. Pauli, replied that to him the whole issue seemed a truly disgusting piece of crap! E. Rutherford distinguished: *physics or stamp collection*. Ironically, he was awarded the Nobel Prize for chemistry! Not to mention C. D. Anderson who, after the discovery of positron, adjudged that "*the rest is chemistry*".

SALVIATI: Forget the gossip and jokes! Let's return to our problem. What Simplicio says is partly right; it remains nevertheless true that in order to understand systems more complicated than helium, physicists and chemists have gone well beyond the methods of the quantum mechanics of Dirac times. Surely you know that, in 1998, the physicists Kohn and Pople were awarded the Nobel Prize for Chemistry, because of their computational techniques that allowed them to find the approximate solution of quantum systems with very large numbers of particles, but this…

SIMPLICIO: Precisely my point! Two physicists got the Nobel Prize for chemistry!

SALVIATI: Yes, of course. I was saying that in this context it is standard to use concepts of a statistical nature that arise because of the large number of interacting particles. In brief, to use something that is not part of quantum mechanics: the counterpart of the statistical laws, discussed by Landau and Lifshitz. I guess, even in this case, an epistemologist would say that we are dealing with emerging properties.

SIMPLICIO: All this fuss over emerging properties! I do not find anything mysterious about them at all. We have seen that in systems with many degrees of freedom, $N \gg 1$, the emerging properties are statistical laws. Well, I do not see why these properties cannot be explained from the laws which govern the behaviour of the elementary constituents of the system.

SALVIATI: Of course it is almost tautological to say that if you perfectly understand a system, then you also understand its emerging properties. The far from trivial point, here, is that emerging properties are not a mere consequence of laws that govern the physics of the system at lower levels: for the statistical laws to be derived, the equations of mechanics do not suffice: one has to resort to an ingredient which is not part of mechanics, which is $N \gg 1$. When we approach the "border" between two domains of science, we find ourselves in a strange no-man's land, where we need some guidance from the higher-level science, e.g. from the macroscopic theory, if we do not want to get lost, even though we are still walking in the microscopic landscape.

SIMPLICIO: I missed something in your argument. Let's go back to chemistry and help me understand in which sense Kohn and Pople's approach does not agree with Dirac's [1929] philosophy.

> The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be solved.

Today we have methods suitable for computing good approximations of such solutions. This a significant technical progress, isn't it?

SAGREDO: A weak aspect of Dirac's approach is that it underestimates the computational difficulties that make the "ab initio" reconstruction of all possible entities and chemical processes, starting from the dynamics of the elementary constituents (i.e. starting from the equations of quantum mechanics), impossible. More importantly,

Dirac did not understand that the terms of the fundamental theories do not suffice to recover the description of complex systems; for that, one has to introduce ingredients which are not included in the first principles.

SIMPLICIO: Pardon me, but I do not understand your argument. You claim that Dirac's reductionism was excessive; but it is a fact that experts of quantum many-body systems obtain results in excellent agreement with experiment, starting from the Schrödinger equation and with the aid of modern technical tools. Now, it is true that the description of complex systems which we have today is based on concepts not derived from first principles, which are called emerging properties by epistemologists. However, in my opinion, this is nothing but a technical artifice which leads to useful approximate descriptions, but ultimately these are the result of the dynamics of the elementary constituents of matter.

SAGREDO: Let me say that I do not believe that the reduction of theories has ever occurred, except in rare situations.

The essential point, in my view, is that the transition from a more fundamental theory to a more phenomenological one almost never happens as a mere process of approximation, such as that of the classical limit of classical mechanics from special relativity (Berry 1994; Primas 1998). To continue the discussion on chemistry and its connections with quantum mechanics, it is appropriate to highlight this technical point, which reveals that it is not just a matter of mathematical simplifications. Indeed, to obtain the laws of chemistry, the basic quantum mechanical theory has to be replaced by another theory which, although shaped from the original one, is qualitatively different.

Consider, for instance, one of the most common procedures in molecular physics Primas (1998): the *Born-Oppenheimer approximation*. In a nutshell, it starts from studying the nuclei neglecting the presence of their electrons and proceeds by solving the Schrödinger equation for the electrons in the field generated by the nuclei. The physical motivation of this procedure rests on the fact that the mass of the nuclei is much larger (a few thousand times) than that of electrons. This approximation is the basis of the concept of molecular structure. For instance, the benzene molecule in the ground state is assigned a hexagonal structure whose vertices are occupied by the nuclei of the different chemicals. Then, the electrons are delocalised and move around this structure. This does not follow from quantum mechanics; as a matter of fact, because of the Einstein-Podolsky-Rosen correlation between electrons and nuclei, this scenario is inconsistent with quantum mechanics.[8] Therefore, the term *approximation*, in this case, is misleading: technically, the Born-Oppenheimer procedure is the result of a *singular limit* (Berry 1994; Primas 1998) obtained taking the $m/M \to 0$ limit, where m and M are, respectively, the mass of the electron and of a nucleus. This is conceptually different from solving the original Schrödinger equation for the complete system made of nuclei and electrons in interactions Scerri (2008).

SALVIATI: Despite Dirac's reputation, many chemists are not impressed by his views on these matters. The reduction of chemistry to physics appears impossible to

[8] See Chap. 6 for a short discussion of the Einstein-Podolsky-Rosen paradox.

them also because, to date, quantum mechanics has almost exclusively produced descriptions and rationalisations of results previously obtained by other means. Provocatively, Woolley (1978, 1986) raised the question *Must a Molecule Have a Shape?*, stressing that the concepts used in chemistry (e.g. the length of a chemical bond) are inconsistent with quantum mechanics.

SAGREDO: The theoretical chemist Primas has studied with great care and deep critical insight the implications of quantum mechanics for the foundations of chemistry. In disagreement with Dirac, he does not believe at all— not even in principle —that chemistry is applied quantum mechanics. According to Primas, in general this theory cannot predict the molecular structure observed experimentally, even if used in its "exact" form, e.g. without approximations such as the Born-Oppenheimer approximation. For instance, quantum mechanics would deny the existence of pyramidal molecules such as AsH_3.

SIMPLICIO: I am astonished! Do you mean that the existence of certain molecules invalidates quantum mechanics?

SAGREDO: This is a rather subtle issue. In the case of the molecule AsH_3, quantum mechanics provides a delocalised configuration, rather than the experimentally observed configuration, with the nuclei around the vertices of a tetrahedron (a triangular based pyramid). According to some authors, such as Woolley and Primas, molecular structure is an emerging property, which results from the interaction between different molecules. A beautiful paper of Claverie and Jona-Lasinio (1986) shows how small perturbations (which could be due to the interaction with the environment) produce the localised structure known in chemistry. We can therefore conclude that the molecular structure cannot be obtained by a "naive" application of quantum mechanics, i.e. merely from the Schödinger equation for the nuclei and electrons of the molecule. The interaction of the molecule with the external environment, which causes the emergence of the known classical properties, must also be taken into acount.

SALVIATI: At this point, it is plain that the disagreement between you, my friends, is not just terminological but substantial: your own viewpoints are totally different and incompatible, I would say. Allow me summarise Simplicio's.

I understand that you, Simplicio, believe that the structure linking the various sciences $X_1, X_2, \ldots, X_n, \ldots$ (ordered from less to more fundamental) is both hierarchical and inclusive:

$$X_1 \subset X_2 \subset \cdots \subset X_n \subset \cdots$$

Furthermore, you think that a real scientific explanation is found only when theory X_n (chemistry for example) is reduced to the more fundamental X_{n+1} (e.g. quantum mechanics).

SIMPLICIO: You have perfectly represented my thought... which is certainly not mine alone. Again, to express this view, I cannot find better words than Weinberg's

> One can illustrate the reductionist world view by imagining all the principles of science as being dots on a huge chart, with arrows flowing into each principle from all other principles by which it is explained... they are all connected, and if followed backward they seem to branch outward from a common source, an ultimate law of nature.

SALVIATI: Sagredo, now is your turn. Illustrate concisely your point of view.
SAGREDO: Unlike our friend Simplicio, I do not believe in a hierarchical structure of the various sciences. I find more appropriate to talk of a partial overlapping between "close" sciences, and an almost complete disconnection between "distant" sciences

$$X_{n-1} \cap X_n \neq \varnothing; \quad X_n \cap X_{n+1} \neq \varnothing; \quad X_{n-1} \cap X_{n+1} \simeq \varnothing.$$

I am well aware that this type of structure dissatisfies the *fundamentalists* of ultimate truth. In my opinion, it is precisely because ''close" sciences overlap only partially that science exists and that the great scientists of the past could say something at a given level of description of phenomena (e.g. hydrodynamics), which retains its validity today, despite the fact they did not have a perfect grasp of more fundamental levels of description (e.g. the molecular dynamics). Moreover, at the borders between different disciplines often the most interesting features appear. Think of the semiclassical limit of quantum mechanics (Bokulich 2008), for example. Even in quantum field theory, often considered a paradigmatic example of the reductionist approach, the actual research work is based on the existence of different levels of description. In that realm, one is forced to adopt different effective theories for different energy scales (Castellani 2002).
SALVIATI: As far as I understand, you share Anderson's thoughts (Anderson 1972). He too was awarded the Nobel Prize, but defends positions very different from Weinberg's. For instance, he argues that:

> The ability to reduce everything to simple fundamental laws does not imply the ability to start from those laws and reconstruct the universe.

Why don't you elaborate further, to explain your view in greater details? Help me understand: when you say that the various sciences are almost disconnected, are you alluding to a real property of nature, or to a sort of necessary construct which we are forced to adopt in order to make, say, approximate predictions?
SAGREDO: I believe that these are two almost inseparable aspects of our understanding of things. Fortunately, Nature has a hierarchical structure with different length, time and energy scales and this is precisely why we can investigate and describe (almost) separately, and with satisfactory accuracy, the different levels of reality.

I cannot really say whether this is a necessity inherent in the structure of our universe, or a mere (indeed fortunate) coincidence.

Some cosmologists have recently been struck by the amazing coincidences thanks to which life on Earth is possible. Among these, we may recall the distance of our planet from the Sun, the flow of energy absorbed and emitted by the Earth and, especially, the values of the fundamental constants of physics, such as Planck's constant, the gravitational constant, the electron and proton masses, etc. There is a school of thought whose followers do not believe that all this is pure coincidence, but should be made part of the laws of physics, in terms of a universal principle, which has been called the anthropic principle (Barrow and Tipler 1988). This school attributes fundamental importance to the fact that all characteristics of the physical universe and of us, intelligent beings, crucially depend on the precise numerical values of

the universal constants. Minimal variations in their values would have prevented the formation of an environment suitable for life as we know it.

SALVIATI: I am not completely convinced by these arguments. Not being a cosmologist, perhaps I am too naive, but I think that this presents a kind of circular argument: without these fortunate "coincidences", there would be no life on Earth, hence we would not be here to debate whether the anthropic principle is a profound truth or not. Analogously, had Nature not had a hierarchical structure with largely independent levels of reality, I believe that it would never have had what we call science. Let's admit it: we have been lucky to be born on a "cold" planet, in which many phenomena are approximately linear and easy to formalise in mathematical terms. This has certainly favoured the development of deterministic concepts, which are the starting point of modern science.

On a hot celestial body, such as the Sun, or in a world with a prevalence of chaotic and irregular phenomena, a hypothetical intelligence compatible with those conditions might not have ever developed science and, in any event, regularities and determinism would have deserved much less attention.

SIMPLICIO: I think I begin to understand what you mean by the different levels of reality and of the corresponding descriptions which have been developed. To tell the truth, it seems to me that you have given evidence only of the partial overlap of chemistry and quantum mechanics. What about the science you like so much, fluid dynamics? You have only argued, referring to recent studies, that the laws governing the behaviour of fluids can be obtained even without recourse to microscopic dynamics.

I would like to better understand the extent to which your idea of the various levels of description is general and does not merely imply that there is a microscopic world (accessible only to scientific eyes) and a macroscopic one immediately accessible to our senses, as Democritus had already intuited long ago.

SAGREDO: Then, let me expand on fluid dynamics. You know that the behaviour of our climate is of great importance for human activity.

In principle, we know how to study a complex phenomenon such as the fluctuations of the aperiodic climate in tropical regions, commonly known as El Niño. We should start from the equations of fluid dynamics for the velocity field, the pressure, the temperature, the density, the humidity etc. of small volumes of the atmosphere and of the oceans. Thermodynamics must of course be included in the description. The solution of all the relevant equations describes the evolution of atmosphere and ocean as interacting systems. This however is not feasible for two reasons: in the first place, the formidable calculations are actually impossible to perform. Second, even if one day we could do that, the results we would obtain would prove useless. This may seem paradoxical, but like in grasping the message of a movie, we have to neglect a large fraction of details contained in the film, such as the position of the single grains of color of each frame, so to understand climate we have to dispose of a large amount of unnecessary information.

It appears that Von Neumann, one of the founders of modern computer science, had already realised this: his pioneering attempts at numerical studies of the evolution of the atmosphere were carried out by solving the so-called quasi-geostrophic

equations, which encompass a number of (non-trivial) assumptions and approximations (Charney et al. 1950), rather than the primitive equations. However, even these equations, though much simpler than the original ones, are still too complicated and detailed to efficiently describe El Niño. Therefore, experts go one step further and introduce effective equations built upon decades of experience of geophysicists and mathematicians (Imkeller and Von Storch 2001). In summary, we are not dealing with the simple

$$\text{microscopic (molecules)} \rightarrow \text{macroscopic (fluid)}$$

scheme; rather we have a more articulated situation, which can be schematised as follows:

$$\text{Molecular level} \rightarrow \text{fluid dynamics} \rightarrow$$
$$\text{quasi-geostrofic equations} \rightarrow \text{effective equations(El Ni}\tilde{\text{n}}\text{o)}$$

SIMPLICIO: I recognise that this approach is practically very useful and efficient for describing non-trivial phenomena that occur on spatial scales much larger than molecular ones and with typical times much longer than microscopic times. However, let me point out that your example, in fact, is but an realisation of what Weinberg said, that is theories connected by arrows (which involve the reduction of a theory to another) starting from the most fundamental one.
SAGREDO: Once again I must contradict you! The transitions between different levels of description are not simple at all. Mind you, I am not talking of technical difficulties, which may be practically insurmountable but conceptually obvious. What I mean is that almost every transition step introduces something technically and conceptually absent from the theory of departure.

In principle, given the equations of the phenomenon to be studied, e.g. the quasi-geostrophic equations, one could proceed by brute force, using numerically efficient and powerful machines. Some believe that this is the only truly valid approach. There are practical difficulties that are hard to overcome at present, such as those concerning the insufficient memory and speed of current computers, for those kinds of calculations. Therefore, if there is no other valid alternative, one either waits for the advent of sufficiently powerful computers, or introduces appropriate simplifying methods, including special numerical methods or parametrisation techniques, which allow one to save computation time and memory.

The proper orthonormal decomposition (POD) method of Holmes et al. (1998), which leads to effective equations, is a semi-empirical procedure that exploits in a self-consistent fashion (also called a bootstrap) both experimental observations and the full set of detailed equations. I cannot enter into technical details, here. They are, indeed, rather complicated. However, to describe in a mathematically consistent way the time evolution of turbulent coherent structures, the idea is to limit the analysis to the *cold fronts*; something that Weinberg (1995) mentioned sarcastically as a nuisance, as far as our understanding of the laws of nature is concerned.

To be honest, I'm afraid the previous discussion may have mistakenly suggested that the transition from one level of natural phenomena description to another, which is highly complicated both technically and conceptually, is only a one-way street, which goes from the more fundamental to the less fundamental.

The actual situation is rather more complex and the arrows connecting different theories sometimes are not unidirectional at all. I think the most important example in this context is afforded by Brownian motion.

SIMPLICIO: I'm really surprised. You know how the theoretical studies carried out by Einstein and Smoluchowski , and the later experimental work by Perrin, convinced even the most irreducible opponent of the atomistic hypothesis, allowing a direct determination of Avogadro's number.[9]

In my eyes, the explanation of Brownian motion from kinetic theory is a perfect example of the success of reductionism: to explain complicated visible phenomena in terms of simple invisible elements.

SAGREDO: I do not think that things are so simple. How is it that outstanding scientists like Maxwell and Poincaré knew of Brownian motion, but were not able to interpret it in the context of kinetic theory? Others didn't even understand its importance. Only in the early twentieth century did Einstein and Smoluchowski, who were young men at the time, formulate the brilliant and very audacious hypotheses which finally allowed the problem to be solved (Chandrasekhar 1943).

SALVIATI: The argument is so important and beautiful that a brief introduction will be useful to our readers. Avoiding too much detail, let us highlight the importance of the hypotheses proposed by Einstein and Smoluchowski, and explain the historical significance of Brownian motion within the disputes between atomists and energetists.

SAGREDO: Brownian motion is one of the topics of theoretical physics that I treasure the most, so I accept your invitation with great pleasure.

In 1827, the Scottish botanist Robert Brown observed under a microscope that pollen grains suspended in water show a rapid and irregular movement. At first, he thought that the phenomenon was linked to the organic nature of pollen grains; however it was soon realised that the same phenomenon occurred even for small pieces of glass or stone. After the initial indifference of physicists, and some interpretation in terms of temperature differences quickly abandoned, the importance of Brownian motion and its connection with thermodynamics was realised. In particular, it was noticed that the speed of motion grew with increasing temperature and with decreasing grain size, whereas it decreased with the viscosity of the fluid. The dependence of the velocity to the grain temperature suggested a possible link between Brownian motion and the kinetic theory of heat. On the other hand, a simplistic explanation of the irregular motion of the pollen as a result of molecular collisions, could not be trusted, because the smallest observable particle is in any case vastly heavier than molecules are. Therefore the individual impacts of molecules on pollen grains could not possibly justify the observed motion.

[9] For a clear discussion of Brownian motion see Chandrasekhar (1943). Einstein's papers have been translated in Einstein (1956).

At this point, Einstein and Smoluchowski put forward their revolutionary theories, which were later elaborated upon by Langevin, in terms of stochastic differential equations (using modern terminology). In short, following Langevin's formulation, the idea is that the motion of the pollen grains is determined by two contributions:

(a) the friction in the fluid; and
(b) the molecular impacts.

Contribution (a) is expressed by Stokes' law. In modern terminology, contribution (b) is assumed to be described by a Gaussian process, with variance obtained from kinetic theory, and without temporal correlations. These hypotheses of normal distribution and immediate loss of correlations are physically justified by kinetic theory, and by the fact that the ratio of the typical times of the Stokes and microscopic contributions is very large ($10^4 - 10^5$).

Once the phenomenon is formalised in this way, everything is clear, and the Einstein-Smoluchowski law of diffusion is easily obtained:

$$\langle \mathbf{x}(\mathbf{t})^2 \rangle \simeq 6Dt, \quad D = \frac{k_B T}{6\pi \eta R},$$

where $\mathbf{x}(t)$ is the position of the particle at time t, k_B is the Boltzmann constant, which is the ratio of the the constant of gases $\mathscr{R}$ and Avogadro's number N_A, $k_B = \mathscr{R}/N_A$, η is the viscosity of the fluid and R the radius of the particle.

The Einstein-Smoluchowski law is so important because it links the easily experimentally measurable value of the macroscopic quantity D, with Avogadro's number. The determination of N_A from the measurement of D, i.e. from the observation of the spreading of pollen suspended in a liquid, and its agreement with values obtained independently, definitively closed the heated controversy about the existence of atoms between Boltzmann, on one side, and Mach and Ostwald, on the other.

SIMPLICIO: I greatly appreciate your historical reconstruction, but I'm still puzzled. I don't understand why you maintain that the works of Einstein, Smoluchowski and Langevin do not support the reductionist approach.

SAGREDO: Clearly I didn't succeed in demonstrating the depth and ingenuity of these achievements. I will try to be more explicit. In the discussion of Brownian motion, different levels of description were mixed up several times. First, the force acting on pollen grains was assumed to be made up of two contributions: one of macroscopic and one of microscopic origin. Moreover, the macroscopic term was given by Stokes' law, in which the viscosity may be determined via kinetic theory (which describes a microscopic realm).

In short, the poor pollen grain is asked to do something seemingly absurd: to behave in accordance with both kinetic theory and hydrodynamics. This is the ingenious idea developed by Einstein and Smoluchowski. Their bold "bootstrap" procedure proved definitively that the atomic hypothesis was correct. However, this is not the explanation of a complicated visible world in terms of simple invisible elements. If anything, this is a rather intricate interweaving of observable macroscopic and indirectly observable microscopic levels of reality.

It is also significant that the theory suggested to experimentalists like Perrin that the quantity to be measured was the mean square displacement and not the velocity, as may have seemed natural within the framework of kinetic theory.

SALVIATI: Cercignani (2006) noted that Brownian motion, one of major chapters of modern physics, has so been poorly understood, and even misconstrued as incompatible with the second law of thermodynamics, by leading philosophers of science, for instance Popper (1958) stated that Brownian motion is a serious problem for the second law. It was well known to both Maxwell and Boltzmann that the second law of thermodynamics applies to macroscopic objects and not to systems with few degrees of freedom:

> Hence the second law of thermodynamics is continually violated, and that to a considerable extent, in any sufficiently small group of molecules belonging to a real body (Maxwell 1878, 2011).

SAGREDO: From what you say, I understand that the essential fact to be underlined is the *fortunate* enormous separation between the microscopic and macroscopic scales. The mass of the pollen grain is so much bigger than that of the individual molecules of the fluid. Likewise, the characteristic relaxation time determined by Stokes' law is so much longer than the typical time of molecular collision. Without this separation of scales, Brownian motion would simply not exist, because it would be nothing more than the motion of a molecule among many other molecules. It would not be easy (perhaps impossible?) to formulate a statistical theory such as that given by the Langevin stochastic differential equation.

SALVIATI: You have grasped perfectly the essence of the passage from one level of description to another, which is also called—misleadingly, in my opinion—reduction of a theory. As previously seen in the framework of chemistry, with regard to the Born-Oppenheimer "approximation", the transition from one theory to another in the case of Brownian motion is linked to a small dimensionless parameter δ, the ratio between the mass of the molecule and that of the pollen grain. The reduction of the theory is not at all achieved through an approximation, but rather through a non-trivial limit: a singular limit. As noted by Berry (1994) and Primas (1998), the emergence of new properties in the transition from one level to another lies exactly in this singularity.

References

Anderson, P.W.: More is different. Science **177**(4047), 393 (1972)

Barrow, J.D., Tipler, F.J.: The Anthropic Cosmological Principle. Oxford Univerisity Press, Oxford (1988)

Berry, M.V.: Asymptotics, singularities and the reduction of theories. In: Logic, Methodology and Philosophy of Science IX, pp. 597 (1994)

Bokulich, A.: Reexamining the Quantum-classical Relation. Cambridge University Press, Cambridge (2008)

Campbell, L., Garnett, W.: The Life of James Clerk Maxwell. MacMillan and Co., London (1882)

Castellani, E.: Reductionism, emergence, and effective field theories. Stud. Hist. Philos. Mod. Phys. **33B**, 251 (2002)
Cencini, M., Cecconi, F., Vulpiani, A.: Chaos: From Simple Models to Complex Systems. World Scientific, Singapore (2009)
Cercignani, C.: Ludwig Boltzmann: The Man Who Trusted Atoms. Oxford University Press, New York (2006)
Chandrasekhar, S.: Stochastic problems in physics and astronomy. Rev. Mod. Phys. **15**(1), 1 (1943)
Charney, J.G., Fjörtoft, R., Neumann, J.: Numerical integration of the barotropic vorticity equation. Tellus **2**(4), 237 (1950)
Claverie, P., Jona-Lasinio, G.: Instability of tunneling and the concept of molecular structure in quantum mechanics. Phys. Rev. A **33**(4), 2245 (1986)
Dirac, P.A.M.: Quantum mechanics of many-electron systems. Proc. R. Soc. Lond. Ser. A **123**(792), 714 (1929)
Einstein, A.: Investigations on the Theory of the Brownian Movement. Dover Publications, New York (1956)
Fodor, J.A.: Special sciences (or: the disunity of science as a working hypothesis). Synthese **28**(2), 97 (1974)
Frisch, U., Hasslacher, B., Pomeau, Y.: Lattice-gas automata for the Navier-Stokes equation. Phys. Rev. Lett. **56**(14), 1505 (1986)
Holmes, P., Lumley, J.L., Berkooz, G.: Turbulence, Coherent Structures, Dynamical Systems and Symmetry. Cambridge University Press, Cambridge (1998)
Humphreys, P.E., Bedau, M.A.: Emergence: Contemporary Readings in Philosophy and Science. MIT Press, Cambridge (2006)
Imkeller, P., Von Storch, J.S. (eds.): Stochastic Climate Models. Birkhäuser, Boston (2001)
Kadanoff, L.P.: On two levels. Phys. Today **39**, 7 (1986)
Landau, L.D.: On the turbulence. Dokl. Akad. Nauk. SSSR **44**, 339 (1944)
Landau, L.D., Lifshitz, E.M.: Statistical Physics, Part 1. Pergamon, New York (1980)
Laplace, P.S.: Essai philosophique sur les probabilités. H. Remy (1829)
Laughlin, R.B., Pines, D.: The theory of everything. P.N.A.S **97**, 28 (2000)
Lovelock, J.: The Revenge of Gaia: Earth's Climate Crisis and the Fate of Humanity. Perseus Books Group, New York (2007)
Maxwell, J.C.: That's thermodynamics. Nature **17**, 257 (1878)
Maxwell, J.C.: The Scientific Papers of James Clerk Maxwell. Cambridge University Press, Cambridge (2011)
Nagel, E.: The Structure of Science. Hackett Publishing Company, Indianapolis (1979)
Popper, K.R.: Irreversible processes in physical theory. Nature **8**, 51 (1958)
Primas, H.: Emergence in exact natural science. Acta Polytech. Scand. **91**, 83 (1998)
Putnam, H., Oppenheim, P.: Unity of science as a working hypothesis. Minn. Stud. Philos. Sci. **2**, 3 (1958)
Scerri, E.R.: Collected Papers on Philosophy of Chemistry. World Scientific, Singapore (2008)
Thom, R.: Structural Stability and Morphogenesis. Westview Press, Boulder (1993)
Weinberg, S.: Newtonianism, reductionism and the art of congressional testimony. Nature **330**(6147), 433 (1987)
Weinberg, S.: Reductionism redux. N.Y. Rev. Books (1995)
Woolley, R.G.: Must a molecule have a shape? J. Am. Chem. Soc. **100**(4), 1073 (1978)
Woolley, R.G.: Molecular shapes and molecular structures. Chem. Phys. Lett. **125**(2), 200 (1986)

Chapter 2
A Random Journey from Monism to the (Dream of) Unity of Science

If the world were clear, art would not exist.

Albert Camus

2.1 History

Among the founders of Western culture, the Presocratic thinkers sought explanations of the natural world that required only one or very few fundamental principles or substances. Hence, Presocratic philosophers were the first (in Western culture, at least) to search for one substance capable either of explaining every natural phenomena or of unifying all phenomena within a single framework. In this endeavour, they put forward the first reductionist vision of the world, which we may interpret as a search for some kind of harmony in the world and, therefore, for a unifying picture. This idea has largely dominated scientific progress and development ever since and proven undeniably fruitful. Their motivation was that the marvelous and amazing variety of nature, experienced by the human senses, needs some kind of order to be appreciated and analysed. In particular, the Presocratics were attracted by a form of monism which may be considered the philosophical precursor to the most extreme version of reductionism. This search for first-principles explanations of "everything" has greatly influenced the subsequent developments of science and knowledge and, in particular, has become a cornerstone of physics, until the present day (Klein and Lachièze-Rey 1996).

Thales proposed water as the primary element: he thought of water as the material origin of everything. For the first time, an explanation of all phenomena was proposed in terms of a unique element, creating a departure from the mythic tradition, which had multiplied causes and ad-hoc explanations. The tragic expedient of the *Deus ex machina* is a splendid illustration of the former approach. Thus, for the first time, Thales expressed three ideas which remained thereafter at the core of scientific endeavour:

S. Chibbaro et al., *Reductionism, Emergence and Levels of Reality*,
DOI: 10.1007/978-3-319-06361-4_2,

- the search for the material causes, or mechanisms, underlying natural phenomena;
- the use of rational arguments to understand the world; and
- the reduction of answers to few or just to a sole fundamental substance.

Following this path, Anaximenes proposed a different cosmogony, in which air was regarded as the fundamental substance. Interestingly, from a methodological point of view, he put forward a logical framework in which every known natural phenomenon (earth, wind, clouds, etc.) could in principle be deduced from his primary substance.

A while later, Heraclitus chose fire as the fundamental element. He preferred an intrinsically more unstable element, in order to underline the dynamical features of the universe: evolution and becoming. Heraclitus clearly sided with "*becoming*" rather than with "*being*". To complete his cosmogony and explain the huge variety of natural phenomena, he complemented his fundamental substance with a fundamental law: a dialectic law asserting that a hidden conflict between opposites pervades nature. This fight was supposed to explain the observed harmony and stability of the world, despite the unstable nature of the fundamental element. As pointed out by Heisenberg (1958), Heraclitus' view appeals to contemporary physics: *if we replace the word "fire" by the word "energy" we can almost repeat his statements word for word from our modern point of view.*[1]

Some years later, Empedocles closed out the pioneering era of the monistic approach, proposing a multiplicity principle, which adds earth, as a fundamental element, to the elements considered by his predecessors. Water, air, fire and earth thus gave birth to the four-element cosmogony, which prevailed in the christian view of the whole Middle Ages, until the Renaissance. Interestingly enough, the idea of a small family of elements responsible for all phenomena constitutes the standard approach of theoretical particle physics, for which the whole world is made up of three families of elementary particles (6 quarks + 6 leptons).

Still in the Classical period, philosophers envisioned another concept which complemented the primordial form of monism described above and which remains at the heart of modern science: considering an abstract principle as a fundamental substance. This conceptual progression was achieved in the sixth century B.C. by Anaxagoras, who advanced intelligence as the first principle and organising force of the universe. At the same time, the Ionic philosopher Anaximander abandoned the idea of fundamental natural elements and made another leap forward, concluding that the fundamental substance should be a complete abstraction. To this purpose, he proposed the $\alpha\pi\epsilon\iota\rho\sigma\nu$, the not-finished or the infinite. This was the first attempt to define an ultimate *One*, as the fundamental and primordial reason of everything. New insights into the relation between the One and the Multiple were given by Parmenides who declared: "*the being is, the not-being is not*" (Diels and Kranz 1951–1952). In this way, the metaphysical framework was introduced, for the first time, and lasted almost intact until modern times. For instance, the laws of physics are thought to be

[1] With an even more audacious analogy, we can see the conflict between opposites essentially as the *in fieri* quantum wave/particle duality.

eternal[2]: their constancy permits a unitary vision of nature. Unstable and fluctuating truths do not allow a rational prediction of phenomena. This vision was shared by Poincaré (1913), who thought it necessary that the laws of nature not be subject to modification. He proposed this view under a methodological assumption, but he admitted that these laws ought to be generalised or transformed in case they were falsified.

Another approach and *Weltanshauung* was developed in parallel to the monistic approach: atomism. Leucippus and Democritus were the foremost proponents of this concept. While atomism avoided the definition of a unique fundamental substance, it established for the first time a reductionist approach to nature, in which every natural phenomenon is explained through the features of a small set of elementary objects.[3]

On the other hand, other great philosophers showed an anti-reductionist attitude, even mocking the reductionist attempts to describe the world. Among these, Socrates was probably the most influential and the most scathing: *It is strange that men wanted to understand the first principles of things, and from that understand everything, with conceit as infinite as their object of study* (Klein and Lachièze-Rey 1996).

The fundamental role of numbers is another important idea that has been most influential from ancient times through the Middle Ages. First underlined by Pythagoreans, the essential relevance of numbers as creative principles has been long recognised from a Mystic point of view. While the Mystic approach has not given important scientific results, it has been significant in the methods of modern scientists like Einstein and Dirac. They admitted the creative and scientific role of mathematical aesthetics and formal consistency, together with the empirical verification. It is plain that the search for a unified and harmonious account of the world has always animated and motivated scientists' and philosophers' works, even if in most diverse fashions. Furthermore, this metaphysical spirit transformed into a search for order and eventually for unity in nature, hence can be seen as one of the sources of the reductionist approach. A striking example (as well as one of the most important) of the step from the Mystic to the scientific point of view is afforded by Kepler, who significantly entitled one of his main works "*harmonia mundi*". Kepler desired primarily to point out the harmonious side of nature, by discovering the laws of the motion of planets. In order to accomplish this aim, he mingles mathematics, which continues to be the first principle underlying nature, with geometry, music, astrology and astronomy. In this way, he was able to reduce to a small set of statements (in this case three mathematical laws) a complex and in some sense universal natural behaviour (planetary orbits). This work prepared the formalisation established by Galileo, Descartes and Newton who, disposing of the spiritualistic and metaphysical elements, moved toward an actual reductionist vision of science. Ultimately, Kepler's work definitively sanctioned the link between mathematics and the harmony investigated by science, while dispensing with Mystical visions.

---

[2] This vision is not naive and is compatible with possible major changes occurring in nature on the scale of the age of the universe.

[3] Of course, it is meaningly to try to establish whether Leucippus or Democritus were "reductionists" in modern terms. Our historical references should be considered *cum grano salis*.

Another important view emerges along with these developments of modern science. The synthesis motivated by the search for harmony in the world accompanies the attempt to unify our knowledge of all phenomena, as disparate as they may be. Thus, it is possible to appreciate how reductionism is intimately related to unification. An example of the unifying qualities of the quest for harmony is given by the Balmer series (Cartier 1995),[4] proposed to give order to a series of experimental results, reducing them to a rigorous and elegant mathematical formula. Yet, the harmony underlying this formula was recognised by Bohr in the quantum atom, and the Balmer series has been shown to be a part of the more general (and thus unifying) theory of quantum mechanics. It is interesting to note that even the quantum theory of fields, in particle physics, appears in some sense to remain in the wake of this harmony representation of the world, although it is an atomistic theory by definition. Indeed, symmetries are used as principles of mathematical harmony, related to different kinds of particles.

Loosely speaking, two main streams have animated the scientific and metaphysical developments since the ancient Greek philosophers, in very different forms: on the one hand, the demand for unity and order (in some sense the necessity of a reduction of phenomena to basic simple levels); on the other hand, the desire and search for harmony in the world, also with regard to the existence of many levels of complexity. Several further important elements have been put forward since then: the importance of intuition and creativity in the process of scientific and philosophical thinking, as well as the formidable instrument of unification that is mathematical language (Jona-Lasinio 2005; Dorato 2010).

2.2 Reductionism: The Philosophical Point of View

It is impossible to discuss, even only superficially, the vast literature concerning the subject of reductionism and liminal domains. In order to complete the present brief overview, we direct the reader to the following important references (Humphreys and Bedau 2006; Boyd and Gasper 1991; Adler et al. 2002).

2.2.1 General Introduction

As suggested by the historical sketch above, reductionism's roots can be traced back to the desire to unify different parts of the same science or even different sciences, within a more general scheme encompassing them. Thus, reductionism and the unity of science have commonly been associated with each other. In particular, the idea and strategy was usually to reduce some higher-level science (such as biology) to

[4] It relates the classification of spectral rays of hydrogen following the frequency, afterwards generalised by Rydberg. It remains true in a given approximation.

a lower-level science such as chemistry or physics; or, more basically, to reduce a theory to some other theory, considered more fundamental. Paradigmatic examples in physics include the relations between classical thermodynamics and statistical mechanics (Nagel 1979), between classical Newtonian physics and the theory of relativity, or between classical mechanics and quantum mechanics (Dirac 1929). Thus, philosophers of science and scientists (who have often been the same persons until as recently as the first half of the nineteenth century) have usually been concerned with inter-theoretic relations (i.e. with relating theories of one domain with those of some other domain). This form of reductionism can be therefore called "epistemic reductionism". This clarifies an important point, related to the epistemic reductionistic approach but, at the same time, distinct from it: since the birth of modern thought, which for philosophical and scientific purposes can be traced back to the rationalist work of Galileo and Descartes, it has seemed natural and obvious to regard the world as being hierarchically structured. That means that we implicitly consider nature to be structured in several more or less fundamental levels, which are related to each other. It is worthwhile reflecting on the fact that every modern attempt to understand natural and human facts or evidence is associated with the effort to classify them, in order to make some sense out them.[5] Indeed, it seems that taxonomy is necessary to understand a given phenomenon, at least in western culture. This metaphysical disposition was further formalised by Descartes (1987), who proposed a rigorous method to face reality and its problems. He put forward a recipe based on the following metaphysical analytical approach: every problem or phenomenon should be decomposed into its smallest parts that preserve its properties. Here the idea is that it is always possible to dissect a phenomenon into smaller pieces, keeping the sum of the parts equal to the whole. The method consists in subdividing the problem into parts which are individually tractable, in order to eventually solve the original problem. This strategy can be related in some sense to reductionism, yet it is simply a manifestation of the human failing of not being able to deal with too many different issues at the same time. This approach may be seen as the basis of so-called methodological reductionism, which states that the only (or best) way to generate scientific knowledge is to decompose complex problems into simple tractable ones. However, it does not claim to be a philosophical position concerning inter-theoretic relations. Besides, Descartes' method remains even nowadays an essential ingredient of every scientific or simply rigorous approach, given that it seems quite hard (perhaps meaningless) to analyse complex phenomena holistically. In Descartes there is an idea of unified science but not necessarily in a hierarchically sense, from higher (less fundamental) sciences to lower (more fundamental) ones. Descartes and Galileo pointed out a humanistic vision of the unity of science, in which the method is the same, but

[5] If this is strikingly so in modernity and in sciences, it is also true in earlier literary and philosophical studies: the highly symbolic, complex but, at the same time, very organised hierarchy proposed by Dante for heaven and hell; Aristotle's metaphysical vision of sciences organised in three areas (theoretical, practical and productive) which were devoted to different purposes and formed part of a unified hierarchy with the theoretical at the top.

total freedom and independence is left to each discipline. For natural sciences they proposed a unified language as well: the language of mathematics. The precise role of mathematics in natural sciences is a very subtle and deep issue (Bouveresse 2011; Dorato 2010), which remains debatable. However, in a loose sense, this kind of unity of science seems difficult to question, and has proven to hold in general (Klein and Lachièze-Rey 1996).

Before discussing in some detail the contemporary formalisation of intertheoretic reductionism, it is worth discussing a philosophical current that has strongly influenced scientific and philosophical thought in the twentieth century: logical empiricism or positivism. In a very broad sense, this philosophical current aims to completely systematise and unify sciences, even humanistic sciences, appealing to logic as its main instrument. The birth of logical empiricism can be traced back to the 1920s and the vivacious atmosphere of Germanic culture; notably the Vienna Circle and the Berlin Circle (Bechtel and Hamilton 2007). The philosophy of the logical positivists is related to the positivism introduced by Comte, the early nineteenth-century French philosopher who was sceptical about philosophical systems and of metaphysics in general. For this reason, he emphasised the importance of "positive knowledge"—that is, knowledge grounded in observation and experimental verification—even for the disciplines that had remained mostly speculative. In this context he is considered the father of modern social sciences. Logical empiricism can be seen as a much more radical version of classical empiricism, first developed by Hume, which has been also influenced by Mach's positivism, which implied a radical empiricism which considered sensorial experience as the only source of knowledge (Bechtel and Hamilton 2007). The adjective *logical* refers to the main tool considered by these empiricists, to proceed from individual observations to generalised scientific claims.[6] Some also argued that the different sciences could be unified through theory reduction. In that sense, logical empiricists appear to accord with Dirac's and Bohr's formal interpretation of the physical world (Bohr 2011; Dirac 1929) and tried to offer this physicists' approach to encompass all knowledge.[7] Their project culminated in the attempt to provide a common account of the methodology of all sciences and link them into a unique theoretical construct, which gave rise to the International Encyclopedia of Unified Science, edited jointly by Neurath, Carnap, and Morris (Bechtel and Hamilton 2007). The goal, according to Neurath (1938), was to dovetail the scientific disciplines, so that advances in one of them would bring about advances in the others as well. As previously explained, the main tool for such dovetailing of different sciences was logical analysis, necessary to formalise and systematise all the concepts of the different sciences and, eventually, the global theoretical claims of various sciences. The editors of the International Encyclopae-

[6] In this sense, they were strongly influenced by the advances of mathematical logic of the late 19th and early 20th centuries due to Frege, Peano, Russell, Whitehead, and others.

[7] In Chap. 6, we will see how the philosophical ideas of the fathers of quantum mechanics are elaborate and not monolithic. On the other hand, a logical empiricist like Schlick developed a realistic approach.

dia envisioned an axiomatised integration of the whole body of knowledge provided by the various sciences and were convinced that this project would lead to a global improvement of each science as well as to prospects for their integration.

2.2.2 Philosophical Model of Theory-Reduction

2.2.2.1 Hempelian Explanation

Inter-theoretic reductionism entered naturally into the doctrine of logical empiricists, as one of the steps needed to set science free from metaphysical issues, and in relation to the issue of the unity of science (AA.VV 2011). In order to accomplish this difficult task, logical empiricists emphasised the role of logical analysis, trying to study by this means all claims made by the different sciences. The crucial idea was to represent scientific claims, including observations and theoretical statements, within a unique framework, whenever possible. In particular, Nagel identified "experimental laws" as a possible relevant candidate for this purpose, regarded as the middle way between theory and experiments and which were able to provide an empirical summary of the phenomena observed. Galileo's law of the quadratic dependence on time of the distance travelled by a falling object is considered an example of experimental law by Nagel. Moved by aversion to metaphysical and ontological issues, logical empiricists wanted to minimise the definition of theoretical objects needed to account for empirical evidence. In this framework, new predictions in the form of unknown observation statements are deduced directly from laws (with the addition of particular conditions), and thus their generality is measured by the number of new statements effectively included. This is the well-established deductive–nomological (D–N) or covering-law model of explanation (Hempel and Oppenheim 1948; Hempel 1965).

Logical empiricists proposed the generalisation of this approach toward the relation between different theories (Kemeny and Oppenheim 1956; Putnam and Oppenheim 1958), requiring that the same observable predictions be obtained within the less general theory (in their language, the reduced theory) as well as in the more general (the reducing) one. This approach was later formalised to show how to derive (hypothetically, at least) the laws of one discipline or science from those of another (Woodger 1952; Nagel 1979; Quine 1964; Kuipers 2001).

2.2.2.2 Nagelian Model of Reduction

The reductionist schema used in most philosophical models of reduction is based upon the seminal and most influential work of Nagel (1979), who stated that: *A reduction is effected when the experimental laws of the secondary science (and if it has an adequate theory, its theory as well) are shown to be logical consequences of the theoretical assumptions (inclusive of the coordinating definitions) of the primary science.*

It is worth recalling that, in all these philosophical discussions, reduction is concerned with a *secondary* science (less fundamental or higher-level) which can be reduced (in some sense) to a *primary* science, considered more fundamental or lower-level.

Thus, roughly, the Nagelian model of reduction can be summarised as follows: given two theories T and T1, T1 is reduced to T if the laws of T explain the laws of T1, where the explanation is interpreted in terms of Hempelian deductive-nomological model. In other words, T1 is reducible to T if the laws of T1 are derivable from the laws of T.

A first issue arose immediately from the obvious fact that laws in different sciences or even at different levels of the same science make use of different vocabularies. Therefore, one should ask whether the relation between semantically different domains is meaningful and to what extent. To answer this question Nagel first proposed that the homogeneous (same vocabulary) relations be distinguished from the heterogeneous ones. In addition, he advocated the use of some "*rules of correspondence*", now commonly called *bridge principles*, that equate the vocabulary. Let us follow Nagel's words: *If the laws of the secondary science contain terms that do not occur in the theoretical assumptions of the primary discipline . . ., the logical derivation of the former from the latter is prima facie impossible*. The two supplementary conditions to be met in order for the reduction to take place are:

1. connectability: assumptions of some kind must be introduced which postulate suitable relations between whatever is represented by "A" (the missing term in the reduced theory) and traits represented by theoretical terms already present in the primary science; and
2. derivability: with the help of these additional assumptions, all the laws of the secondary science, including those containing the term "A", must be logically derivable from the theoretical premises and their associated coordinating definitions in the primary discipline (Nagel 1979).

As an example of heterogeneous reduction, Nagel takes the reduction of thermodynamics to statistical mechanics. In this case, the concept of temperature is purely macroscopic and simply does not exist in the microscopic world.

The second issue concerning this theory reduction model is the fact that the regularities captured in higher-level laws exist only under certain conditios. To overcome this, he proposed that reduction also required statements of additional possible elements (in mathematical language they are boundary conditions). Therefore, in conclusion, a reduction scheme is then conceived to have the the following form:

- Lower-level laws (in the basic, reducing science)
- Bridge principles
- Boundary conditions
- Higher-level laws (in the secondary, reduced science).

A very standard example is the derivation of Gay-Lussac law from the kinetic theory of gases, as part of an overall reduction of classical thermodynamics to the newer and more basic science of statistical mechanics (Nagel 1979).

Feyerabend, however, argued that meaningful bridge principles cannot be established (Feyerabend 1962, 1985), since words in different theories have different meanings hence they remain *incommensurable* even when they have the same form. Kuhn too insisted on this difficulty, focussing on the so-called reduction of Newtonian to Einsteinian mechanics (Kuhn 1996). In order to address these criticisms, Schaffner (1967, 1969) revised the Nagelian model, describing "reduction functions" rather than bridge laws and proposing a revised Nagelian-type model, encompassing in a formal way the examples given by Feyeraband and Kuhn.[8]

Several comments are in order.

- As pointed out by Sklar (1967), the history of science strikingly demonstrates a point often overlooked by both scientists and philosophers—there are no actual successful homogeneous reductions of theories which do not concern old issues of scarce interest in present-day science. For instance, one might propose the inclusion of Galilean experimental results on falling objects within Newtonian mechanics. This obliges us to consider two other important questions.
- The project of the logical positivists does not seem to be grounded much in scientific facts.[9] Their ideal of knowledge is not necessarily related to reality but unified by formal logics and in practice reduced to physics. Let these enthusiastic reductionists talk: Reichenbach (1959) maintained that "*today it is possible to say that chemistry is a part of physics, just as much as thermodynamics or the theory of electricity*", and Putnam and Oppenheim (1958) argued for "*the possibility that science may one day be reduced to microphysics (in the sense in which chemistry seems today to be reduced to it* ..."). Or consider the claim of Nagel (1979) that "*certain parts of 19th century chemistry (and perhaps the whole of this science) is reducible to post-1925 physics*". Such statements are not made by philosophers alone. Dirac (1929) wrote that "... *the underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble*". And Feynman et al. (1964) celebrated the Schrödinger equation as "... *one of the great triumphs of physics. By providing the key to the underlying machinery of atomic structure it has given an explanation for atomic spectra, for chemistry, and for the nature of matter*". Nonetheless, this strict hierarchy of levels (mind, psychology, biology, chemistry ...) reducing eventually to physics leaves little room for the independence of higher levels, and the ontological significance of those levels' properties and elements[10] appear to be swept away. We will shortly return to this point.

[8] Other Nagelian-type models have been suggested and discussed subsequently (Hull 1974; Ruse and Wilson 1986). However, it is worth emphasising that these models all differ from Nagel's original one, but not from a substantial and philosophical point of view. In particular, they all contain some kind of (not clearly specified) bridge principle.

[9] However, contrarily to a certain *vulgata*, logical positivists **were well aware** of these limits (Bouveresse 2011).

[10] In a different view, their causal powers.

- The reduction of thermodynamics to statistical mechanics is of a different nature and deserves much more attention: it will be discussed in detail later. For the time being, it suffices to note that this supposed success has unfortunately become paradigmatic and even today it is presented in many philosophical discussions and books almost like a dogma of successful reduction. However, this belief is completely misguided, as we shall show [see also (Sklar 1995)].
- Another issue concerning the philosophical position of logical empiricists deserves consideration. It is widely accepted that connecting theories are problematic and that bridge laws are questionable (Fodor 1974). Nevertheless, the prevalent idea in the philosophical literature is that one should rely on identity relations grounded in logical argument. A lucid presentation of this approach is given by Suppes, one of the first proponents of a theory-reduction model: "*To show in a sharp sense that thermodynamics may be reduced to statistical mechanics, we would need to axiomatise both disciplines by defining appropriate set-theoretical predicates, and then show that given any model T of thermodynamics we may find a model of statistical mechanics on the basis of which we may construct a model isomorphic to T*" (Suppes 1957). Consequently, Suppes required an isomorphism between each model of the reduced theory and a corresponding one of the more general one; therefore before proceeding to a reduction between two theories, one should logically formalise and axiomatise both of them, determine the suitable identity relations and eventually carry out a deductive-nomological reduction.
  This is important since the current approach of philosophers of science rests by and large on the logical analysis of theories (Humphreys and Bedau 2006). In this work, we mainly address this kind of reductionnism: deducibility in a broad logical sense. This can be said to be of Nagelian type, even considering its many recent improvements (Kim 1993, 2000; Butterfield 2011a,b; Bouveresse 2011). In the following, we will give explicit examples which show that reduction both in Suppes or Nagel terms is in fact impossible. However, even admitting that certain "technical" difficulties could be overcome sooner or later, the axiomatisation of theories remains necessary. Now, by definition, science is a work in progress, thus any attempt to axiomatise any of its branches, in order to effect a logical reduction of current theories, appears unrealistic. In many cases, dialectical or dynamical relations, rather than formal logical ones, appear more relevant (Sève 1998; Sève and Guespin-Michel 2005; Sanchez-Palencia 2013).

2.3 Reduction in Physics and Philosophy

Physicists look at inter-theoretic relations in the opposite manner. As recognised by Nickles (1973) in a philosophical paper, there are two notions of reduction:

1. Reduction$_1$, namely the reduction of philosophers discussed so far;
2. Reduction$_2$, namely physicists' reduction of theories (Batterman 2002).

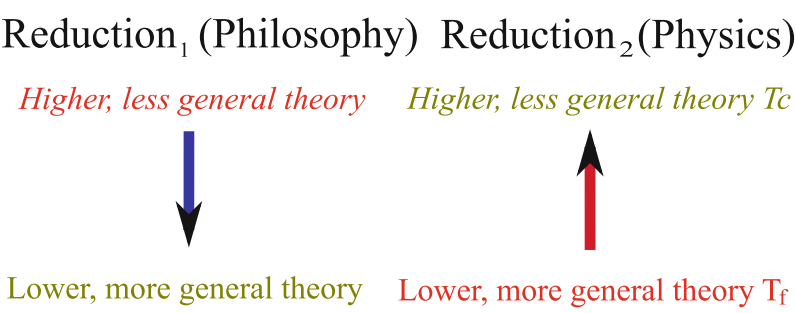

Fig. 2.1 Illustration of the two types of reduction: reduction$_1$, the "philosophical" notion; reduction$_2$, the "physical" notion. The first concerns the reduction of the more specific theory to the more general, in the attempt to unify and explain the phenomena of interest; the latter views the more general theory as encompassing the more specific theory, thus reducing to that in the overlapping areas. Philosophers have noticed that the physicists' notion of reduction implicitly entails a horizontal time direction, meaning that the more general theory is often more recent than the more specific one, since it is developed in the light of new observations incompatible with the previous theory

In physics, the reduction of a theory (or law) T_f (f for finer) to another T_c (c for coarser) means that the theory T_f is more general and encompasses theory T_c and, in some particular case, T_f implies T_c. Furthermore, theory T_f is usually more recent than T_c, so that it can be viewed as a generalisation of the coarser theory T_c. Thus the meaning of reduction is similar to the deductive one (Nagel 1979) but the roles of reduced and reducing theories in physics are reversed with respect to the philosophical notion of reduction, Reduction$_1$. Despite possible semantic or ontological issues, it is always highly desirable for a new proposition to be consistent with the existing more specific theory, if the latter is still valid. Ineed, the reproduction of previously established results is the first, necessary step in the validation of a new theory or law. In Fig. 2.1, the two kinds of reductions are depicted. To give a simple, classical example that illustrates this point, consider the relation between Newtonian and relativistic mechanics. Newtonian mechanics was, and is, known to be in excellent agreement with macroscopic experiments, hence with basically all experiments carried out till the first half of the 19th century. The special theory of relativity, as well as the subsequent general theory of relativity, intended to address open issues in Newtonian mechanics. However, they clearly were generalisations of Newtonian mechanics, supposed to reduce to that in all cases in which Newtonian mechanics makes correct statements. In this sense, physicists refer to Newtonian mechanics as a limiting case of the theory of relativity: relativity (T_f in this case) reduces to Newtonanian mechanics (T_c) when the speed of light is considered infinite.

This shows that, while in Reduction$_1$ the more specific upper-level theory is reduced to the more general lower-level one (e.g. the reduction of gas laws to the more general theory of statistical mechanics), in Reduction$_2$ the more general theory is newer and reduces to the older theory, now recognised to be partially incorrect

(e.g., the reduction of Einstein's formula for momentum to Newton's formula). In brief, Reduction_1 moves from special to general, whereas Reduction_2 moves from general to special (See Fig. 2.1).

Let us begin to formalise this relation of limits which is important indeed to understand inter-theoretic relations. Since natural sciences, particularly physics, are highly mathematical in nature, a characteristic parameter (typically dimensionless and called an order parameter) suggesting the correct limiting cases in which a theory is valid, possibly recovering the domain of validity of a coarser one, may usually be identified. The coarser theory can be seen as an asymptotic limit of the finer one, in a mathematical sense. Symbolically we can express this limiting operation as:

$$\lim_{\delta \to 0} T_f(\delta) = T_c \tag{2.1}$$

It is important to underline here that first, the limit operation applied to theories should only be taken in a formal symbolic sense, not in a rigorous one; and second, we are leaving aside for the moment all epistemological issues (such as "incommensurability") which may arise (and indeed usually do) in such a reducing limit. Let us reconsider the case of Newtonian mechanics and special relativity. As mentioned above, the relevant parameter is c, which must tend to infinity for special relativity to reduce to the Newtonian mechanics. Here, m_0 is the rest mass and the classical expression $p = mv$ is therefore recovered in the asymptotic case $c \to \infty$. The same limit procedure can be explicitly found in many other inter-theoretic relations. Following Berry (1994), we propose a list of asymptotic limits[11] which connects different theories:

- special relativity $\to$ Newtonian mechanics: $\delta = \frac{v}{c}$, where v is the velocity, c is the speed of light.
- general relativity $\to$ special relativity: $\delta = \frac{Gm}{ac^2}$, where G is the universal gravitational constant, m is the mass and a is the relevant length.
- statistical mechanics $\to$ thermodynamics $\delta = \frac{1}{N}$ and $\tilde{\delta} = \frac{1}{V}$ with $\delta/\tilde{\delta} = const.$ where N is the number of microscopic particles in the volume V.
- viscous fluids $\to$ non-viscous fluids: $\delta = \frac{1}{Re}$; where Re is the Reynolds number ($Re = \frac{UL}{\nu}$ is the relevant adimensional number, with ν the kinematic viscosity, U and L are the typical velocity and length of the flow respectively).
- wave optics $\to$ geometrical optics: $\delta = \frac{\lambda}{a}$; where λ is the wavelength and a the relevant macroscopic length.
- quantum mechanics $\to$ classical mechanics: $\delta = \frac{h}{S}$; where h is the Planck's constant and S is the relevant action of the system.

[11] These limits are often called asymptotic since asymptotic analysis (Bender and Orszag 1978) is commonly used to describe the limit behaviour, which in our case is how a theory behaves when the appropriate order parameter approaches the limit value. The theory is trivial when the limit is regular but can be quite sophisticated and complex when it is singular (Primas 1981).

Some of these limits will be analysed in detail in the following chapters; the last two have also been discussed at length elsewhere (Berry 1981, 2001; Batterman 2002).

It is important, however, to note that the limit concerning relativistic mechanics is *regular*: there are no singularities, and therefore no infinities arise. In this case, it is appropriate to talk of inter-theoretic reduction, even though there is room for subtle discussions from the semantic and ontological points of view illustrated e.g. by Feyerabend and Kuhn. More precisely, let us consider one example given by Batterman (2002). In the symbolic Eq. 2.1, if formulas in T_f smoothly approach the corresponding formulas in T_c, the limit can be said to be regular. In these cases, the result obtained by taking $\delta = 0$ equals that given by the $\delta \to 0$ limit. The limit is said to be singular if the behaviour for $\delta = 0$ differs from the behaviour for $\delta \to 0$. In these cases the finer theory cannot very simply and directly reproduce the coarser one. A simple and clear illustration of these cases is afforded by the quadratic equation

$$x^2 - x + \delta 9 = 0 \tag{2.2}$$

which has two roots for any value of δ, and these two roots smoothly converge to the two roots of the case with $\delta = 0$, which are 0 and -1. Differently, the equation

$$\delta x^2 - x + 9 = 0 \tag{2.3}$$

has two roots only if $\delta > 0$, whereas at $\delta = 0$ the equation changes drastically its nature, becoming an equation of first degree with 9 as its unique root. This is an example of a singular limit, and even though not all singular limits are due to changes in the degree or order of the equations, this kind of singularity is paradigmatic. In practice, the relevant equations, which are usually partial differential equations, either change in order or become ill-posed in the asymptotic limit. Berry (2002) explains the essence of singular limits with an amusing example. Biting into an apple and finding half a maggot is unpleasant, but finding one-third of a maggot is worse. The less you find, the more you might have eaten. However a small maggot fraction $\delta \ll 1$ is qualitatatively different from no maggot ($\delta = 0$).

It is also important to stress that regular limits are almost absent in inter-theoretic relationships, the relativistic-Newtonian mechanics case being a fortunate exception. On the other hand, singular limits are much more interesting than regular limits from a scientific and philosophical point of view, because they are typically related to the discovery of new properties.

In the list given above, the inter-theoretic asymptotic limits are singular apart from general relativity, which approaches special relativity regularly, and special relativity, which in turn approaches Newtonian mechanics regularly. The corresponding emerging features, among others, include: (a) critical phenomena, when statistical mechanics approaches thermodynamics; (b) turbulence in fluids; (c) interference and caustics when wave theory approaches geometric optics and (d) fluctuations universality and chaos when quantum systems reduce to classical ones. In all of those cases

in which a regular asymptotic limit cannot be taken, one cannot speak of reduction and new features appear. In those situations, it is possible to talk of *emergent* properties and hence of some kind of *emergence* (Batterman 2002).[12]

2.4 Emergence

2.4.1 Introduction

The issue of *emergence* is the appearance of novel phenomena, even though this definition is far from clear-cut. The word comes from latin and appeared in the french and english vocabularies in the fifteenth century (Adler et al. 2002). It indicates *appearance* (of novel properties). Since the eighteenth century the word has been used as a technical term in physics, geology and in evolutionary biology, where emergence indicates the appearance of a new and functional organ in a vegetal or animal line (Adler et al. 2002). The philosophical use is more recent and indicates effects which are not mechanically explained by their causes. The first to use the word in this sense is Lewes, who began the movement called "British emergentism" (Alexander 1920; Morgan 1923; Broad and Paul 1925). He distinguished emergent facts, which cannot be predicted on the grounds of past experience, from resultant facts, which can. In the twentieth century the word emergence has been widely (perhaps too widely) used for epistemological purposes and not only in inter-theoretic relations. From the very beginning, emergence opposed reductionism. It is possible to bring this opposition to life by emphasising two contrasting views of emergence and reduction given by two biologists: *nonoverlapping magisteria* (keeping different levels separate), by Gould (1997), and *consilience* (trying to connect all the levels), by Wilson (1998). It is also fair to say that a new furore about emergence, reductionism and supervenience has characterised the last decades, notably within the debates on the philosophy of the mind, in which the concept of supervenience has been differentiated from that of emergence (Humphreys 1997b; Humphreys and Bedau 2006).

Before entering into details about reduction and emergence, it is worth saying something about the relationship between fundamental scientific theories and contingent conditions or *contexts*, which are bound to play an important role in the emergence of new features within inter-theoretic relations.[13] A context fixes the

[12] For readers more interested in philosophy, it is worth noting that emergence and Nagelian reduction are distinct notions but, in our opinion, they are related and it seems reasonable to argue that the presence of emergent properties make inter-theoretic reduction in the Nagelian sense impossible to carry out. Even though deeper analysis is certainly required, the arguments discussed in the next chapter appear in favour of this thesis.

[13] This fact, recognised by logical empiricists like Nagel, has been rigorously expressed by Primas (1981, 1998) and has been made more palatable to scientists in recent studies (Bishop and Atmanspacher 2006).

boundaries between what can be considered relevant or irrelevant in a given situation, notably in an experiment or observation. Therefore, the context allows us to fix the relevant level of description of a specified reality, hence the appropriate theory or model. Fundamental theories are generally formulated in terms of universal principles, assuming that these principles can apply to many different phenomena and possibly to everyone, as the idea of universality implies. It is plain that, given this broad scope, these theories are constructed in a context-independent way. In this sense, one may be led to accept the sceptical view, (Van Fraassen 1989; Cartwright 1983) according to which *the fundamental laws of physics do not describe true facts about nature*, and therefore the laws of physics do not state true facts, which is another way of saying that there are no laws of nature at all. However, this is due to the fact that fundamental laws refer to independent reality whereas phenomenological laws refer to empirical reality (Primas 1998; Bishop and Atmanspacher 2006). Starting from fundamental laws, one can build an operational theory which is a phenomenological model and can be used in the empirical context, disregarding all the details that are not relevant. The existence of a context which in turn identifies the relevant model is related to the emergence of new features. In practice, in inter-theoretic approaches, one starts from a given fundamental theory and restricts it, according to the theoretical context of interest, to obtain a coarser, less fundamental theory. From a formal point of view concerning the mathematical sciences, this operation amounts to an asymptotic expansion which is often singular, context-dependent and leads to the emergence of features qualitatively different from those characterising the fundamental theory. This point will be discussed further in the following chapters.

2.4.2 Reduction versus Emergence

A good collection of views on emergence, along with many influential and respected papers can be found in Humphreys and Bedau (2006). It is instructive to go deeper in the definition of emergence, following Kim (2000), who gives the following list of points, meant to illustrate the tenets of the "central doctrine of emergentism":

1. *Emergence of complex higher-level entities*: Systems with a higher level of complexity emerge from the coming together of lower-level entities in new structural configurations.
2. *Emergence of higher-level entities*: all properties of higher-level entities arise from the lower-level properties and relations that characterise their constituents. Some properties of these higher, complex systems are "emergent", and the rest merely "resultant".
3. *The unpredictability of emergent properties*: emergent properties are not predictable from exhaustive information concerning their "basic conditions". In contrast resultant properties are predictable from lower-level information.

4. *The inexplicable/irreducibility of emergent properties*: Emergent properties, unlike those that are merely resultant, are neither explicable nor reducible in terms of their basal conditions.
5. *The causal efficacy of the emergency*: Emergent properties have causal powers of their own; novel causal powers irreducible to the causal powers of their basal constituents.

Some comments are in order.

- Points 1–2 insist on the importance of the *whole*-to-*part* relationship, which is related to a hierarchical vision of the world subdivided into different levels. These points are hence related to the intuitive definition of reduction: the whole is nothing but the sum of the parts. While the hierarchical world is in general acceptable and difficult to question, it is not always relevant for inter-theoretic relations. It has been shown convincingly, in physics at least (Batterman 2002), that emergent properties without a whole-to-part relationship may be found when inter-theoretic relations can be formalised as asymptotic limits, and when asymptotic analysis holds. On the contrary, what seems really important for formalised sciences is the presence of a singular limit in the inter-theoretic relation underlying the emergence of new properties.
- Points 3–4 try to identify the characteristics of a genuinely emergent property. These two issues seem indeed to be relevant. It is difficult to conceive of emergent properties without thinking of something that cannot be predicted or explained in terms of the finer theory when approaching the coarser one. This point will be discussed at length in the following chapters in order to clarify the sense in which this explanation can be given, when it can be given.
- Point 5 is profoundly grounded in the debates on the philosophy of the mind. In this book, however, only inanimate matter is considered, which makes it hard to attribute causal power to finer or coarser levels or theory. In these cases it seems better (and also safer) to talk of explanation rather than of causal relations.

We can thus use here the scheme of classifications proposed by Bishop and Atmanspacher (2006), allowing different kinds of reduction and emergence to be discussed consistently:

1. At a certain level, the description of properties (including its laws) offers both necessary and sufficient conditions to rigourously derive the description of properties at a higher level. This is the strictest possible form of reduction.
2. At a certain level, the description of properties (including its laws) offers necessary but not sufficient conditions to derive the description of properties at a higher level. This version indicates that contingent contextual conditions are required in addition to the lower-level description for the rigourous derivation of higher-level properties. In this case, we speak of emergence.
3. At a certain level, the description of properties (including its laws) offers sufficient but not necessary conditions to derive the description of properties at a higher-level. This version includes the idea that a lower-level description offers multiple

realisations of a particular property at a higher level—a feature characteristic of supervenience.

4. At a certain level, the description of properties (including its laws) offers neither necessary nor sufficient conditions to derive the description of properties at a higher-level. This represents a form of radical emergence, insofar as there are no relevant conditions connecting the two levels whatsoever.

Note that class (2) complements class (3), since in many cases it turns out that higher-level features both supervene on and emerge from lower-level properties. Class (4) is not particularly attractive to those interested in explanatory relations between different levels of description, since it regards as loose the implications between those levels of descriptions. By contrast, class (1) represents the conventional wisdom of reduction: lower-level theories imply and completely set higher-level theories. Even though some examples in the literature were originally thought to exemplify class (1), in reality they do not bear closer scrutiny, and this class seems to be almost empty. Therefore, fewer and fewer philosophers admit class (1), a notable exception being Kim (1993). However, this vision remains probably the received view among some physicists.

2.4.3 Emergence and Reduction in Natural Sciences

This book concerns formal natural sciences, or exact natural sciences, namely physics and chemistry. In the following chapters, we shall discuss several examples of inter-theoretic relations which are non-trivial and for which the tenet (1) of strict reduction will be shown to be clearly inappropriate. Furthermore, it will be shown that some new features emerge *en route* from the basic theory to the higher-level theory.

Natural sciences, theoretical physics in particular, are highly mathematicised. Nowadays, it is frequent to see the cross-fertilisation from the frontiers of theoretical physics in some fields of mathematics. This allows a detailed analysis of inter-theoretic relations and thus may constitute an important testing ground for more general speculative arguments put forward by philosophers. Unfortunately, too often philosophers are anchored to a very simple idea of exact scientific theory and miss the mathematical and physical subtleties underlying them. This is particularly damaging, since these subtleties are nevertheless used to formulate examples for reduction and emergence from physics and chemistry (Nagel 1979; Kemeny and Oppenheim 1956; Feyerabend 1962; Bunge 1985; Humphreys 1997b; Batterman 2002) among the best known. However, the powerful method of investigation permitted by mathematics comes at a cost: **physics and chemistry** treat only the simplest natural object: inanimate matter. The more complex the phenomena considered, the less formalised description the that must be used. Nevertheless, we hope and believe that considerations arising from natural sciences are useful also in more complex domains of study and perhaps also for speculative analysis.

In inter-theoretic relations concerning phenomena at the boundary between two levels of representation of reality, the fundamental equations (the lower-level theory) do not suffice to represent the higher-level context-dependent empirical reality, and have to be complemented by suitable context or boundaries. This operation is generally accomplished via some multi-scale mathematical approach. Through an asymptotic expansion it is possible to produce the limit of the basic lower-level theory which should hold in the higher-level one. Nevertheless, this limit most often turns out to be singular, with the relevant terms diverging to infinity, which makes the desired asymptotic expansions impossible in the basic theory.

This singularity shows that the basic theory is not sufficient to cover higher-level phenomena. This means that there is a gap between the levels that cannot be bridged by the sole "language" taken from the fundamental theory. New semantics should be introduced, in order to obtain a closed and complete representation of the higher level. One then wonders whether natural science teaches anything about category (3), supervenience. It is fair to say that this category has been put forward in the framework of the philosophy of the mind, and is associated in particular with Kim (1984). We shall discuss this point in the context of "special sciences" in the next chapter.

2.4.4 Emergence and Reduction in Special Sciences

Let us now briefly review some arguments about reductionism in sciences other than physics, particularly in the philosophy of the mind, which has witnessed heated debates over the last decades. The implications of our study for these issues will be discussed in the conclusion.

As pointed out in the last section, the subjects of emergence and reduction are strongly related, and are also related to a very general issue: that concerning the unity or plurality of science. Indeed, specific scientific theories try to explain a vast variety of phenomena in terms of a very large number of different disciplines, sub-domains, etc. One may then ask whether this structural differentiation reflects a real and deeper differentiation at the level of objects and properties or whether, on the contrary, it is simply due to an effort to classify and organise the topic of interest, which would imply the existence of a fundamental unifying science, capable of encompassing all the sciences. It is clear that this view is related to some form of reductionism.

Let us state precisely a subtle but fundamental point, the problem of the status of a scientific theory. Science aims to describe the world and thus gives an interpretation that refers to our understanding of observable behaviours or patterns. This form of interpretation of the empirical reality is epistemic. On the other hand, a "realistic" interpretation deals with the nature of existence and refers to a theory about "real things", i.e. about those objects which exist independently of any observational context. A particular metaphysical doctrine which seems very reasonable and widely accepted (probably by all physicists) is physicalism. Physicalism states that all the entities in the world are physical and that all properties are either physical or related

to physical properties. For instance, "*All individuals are constituted by, or identical to, microphysical individuals, and all properties are realised by, or identical to, microphysical properties*" (Gillett 2003).

Here microphysical means the lowest-level of the physics description, hence could be that of elementary particles, but this notion, commonly used in philosophical discussions, is vague and often higher-level properties like chemical ones are included in the microphysical ones. Nevertheless, the physicalist thesis seems to be confirmed in actual science and thus to refuse it appears smacks of sophistry. Indeed, in natural sciences at least, there is no example of a non-physical entity. All experimental work in the different areas of physics (from particle physics to astrophysics) and chemistry agree on these grounds.

Given the physicalist framework, it is opportune to give an account of direct reductionist views, which can be referred to as reductive physicalism or ontological reductionism, and which naturally lead to an ontological minimalism. Reductive physicalism means: (a) there are a small number of different fundamental constituents of the world; (b) every other object, state, process or property is composed of these fundamental entities, and are nothing but fundamental physical entities. It is fair enough to say that many physicists share this view still today [(Weinberg (1987) being one of its champions].

Although ontological minimalism has the merit of order and simplicity, it dissatisfies those who think that biological processes and particularly human actions are not so directly related to the properties of the fundamental building blocks of matter. It seems hopeless to counter such strongly rooted beliefs about the ontological status of the world. In some sense, ontological minimalism is a form of metaphysical position which substitutes some kind of god with elementary particle physics. At variance with this position, some authors like Davidson claim that there are no possible bridge laws between the level of mind and that of physics. He asserts that such bridge laws are not merely hard to discover (and hence have not yet been discovered): he states that they do not exist. Davidson indeed believes that mental states and process are "anomalous" in the sense that they cannot be expressed in terms of laws (*nomos* in greek), by their very nature (Davidson and Block 1980). He bases his views on two main facts which differentiate human regularities from natural ones: human regularities are normative and present numerous exceptions, while natural laws are factual and generally true.

This purely metaphysical position is hardly satisfactory and philosophers have pursued the non-reductionist thesis (Darden and Maull 1977; Machamer et al. 2000; Craver 2007; Bechtel 2008). Fodor, in an article of note (Fodor 1974), argue on this neat distinction between physics, which should be the fundamental science by definition, and other sciences named "special sciences".

In this framework, he also maintains that bridge laws between all special science regularities and physical laws are impossible to find. In particular, he claims that physics may describe some of the natural regularities, but not all of them and that, consequently, special sciences enjoy an autonomous status, independent of the laws of physics. It is worth noting that Fodor uses purely qualitative arguments. Complex phenomena, like the economy, include so many intricate physical actions that it

becomes impossible to find any bridge laws. Actually, the main point of that paper is to introduce the idea of multi-realisability, which would later become very popular in the philosophy of the mind, stating that, given a property described by some special sciences, it can be realised by physical properties in several different manners. Assuming that bridge laws between physics and special sciences do not exist, Fodor postulates that the properties of special sciences should be multi-realisable. Fodor's main conclusion is that the world is compatible with a non-reductive physicalism, in contrast with the reductive physicalism or ontological minimalism: properties expressed in terms of special sciences are realised by a combination of physical properties, but they can be realised in various different ways.

Non-reductive physicalism is different from Descartes' dichotomy: on the one hand, it states that there is an ontological unity, since every property is ultimately physical but, on the other hand, it denies unity, because aggregation of physical properties described by special sciences escape a purely physical description. There is therefore a systematic dependence of special properties on physical ones, without that implying the identification of the relevant descriptions of the different phenomena. This kind of dependence has been specified as supervenience (Kim 1993, 2000). A plethora of definitions have been proposed to distinguish various kinds of supervenience: weak, strong, or global supervenience. From a logical point of view, accepting supervenience has direct consequences. Indeed, adopting the physicalist thesis "*all individuals are constituted by, or identical to, microphysical individuals, and all properties are realised by, or identical to, microphysical properties*", means, in terms of causal theory, adopting the causal inheritance principle (Kim 2000) or the "realisation thesis" (Gillett and Rives 2001):

> An instance of a property, or combination of properties, P realises an instance of a property M if and only if P plays the causal role of M by virtue of P having all the causal powers individuative of M, but not vice versa.

These arguments imply that higher-level properties do not contribute any causal powers to individuals. At variance with common sense, all causal powers[14] lie in the microphysical properties. In this sense supervenience leads naturally to a new form of reductionist physicalism. Higher properties may be multi-realised, but are not autonomous, which makes them mere epiphenomena. Some authors have thought of possible ways out, in the same framework. Notably, Gillett (2003) suggests that a possible way to circumvent the loss of causal power at higher level is to show that physics is not causally closed or "complete".[15] However, this would seem a challenging task, for all experiments in physics suggest the opposite.

In light of this argument, supervenience has lost most of its appeal to philosophers engaged in non-reductive approaches. Emergence, a category which became popular at the beginning of the nineteenth century during so-called British Emergentism, has

[14] In the notion given by Kim (2000).

[15] Basically the claim that all microphysical events are determined, in so far as they are determined, by prior microphysical events and the laws of physics.

risen again in popularity (Humphreys 1997a,b; Gillett 2003; Humphreys and Bedau 2006).[16]

Bedau proposed a weak form of emergence based on two admittedly vague but nevertheless useful hall marks of emergent phenomena, defined as follows:

1. Emergent phenomena are somehow comprised of, and generated from, underlying processes.
2. Emergent phenomena are somehow autonomous from underlying processes.

We therefore have a weak form of epistemic emergence. Higher-level properties are difficult or impossible to explain in terms of the theories describing the lower level, and they are completely determined by physical mechanisms. This keeps the ordered and discrete hierarchical structure implicit in reductive-physicalism, and led Gillet to convincingly demonstrate that weak emergence cannot confer causal efficacy on higher-level properties like those concerning the mind.

Humphreys (1997a) has proposed a strong form of emergence in order to reach a true non-reductive physicalism (allowed neither by weak emergence nor by supervenience). Quite uncommonly for a philosopher, he looked for a genuine empirical fact supporting his views. In his vision, higher-level properties are given by a fusion of lower-level properties, which thus realise higher-level properties without being involved in the causal chain. In this sense, when lower-level properties blend to permit a higher-level property to emerge, they no longer exist and, therefore, do not play any causal role.

Concluding remarks are in order. In the following, we shall analyse some examples of foundational nature, concerning theory reduction in physics. In our opinion, these examples also reveal the weakness of epistemic reductionism, intended in a nomological-deductive sense, although attempts to conciliate reductionism with some forms of emergence have been recently made (Butterfield 2011a,b). Furthermore, we indicate as "naive reductionism" a position similar to homogeneous or simple heterogeneous Nagelian reductionism, in which a given level of description can be deduced simply by the underlying lower level. The "extreme reductionist" pushes to the limits this position, claiming that everything can be deduced from the most fundamental level of reality, that pertaining to sub-nuclear particles.

[16] Earlier we discussed emergence in the framework of natural sciences. Philosophers have discussed it much more in terms of logical relations and in the framework of causal efficacy. They differ on views and definitions, but most agree on the point that emergence has to be related to downward causation (that is to say that physics is not causally complete). In order for mental properties to be causally efficient, they have to cause changes in the physical world.

References

Adler, D., Fargot-Largault, A., Saint-Sernin, B.: Philosophie des sciences. Folio Essais (2002)
Alexander, S.: Space, Time and Deity Volume 2. Macmillan, New York (1920)
AA.VV: The Stanford encyclopaedia of philosophy (2011)
Batterman, R.W.: The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence. Oxford University Press, Oxford (2002)
Bechtel, W.: Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience. CRC Press, Boca Raton (2008)
Bechtel, W., Hamilton, A.: Reduction, integration, and the unity of science: natural, behavioral, and social sciences and the humanities. Philos. Sci.: Focal Issues **1**, 377 (2007)
Bender, C.M., Orszag, S.A.: Advanced Mathematical Methods for Scientists and Engineers: Asymptotic Methods and Perturbation Theory. Springer, Berlin (1978)
Berry, M.V.: Singularities in waves and rays. In: Klacman, M., Balian, R., Poirier, J.-P. (eds.) Les Houches Lecture Series Session XXXV, p. 453. North-Holland: Amsterdam (1981)
Berry, M.V.: Asymptotics, singularities and the reduction of theories. In: Logic, Methodology and Philosophy of Science IX, p. 597 (1994)
Berry, M.V.: Chaos and the semiclassical limit of quantum mechanics (is the moon there when somebody looks? In: Russell, R.J., Clayton, P., Wegter-McNelly, K., Polkinghorne, J. (eds.) Quantum Mechanics: Scientific Perspectives on Divine Action, p. 41. Vatican Observatory CTNS Publications (2001)
Berry, M.V.: Singular Limits. Phys. Today. **55**, 10 (2002)
Bishop, R.C., Atmanspacher, H.: Contextual emergence in the description of properties. Found. Phys. **36**(12), 1753 (2006)
Bohr, N.: Atomic Theory and the Description of Nature: Four Essays with an Introductory Survey. Cambridge University Press, Cambridge (2011)
Bouveresse, J.: Essai VI. Les lumières des positivistes. Éditions Agone (2011)
Boyd, R., Gasper, P.: The Philosophy of Science. MIT Press, Cambridge (1991)
Broad, C.D., Paul, T.K.: The Mind and its Place in Nature. Routledge & Kegan Paul, London (1925)
Bunge, M.: Treatise on Basic Philosophy: Epistemology and Methodology III: Philosophy of Science and Technology, Pt I-Formal and Physical Sciences, volume 7. Reidel (1985)
Butterfield, J.: Emergence, reduction and supervenience: a varied landscape. Found. Phys. **41**(6), 920–959 (2011a)
Butterfield, J.: Less is different: emergence and reduction reconciled. Found. Phys. **41**(6), 1065–1135 (2011b)
Cartier, P.: Kepler et la musique du monde. La Recherche **278**, 750 (1995)
Cartwright, N.: How the Laws of Physics Lie. Clarendon Press, Oxford (1983)
Craver, C.F.: Explaining the Brain. Oxford University Press, Oxford (2007)
Darden, L., Maull, N.: Interfield theories. Philos. Sci. **44**, 43 (1977)
Davidson, D., Block, N.: Mental events. Readings Philos. Psychol. **1**, 107 (1980)
Descartes, R.: Discours de la méthode. Vrin, Paris (1987)
Diels, V., Kranz, W.: Die Fragmente der Vorsokratiker. 6th edn. Weidmann, Berlin (1951–1952)
Dirac, P.A.M.: Quantum mechanics of many-electron systems. Proc. R. Soc. Lond. Ser. A **123**(792), 714 (1929)
Dorato, M.: Why is the language of nature mathematical? In: Antonini, G., Altamore, A. (eds.) Galileo and the Renaissance Scientific Discourse, p. 65. Nuova Cultura, Rome (2010)
Feyerabend, P.K.: Explanation, Reduction and Empiricism. Minnesota Studies in the Philosophy of Science. University of Minnesota Press, Minneapolis (1962)
Feyerabend, P.K.: Problems of Empiricism. Cambridge University Press, Cambridge (1985)
Feynman, R.P., Leighton, R.B., Sands, M.: The Feynman Lectures on Physics. Addison-Wesley, Reading (1964)
Fodor, J.A.: Special sciences (or: the disunity of science as a working hypothesis). Synthese **28**(2), 97 (1974)

Gillett, C., Rives, B.: Does the argument from realization generalize? responses to Kim. South. J. Philos. **39**(1), 79 (2001)
Gillett, C.: The varieties of emergence: their purposes, obligations and importance. Grazer Philos. Stud. **65**(1), 95 (2003)
Gould, S.J.: Nonoverlapping magisteria. Nat. Hist. **106**, 16 (1997)
Heisenberg, W.: Physics and Philosophy. Harper, New York (1958)
Hempel, C.G., Oppenheim, P.: Studies in the logic of explanation. Philos. Sci. **15**(2), 135 (1948)
Hempel, P.: Aspects of Explanation and Other Essays in the Philosophy of Science. The Free Press, New York (1965)
Hull, D.L.: Philosophy of Biological Science. Prentice-Hall, Englewood Cliffs (1974)
Humphreys, P.: Emergence, not supervenience. Philos. Sci. **64**, 337 (1997b)
Humphreys, P.: How properties emerge. Philos. Sci. **64**(1), 1 (1997a)
Humphreys, P.E., Bedau, M.A.: Emergence: Contemporary Readings in Philosophy and Science. MIT Press, Cambridge (2006)
Jona-Lasinio, G.: La matematica come linguaggio delle scienze della natura. In: Matematica Cultura e Societá, p. 93. Quaderni della Scuola Normale Superiore di Pisa, (2005)
Kemeny, J.G., Oppenheim, P.: On reduction. Philos. Stud. **7**(1–2), 6 (1956)
Kim, J.: Concepts of supervenience. Philos. Phenomenol. Res. **45**(2), 153 (1984)
Kim, J.: Supervenience and Mind: Selected Philosophical Essays. Cambridge University Press, Cambridge (1993)
Kim, J.: Mind in a Physical World: An Essay on the Mind-body Problem and Mental Causation. The MIT Press, Cambridge (2000)
Klein, E., Lachièze-Rey, M.: La quête de l'unité: l'aventure de la physique. A. Michel (1996)
Kuhn, T.S.: The Structure of Scientific Revolutions. University of Chicago Press, Chicago (1996)
Kuipers, T.A.F.: Structures in Science: Heuristic Patterns Based on Cognitive Structures An Advanced Textbook in Neo-classical Philosophy of Science. Springer, Berlin (2001)
Machamer, P., Darden, L., Craver, C.F.: Thinking about mechanisms. Philos. Sci. 1 (2000)
Morgan, C.L.: Emergent Evolution. Williams and Norgate, London (1923)
Nagel, E.: The Structure of Science. Hackett Publishing Company, Indianapolis (1979)
Neurath, O.: Unified Science as Encyclopedic Integration. University of Chicago Press, Chicago (1938)
Nickles, T.: Two concepts of intertheoretic reduction. J. Philos. **70**(7), 181 (1973)
Poincaré, H.: Dernières Pensées. Flammarion, Paris (1913)
Primas, H.: Chemistry. Quantum Mechanics and Reductionism. Perspectives in Theoretical Chemistry. Springer, Berlin (1981)
Primas, H.: Emergence in exact natural science. Acta Polytech. Scand. **91**, 83 (1998)
Putnam, H., Oppenheim, P.: Unity of science as a working hypothesis. Minn. Stud. Philos. Sci. **2**, 3 (1958)
Quine, W.V.: Ontological reduction and the world of numbers. J. Philos. **61**(7), 209 (1964)
Reichenbach, H.: Modern philosophy of science: Selected essays. Greenwood Press, Westport (1959)
Ruse, M., Wilson, E.O.: Moral philosophy as applied science. Philosophy **61**(236), 173 (1986)
Sanchez-Palencia, E.: Promenade dialectique dans les sciences. Éditions Hermann, Paris (2013)
Schaffner, K.F.: Approaches to reduction. Philos. Sci. 137 (1967)
Schaffner, K.F.: The Watson-Crick model and reductionism. Br. J. Philos. Sci. **20**(4), 325 (1969)
Sève, L.: Sciences et dialectiques de la nature. La Dispute, Paris (1998)
Sève, L., Guespin-Michel, J.G.: Émergence, complexité et dialectique: sur les systèmes dynamiques non linéaires. Odile Jacob, Paris (2005)
Sklar, L.: Types of inter-theoretic reduction. Br. J. Philos. Sci. **18**(2), 109 (1967)
Sklar, L.: Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics. Cambridge University Press, Cambridge (1995)
Suppes, P.: Introduction to Mathematical Logic. Van Nostrana, Princeton (1957)
Van Fraassen, B.C.: Laws and Symmetry. Oxford University Press, Oxford (1989)

Weinberg, S.: Newtonianism, reductionism and the art of congressional testimony. Nature **330**(6147), 433 (1987)
Wilson, E.O.: Consilience: The Unity of Knowledge. Knopf, New York (1998)
Woodger, J.H.: Biology and Language. University Press, Nancy (1952)

Chapter 3
A First Attempt to Tame Complexity: Statistical Mechanics

The scientist does not study nature because it is useful; he studies it because he delights in it, and he delights in it because it is beautiful.

H. Poincaré

People often observe that today's world is *complex*. Although this term is not precisely defined, we instinctively distinguish simple from complex situations and often need to tackle complexity. Politicians deal with it when facing emergencies caused by natural disasters or anticipating the consequences of certain financial operations. Local administrators wonder whether the introduction of a roundabout at a given intersection will improve or worsen traffic conditions, or have to decide how many doses of flu vaccine they should buy for the coming winter. Therefore, a quantitative theory of *complex systems* would be greatly appreciated by them as well as by many others.

Trying to understand the origin of complex phenomena, one finds that they often occur in the presence of many interacting *agents* which may in turn be simply or intricately structured. It is thus instructive to glance at the numbers concerning various phenomena:

- crowds number from $O(10^2)$ up to $O(10^6)$ individuals
- the DNA has $O(10^9)$ nucleotides
- the world economy involves $O(10^{10})$ persons
- the human brain has $O(10^{11})$–$O(10^{12})$ neurons
- a laser requires $O(10^{18})$ atoms
- the water in a glass comprises $O(10^{25})$ molecules.

It appears that complexity does not merely emerge as a consequence of large numbers. The interactions among the constituents of a given system, and between system and outer environment ought to play an important role. For instance, we share more than 99 % of our genes with rats, yet we look quite different from them, and genetists think that such a difference is mainly due to the different ways in which small groups of alleles interact.

S. Chibbaro et al., *Reductionism, Emergence and Levels of Reality*,
DOI: 10.1007/978-3-319-06361-4_3,

Because biological phenomena are complex, the world economy is complex, the global warming is complex, human behaviour is complex etc., complexity is ever finding a place in various intellectual discussions. Indeed, our world is interrelated on all levels to such an extent that it is practically impossible to disentangle any event occurring in any given place from the events that occurred or are occurring anywhere else. Therefore, local analyses can hardly lead to a satisfactory understanding of these phenomena. At the same time, it is exceedingly difficult to contemplate all the factors contributing to them all at once.

Among the familiar complex systems, the human brain is certainly the most striking one. Its structure is similar to that of a network of telegraphers—the neurons—each of which fires an electrical impulse, when prompted to do so by the arrival of another impulse. After each firing, a neuron stays quiescent for a time, thus producing sequences reminiscent of dots and dashes: definitely a simple mechanism for an organ meant to determine our finest motions, to allow us to speak and to conceive abstract thoughts. But the mechanism would look even simpler from the point of view of the atoms of which the brain is made.

This indicates that different amounts of the elementary constituents of a given object, all obeying the same interaction laws, often need to be described in different, irreducible, terms. On the other hand, the transition from one kind of description (e.g. that concerning atomic interactions) to another one (e.g. that required by the behaviour of neurons) should be gradual because, ultimately, neurons are made of atoms.

Curiously, the list above suggests that very large aggregates tend to behave more simply than smaller aggregates. Speaking of the disordered motion of the enormous number of molecules constituting a gas, Poincaré (1914) expresses this idea as follows:

> It seems at first that the orderless collisions of this innumerable dust can only engender an inextricable chaos before which the analyst must retire. But the law of large numbers, that supreme law of chance, comes to our assistance. In the face of a semi-disorder we should be forced to despair, but in extreme disorder this statistical law re-establishes a kind of average or mean order in which the mind can find itself again.

In a sense, this says that complexity can emerge as a property of large systems, which are too big to be understood in terms of the individual behaviour of elementary constituents, but not so large that the effect of the individual contributions is completely washed out in statistical terms. However, no unambiguous definition of complex phenomena has emerged, so far. The term is mostly used to point out the intractability and unpredictability of many phenomena of current interest. Perhaps, the most widespread notion of complex system refers to entities made of many elementary constituents that interact in an elementary fashion, whose collective behaviour is much richer than the simplicity of the elementary behaviours leads one to expect. Admittedly, this is not a very precise definition.

3.1 A Short History of Statistical Mechanics

The previous discussion demonstrates that tackling the complexity of collective phenomena from a quantitative perspective is, in general, a major endeavour. Statistical mechanics (SM), whose original aim was to understand the thermodynamics of macroscopic systems in terms of their atomic constituents, was the first successful attempt. We will see that this success rests on the assumption that the numerosity of the elementary consituents of the objects of interest is that required to avoid Poincaré's "desperate" situations.

The earliest serious attempt in this direction dates back to the first half of the eighteenth century, when Daniel Bernoulli produced the first version of *kinetic theory*: a theory meant to relate the observable properties of macroscopic objects to the behaviour of invisible, ever rushing and elastically colliding molecules. The novelty of Bernoulli's ideas, compared to those of the ancient greek philosophers and those of his more immediate predecessors and contemporaries, such as Newton and Boscovich, is that they were formulated in mathematical terms and could account for various experimentally measurable properties of the pressure of gases (Cercignani 2006). The mid-nineteenth century, when the works of Maxwell and Boltzmann were published, marks the birth of present-day SM.

These works pursued a microscopic interpretation of the observable macroscopic phenomena, based on the atomistic view of matter, i.e. on the assumption that all material objects comprise an exceedingly large number of microscopic elements, obeying the laws of (classical) mechanics. This way, a relation was found between thermodynamic concepts such as temperature, pressure, heat, work and entropy and mechanical quantities, such as potential and kinetic energy, force, and momentum. In other words, the *primitive ontology* of SM, which is the set of fundamental items assumed to exist and from which everything else descends, reduces to:

- a very large number of atoms obeying the laws of mechanics;
- the empty space, within which atoms move.

For this reason, SM is often considered a reductionistic theory, constructed to derive the macroscopic equations governing the observable behaviour of a given object from knowldege of the interactions among the atoms of that object. This view is summarised in Laplace's celebrated "*A Philosophical Essay on Probabilities*" (Laplace 1829). There, Laplace adopts the *principle of sufficient reason*, maintaining that present events are tied to preceding ones, because a thing cannot occur without a cause which produces it. He then argues that:

> We ought then to regard the present state of the universe as the effect of its anterior state and the cause of the one which is to follow. Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it – an intelligence sufficiently vast to submit this data to analysis – it would embrace in the same formula the movements of the greatest bodies of the universe and those of the lightest atom; for it, nothing would be uncertain and the future, as the past, would be present in its eyes.

As an example, Laplace recalls the case of Halley's comet, which had been correctly predicted by Clairaut to return at the beginning of April 1759, and then ventures to postulate that the regularity of astronomic events must be common to all phenomena.[1]

However, while a sufficiently vast intelligence could predict *all* macroscopic phenomena, like an astronomer predicting lunar eclipses from the law of gravity, humans must resort to probabilistic descriptions. This seems to be but a practical issue in Laplace's thought, as he blames it on our ignorance; an attitude that would later be criticised by Poincaré (1914) and which, indeed, must be confronted with numerous questions.

Consider, for instance, Deutsch's evocative argument about a copper atom in the tip of the nose of the statue of Winston Churchill, in Parliament Square in London (Deutsch 1997). That atom got there because Churchill served as prime minister, because his leadership contributed to the Allied victory in the Second World War and because bronze, which contains copper, is a traditional material for such statues. This constitutes an *explanation* of a low level physical observation—the location of a copper atom—through high-level theories about emergent phenomena such as leadership, war and tradition. Deutsch adds that there is no reason, even in principle, to search for any lower-level explanation. Would our understanding of the phenomenon be advanced by knowledge of the probability that such a statue exists, given the initial condition of the solar system? Would knowledge of the trajectories of all atoms of the Universe, from the beginning of life on Earth, explain the Second World War? The conclusion is that, while for militant reductionists the laws obeyed by subatomic particles are the basis of the hierarchy of all knowledge, in the real structure of scientific knowledge, and of our knowledge generally, such laws play much more modest roles.

This example illustrates very clearly what a microscopic mechanical theory of natural phenomena may be expected to predict, but also shows how useless that would be. After all, given the exponential rate at which computer performance continues to increase, simulating all the trajectories of the molecules of air in a cubic centimetre may not be too remote a possibility. But would that advance our understanding of the properties of a gas, beyond what is already implied by Boyle's law?

In the first half of the nineteenth century, mechanics was such a mature theory, so successful in describing quite diverse phenomena, that it had gradually come to be considered capable of illminating all natural phenomena. By regarding astronomical objects as point masses that attract one another with a force directed along the line joining them, with a strength inversely proportional to the square of their separation, the motion of planets and comets could be predicted with impressive accuracy. Assuming that the same force acted between any two given point masses, and applying the laws of motion inferred from observations made on Earth, Newton described

[1] He states that: "*The curve described by a single molecule in air or vapour is regulated in a manner just as certain as the planetary orbits; the only difference between them is that which comes from our ignorance*", and, more generally, that: "*The regularity which astronomy shows us in the movements of the comets doubtless exists also in all phenomena.*"

the motion of astronomical bodies, computed the mass of those passing by the Earth, and developed a theory of tides. These achievements spanned just about all the scales accessible to observation at the time, motivating Newton's successors to seek further mechanical explanations of natural phenomena. Thus Democritus' belief that matter consists of atoms was brought back to life, thanks to the fact that atoms were also assumed to obey the laws of mechanics. The challenge was tremendous. Apart from the inverse-square gravitational attraction experienced by all material objects, the theory had to explain the impenetrability of atoms, the fact that solids do not evaporate even at low pressures and that their volumes increase when they melt (Cercignani 2006). This led Boscovich (1966) to postulate repulsive interactions among atoms, and to realise that it was not necessary to think that the different states of matter were made of different kinds of atoms.

The atomic theory of matter rapidly achieved important successes. Laplace adopted it to describe capillary action, under the assumption that atoms attract each other at short distances and repel each other at long distances. This theory was successful enough to persist until the work of van der Waals. Thompson, in turn, proposed that heat was just a form of energy related to the irregular motion of atoms, extending the principle of conservation of mechanical energy to all phenomena.

However, as recalled by Brush (1986) in the nineteenth century, the eventual acceptance of the atomic hypothesis was mostly due to studies on the nature of light, not thermodynamic experimental evidence. For instance, in his pioneering work, *Reflexions sur la puissance motrice du feu et sur les machines propre a developper cette puissance* of 1824, Carnot treated heat as a substance which might flow from place to place, although in a cautious and critical fashion. Only later, did he accept the mechanical theory of heat, as documented in (Carnot 1897).[2]

The atomistic view was appreciated by many other scientists, such as Clausius, but Maxwell's contributions were the most insightful and paved the way for Boltzmann's work. Specifically, Maxwell developed a preliminary theory of transport processes including heat transfer, viscous drag, and diffusion but, most importantly, he introduced the concept of a probability distribution function, to describe the statistics of the velocities of gas molecules in equilibrium at a given temprature. He then realised that this distribution coincided with the one that Gauss had previously introduced to describe random errors in experimental observations.

Boltzmann, "the man who trusted atoms" according to the characterisation of Cercignani (2006), had always interpreted continuum models of matter as idealised representations of macroscopic aggregates of atoms, but found strenuous opposition from his contemporaries.

When the issue was eventually settled in favour of the atomistic theory, it had become so convincing that Feynman et al. (1964) could later state:

---

[2] *We may be allowed to express here a hypothesis concerning the nature of heat. At present, light is generally regarded as the result of a vibratory movement of the ethereal fluid. Light produces heat, or at least accompanies the radiant heat and moves with the same velocity as heat. Radiant heat is therefore a vibratory movement. It would be ridiculous to suppose that it is an emission of matter while the light which accompanies it could only be a movement.*

> If, in some cataclysm, all of scientific knowledge had to be destroyed, and only one sentence passed on to the next generation of creatures, what statement would contain the most information in the fewest words? I believe it is the atomic hypothesis.

Indeed, the atomistic approach has proven successful in the description of countless phenomena and the techniques of SM are currently applied to practically any circumstance in which the "*microscopic*" or "*elementary*" constituents of a given "*macroscopic*" or "*collective*" entity may be identified.[3] It is however interesting to note that the endorsement of so many influential scientists from the nineteenth century onwards was not enough to dispel the reservations against the atomic hypothesis held by equally influential scientists, such as Mach and Ostwald. Many scientists and philosophers insisted that matter was a continuum and that atoms could at best be taken as mathematically useful in dealing with certain phenomena. Uncontroversial new knowledge was required to clarify the role of atomism in physics, something that was eventually acquired thanks to the study of the remarkable phenomenon known as Brownian Motion.

Later in the twentieth century, the success of atomism and the techniques of SM led scientists to take them as universal and applicable to the most diverse range of fields. Indeed, the overall behaviourof large assemblies of interacting objects requiring statistical descriptions should depend only moderately on the details of the "microscopic" phenomena. Eventually, SM gave birth to that sociologically interesting phenomenon known as the "*Science of Complexity*". This is a collection of studies, which are frequently referred to? by experts of the most diverse fields, because it seeks for a unitary conceptual framework for all disciplines. Thus, in contemporary research, subjects ranging from physics to computer science, from biology to finance, from political science to modern art criticism, are often considered in terms of the fundamental concepts of SM, such as "entropy" and "criticality".

3.2 Towards a Systematic Theory

Once the atomistic view is accepted, the identification of the issues that may benefit from it still poses a challenge. No doubt, matter is made of atoms, and what happens to any piece of matter depends on what happens to its atoms; but which aspects of the behaviour of that object may be understood in terms of what we know about atoms? Which phenomena may be elucidated by the laws which we use to describe the behaviour of atoms? The fact is that our description of the atomic world is a direct extension of classical (Newtonian) mechanics,[4] which has been very success-

[3] Here, the quotation marks hint at the fact that the term "microscopic" may be quite different from our everyday notion of microscopically small. For instance, it could refer to entire galaxies, in cosmological studies, while the term "collective" may be used in nuclear physics to describe the agglomerate of elementary particles in an atomic nucleus, clearly too small compared to any object that we consider "macroscopic".

[4] Conceptually, quantum mechanics does not change the picture (Lebowitz 1993).

fully developed to describe systems made of a small number of objects, while the properties of matter result from the cooperation of exceedingly large numbers of atoms. Therefore, there is no *a priori* reason why this extension should work, and Laplace's conviction that it should looks particularly bold. After all, while theories typically apply over limited ranges of validity, here Newtonian mechanics must hold over a range of particle numbers which spans 24 orders of magnitude or more! In fact, this is precisely what Newton assumed—that falling apples and the Moon obey the same dynamical laws. Even more boldly, the founding fathers of SM ventured to relate the Newtonian dynamics assumed to describe the motions of atoms with the thermodynamics of macroscopic objects.[5]

To obtain such a relationship, one must consider the space $\mathscr{M}$ of all microscopic states of a given system of interest, known as the *phase space*. In the case of the simple monatomic gases, $\mathscr{M}$ may be taken to be the $6N$ dimensional phase space $\mathbf{R}^{6N}$, i.e. the space of all coordinates and momenta $\Gamma = (\mathbf{q}, \mathbf{p})$ of the N atoms.[6] Given a vector $\Gamma \in \mathscr{M}$ representing an initial microstate, the state at a later time may be denoted by the evolved vector $S^t\Gamma \in \mathscr{M}$, corresponding to the changes in positions and velocities of the atoms, where S^t is the evolution operator for a time t.

In the case of an isolated system of particles, this evolution is expressed by Hamilton's equations of motion, which can be compactly written as:

$$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}} \; ; \qquad \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}} \tag{3.1}$$

where H is the Hamiltonian of the system, and $H(\mathbf{q}, \mathbf{p}) = E$ determines the energy surface in which the time evolution $S^t\Gamma$ is confined for all t. Then, one assumes that each microstate $\Gamma \in \mathscr{M}$ implies a particular value $\mathscr{O}(\Gamma) \in \boldsymbol{R}$ for each observable quantity $\mathscr{O}$.

Measurements do not occur instantaneously, but take a certain amount of time: as a consequence of the very rapid and perpetual motion of the atoms, it does not yield a precise value $\mathscr{O}(\Gamma)$, because Γ changes in that time. Instead, a measurement yields the average of the values that $\mathscr{O}(S^t\Gamma)$ takes while $S^t\Gamma$ explores the phase space. For instance, the pressure of a gas measured by a manometer is determined by the average of the variations of the molecular momenta, which occur when the molecules hit the surface of the sensor.

---

[5] The impressive range of validity of classical mechanics should not be confused with the validity of Hamiltonian mechanics, the mathematically elegant formulation of Newtonian mechanics. This, indeed, suffers from various limitations, such as the absence of dissipation, which may instead be incorporated via Lagrangian mechanics.

[6] If the elementary constituents of the system cannot be approximated by point-like particles, but still obey classical mechanics, one would have to refer to a larger phase space, which has more dimensions, but the basic idea does not change.

3.2.1 Boltzmann's Grand Vision

A mathematical expression for a macroscopic measurement, i.e. for the average over the myriad microscopic events performed by our senses or by our measurement tools, is given by:

$$\overline{\mathscr{O}}^T(\Gamma) = \frac{1}{T}\int_0^T \mathscr{O}(S^t\Gamma)\,dt\ , \tag{3.2}$$

where T is the duration of the measurement and Γ is the initial microstate.

In general, the result of a measurement could depend on both T and on Γ. The dependence on T makes the result of the measurement subjective, as T may be varied at will by the experimentalist, while the dependence on Γ makes it stochastic, because the initial microstate can neither be controlled nor identified. For a scientific theory to be conceived, these dependencies must play no fundamental role.

Indeed, the microscopic events occur on time scales are much shorter than the observation scales. Thus, one may hope that the quantity $\mathscr{O}(S^t\Gamma)$ explores the interval of all its possible values so rapidly that it takes a relatively short time T for $\overline{\mathscr{O}}^T(\Gamma)$ to practically equal the average of $\mathscr{O}(\Gamma)$ over all points $\Gamma \in \mathscr{M}$. The appropriate value of T depends on the sensitivity of the measurement tool: it is shorter for higher sensitivity, but it nevertheless remains virtually infinite with respect to molecular charactistic times, such as the average time between two molecular collisions. Therefore, a theory of measurements relies on the existence of the following limit:

$$\overline{\mathscr{O}}(\Gamma) = \lim_{T\to\infty}\frac{1}{T}\int_0^T \mathscr{O}(S^t\Gamma)\,dt \tag{3.3}$$

and on the irrelevance of Γ for its value.[7] Indeed, if thermodynamics applies, the properties of a macroscopic system can be predicted without any knowledge of Γ. This is essentially the content of Boltzmann's celebrated *ergodic hypothesis*.

When the ergodic hypothesis holds, $\mathscr{M}$ may be endowed with a probability density ρ reflecting the frequency with which the different regions of $\mathscr{M}$ are visited during the evolution. This probability distribution, also called *ensemble* (Fermi 1956),[8]

[7] The existence of such limits was proven by Birkhoff in 1931 for a wide class of systems, while the dependence on Γ poses many more problems, from a mathematical perspective.

[8] In his renowned small book on thermodynamics, Fermi says that: *Studying the thermodynamical state of a homogeneous fluid of given volume at a given temperature (the pressure is then defined by the equation of state), we observe that there is an infinite number of states of molecular motion that correspond to it. With increasing time, the system exists successively in all these dynamical states that correspond to the given thermodynamical state. From this point of view we may say that a thermodynamical state is the ensemble of all the dynamical states through which, as a result of the molecular motion, the system is rapidly passing.*

associates higher weights to the regions which are visited more often, and lower weights to those which are visited less frequently so that:

$$\overline{\mathscr{O}}(\Gamma) = \int_{\mathscr{M}} \mathscr{O}(\Gamma)\, \rho(\Gamma) \mathrm{d}\Gamma \equiv \langle \mathscr{O} \rangle_{\rho} \tag{3.4}$$

for all initial conditions $\Gamma \in \mathscr{M}$, except a set of vanishing probability. The quantity $\langle \mathscr{O} \rangle_{\rho}$ is called the phase average, and a given system is called *ergodic* if the time averages of *all* its observables equal their phase counterparts, as in Eq. (3.4).

The practical importance of the ergodic hypothesis is evident: knowledge of the initial microstate Γ and of the solution of Hamilton's equations, which yields $S^t \Gamma$, are unnecessary if the system is ergodic.

Given all this, a question arises: how do the averages of the microscopic properties of a system of particles relate to the thermodynamic properties of the object composed of those particles? The first, in fact, consist of a myriad of mechanical parameters, such as the particles' energy and momentum, while the others consist of just a few measurable quantities, like temperature and pressure. The corresponding diversity of fundamental terminology, qualifies the reduction of thermodynamics to mechanics as "heterogeneous" reduction (Nagel 1979), a condition which may prevent the logical derivation of the former theory from the latter. As recalled in the last chapter, for that to be possible, Nagel requires the existence of relations between the terms of the theory to be reduced and elements of the vocabulary of the reducing theory. Such bridge laws must reflect a kind of identity between the objects of study of the two theories, and must be empirically supported.

The bridge law which associates thermodynamics with the classical mechanics of atoms was proposed by Boltzmann and it is engraved in his tombstone:

$$S = k \log W. \tag{3.5}$$

This celebrated relation connects the thermodynamic entropy S of an object in the macroscopic state X, to the volume W of all microstates in $\mathscr{M}$ which correspond to the same X. For example, considering the macrostate X corresponding to a given energy E, one typically considers the energy shell $E - \delta E \leq H(\mathbf{q}, \mathbf{p}) \leq E$, with small δE, and obtains:

$$W = \int_{E - \delta E \leq H(\mathbf{q}, \mathbf{p}) \leq E} \mathrm{d}\mathbf{q}\mathrm{d}\mathbf{p}\,.$$

The microcanonical probability distribution is constant in the energy shell so that:

$$\rho(\Gamma) = \begin{cases} \frac{1}{W} & \text{if } E - \delta E \leq H(\mathbf{q}, \mathbf{p}) \leq E \\ 0 & \text{otherwise} \end{cases}$$

Equation (3.5) qualifies as a bridge law, because S is a thermodynamic quantity, while W is a microscopic entity. Once it has been introduced, further mechanical properties of our description of the microscopic dynamics may be related to as many other thermodynamic quantities, thus bridging the gap between micro- and macro-descriptions. In particular, in the *entropy represention* of thermodynamics, one obtains the temperature as:

$$\frac{1}{T} = \frac{\partial S}{\partial E} \,, \tag{3.6}$$

and the free energy as $F = E - TS$, where E is the mechanical energy of the system. Note that some authors, including Nagel, take the relation between the temperature T and the average kinetic energy K as the bridge law. This is incorrect, as thermodynamics requires a thermodynamic potential, such as the free energy, which is a function of the other relevant thermodynamic variables. In addition, the expression of T in terms of K, which had already been guessed by Bernoulli, holds only for a special class of phenomena.

The bridge law (3.5) has important consequences supported by empirical evidence, including, in particular, those derived from Einstein's version:

$$P(\alpha) \sim e^{(S(\alpha)-S_{\text{eq}})/k} \tag{3.7}$$

where S_eq is the equilibrium entropy and $S(\alpha)$ is the entropy of a spontaneous fluctuation of some collective variable (α) which may be produced by the cooperation of many microscopic motions. Formula (3.7) is meant to represent the probability of fluctuations about equilibrium states of microscopic mechanical quantities, such as the energy E of a system in contact with a heat bath at temperature T. As these fluctuations are related to observable quantities, they can be characterised by macroscopic equilibrium experiments. For instance, if α is the energy, one identifies the average $\langle E \rangle$ of the energy with the internal energy U of the system, and introduces the standard deviation $\sqrt{\langle E^2 \rangle - \langle E \rangle^2}$, which measures the size of the fluctuations. We obtain:

$$\langle E^2 \rangle - \langle E \rangle^2 = kT^2 C_v \tag{3.8}$$

where C_v is the heat capacity at constant volume. Because C_v is extensive, hence proportional to the number N of particles in the system, the relative size of the energy fluctuations is negligible in large systems:

$$\frac{\sqrt{\langle E^2 \rangle - \langle E \rangle^2}}{\langle E \rangle} \sim O\left(\frac{1}{\sqrt{N}}\right) \to 0 \,, \quad \text{for } N \to \infty \tag{3.9}$$

In other words, the fluctuations of the microscopic mechanical quantity E grow with the system size proportionally to $\sqrt{N}$, but are negligible with respect to the observable internal energy $\langle E \rangle = U$, which is of order $O(N)$. This should not lead one to relegate fluctuations to the set of only marginally interesting phenomena.

Indeed, in his search of an ultimate proof of the existence of atoms, Einstein realised that Eq. (3.8)

> would yield an exact derivation of the universal constant [k or, equivalently, Avogadro's number N_A] if it were possible to determine the average of the square of the energy fluctuations of the system.

He succesfully applied this idea to describe Brownian motion.

3.2.2 *Beyond the Mathematical Limitations of Ergodic Theory*

The effort to turn the physical notion of measurement into an appropriate mathematical one has led to the issue of ergodicity, which seems to cleverly frame the connection between mechanical and thermodynamical quantities. Unfortunately, apparently modest consideration of real-life systems, such as the insensitivity of thermodynamic quantities to microscopic states, raises deep mathematical questions. In particular, one faces the problem of identifying the probability density ρ that describes systems in equilibrium, or evolving towards equilibrium,[9] and the fact that requiring Eq. (3.4) to hold for *all* phase functions is too demanding, compared to the needs of thermodynamics. In addition, the time scales over which the ergodicity of a system of many degrees of freedom would be obtained are astronomically larger than the physically relevant time scales.

The celebrated work by Fermi, Pasta and Ulam, concerning a chain of nonlinear oscillators, further showed that ergodicity may be violated even by the simplest particles systems.[10]

Quite surprisingly, the very simple probability distributions known as *microcanonical, canonincal* and *grand-canonical* ensembles describe very well most equilibrium situations. When legitimate, this is an extremely powerful way of proceeding,[11] whose success rests on our limited knowledge of the microscpic dynamics.

---

[9] A physical system is in an *equilibrium* state if all currents—of mass, momentum, heat, etc.—vanish, and the system is uniquely described by a (typically quite small) set of *state variables* which do not change with time.

[10] In their numerical simulations, known as the FPU experiment, Fermi and coworkers showed that a typical Hamiltonian system is not ergodic. This fact was totally unexpected, at that time, and was only later explained in the sophisticated mathematical terms of KAM theory (Cencini et al. 2009).

[11] However one should be wary of possible misunderstandings. In particular, ensembles are often described as *fictitious collections of macroscopically identical copies of the object of interest, whose microstates differ from each other.* While this maybe a convenient perspective, at times, one should not forget that their purpose is to describe the properties of a single system, whose microstate evolves forever. We can say that the word "statistical ensemble" is nothing but a way to indicate the probability density of Γ.

But why? The foundations of the ergodic hypothesis look shaky, and its success puzzling, if no further explanation is given. To address these issues, Khinchin (1949) pioneered an approach based on the following premises:

(a) statistical mechanics concerns systems with a large number of degrees of freedom;
(b) the physical observables are but a few and quite special functions;
(c) it is physically acceptable that ensemble averages do not coincide with time averages, on a small set of phase space trajectories.

As appropriate for rarefied systems, he considered dynamics whose Hamiltonians are the sum of single particle contributions:

$$H = \sum_{n=1}^{N} H_n(\mathbf{q}_n, \mathbf{p}_n)$$

and restricted the space of observables to the *sum functions*—functions defined as sums of single particle contributions f_n:

$$f(\Gamma) = \sum_{n=1}^{N} f_n(\mathbf{q}_n, \mathbf{p}_n) \, .$$

The pressure and the kinetic energy are examples of such functions. Then, denoting by $\langle \, . \, \rangle$ the microcanonical ensemble average, Khinchin demonstrated that:

$$\text{Prob}\left(\frac{|\overline{f} - \langle f \rangle|}{|\langle f \rangle|} \geq K_1 N^{-1/4} \right) \leq K_2 N^{-1/4} \, ,$$

where K_1 and K_2 are constants. This means that the microcanonical averages of sum functions differ from their time averages by more than a (small) relative tolerance only along a set of trajectories whose probability vanishes in the $N \to \infty$ limit. The problem is that the initial conditions of the microstate must be taken within a proper subset of the phase space, but Khinchin showed that the fraction of volume of phase space which lies outside this subset vanishes in the $N \to \infty$ limit.

Ultimately, from various standpoints, Khinchin's theory ascribes the good statistical properties required for normal thermodynamic behaviour to the fact that N is very large. From his perspective, the details of the microscopic dynamics appear practically irrelevant for the physics of rarefied gases. An important extension of this approach, which goes beyond the low density gas, was obtained by Mazur and Linden (1963). These authors did not require the Hamiltonian to be separable in single particle contributions, but admitted particles to interact only through short-range interaction potentials and, like Khinchin, considered only sum variables. They proved that their systems can be treated as consisting of many non-interacting parts.

Although even this theory is not completely satisfactory, because the set of sum variables is too limited for dense systems, the works of Khinchin and of Mazur and

van der Linden clarify why one should not be surprised that the ergodic hypothesis applies so generally in physics: macroscopic systems are made of very many particles.

3.2.3 Summary

Before proceeding with specific illustrations of the effectiveness of SM in describing natural phenomena, let us briefly summarise the fundamental concepts discussed above. The purpose of SM is the understanding of macroscopic behaviour in terms of the properties of the microscopic constituents of matter. This has been achieved by adopting:

- the ergodic hypothesis
- the bridge law (3.5).

However, strictly speaking, the ergodic hypothesis cannot be verified, except in a few exceptional cases, as evidenced by the FPU numerical experiment and by the KAM theorem. On the other hand, Khinchin's strong mathematical results showed that Eq. (3.4) holds in the $N \to \infty$ limit, for a physically relevant class of observables. Then, the success of SM in describing macroscopic systems may be attributed to the following facts:

- although ergodicity is not exactly verified in realistic models, it does hold in a weak sense, which is sufficient for the purposes of physics;
- the bridge law $S = k \log W$ links the microscopic mechanical quantity W with the emerging thermodynamic quantity S, through the Boltzmann constant k; We stress the fact that the relation (3.5) is a fundamental assumption of the same nature (and importance) as Newton's principles for mechanics;
- macroscopic objects are made of very large numbers of microscopic constituents. The number of particles in macroscopic bodies is of the order the Avogadro's number ($N_A \approx 6.02 \times 10^{23}$). Boltzmann's constant $k = \mathscr{R}/N_A$, where $\mathscr{R}$ is the universal gas constant that takes the mind-boggling value $k \approx 1.38 \times 10^{-23}$ J/K, is an astonishingly physically powerful element of the bridge law. Because it constitutes a unit of entropy (energy divided by temperature), k binds mechanics and thermodynamics together. The very small numerical value of k measures the "distance" between the microscopic world and the macroscopic world.

3.3 The Paradigmatic Brownian Motion

Let us now turn our attention to the remarkable phenomenon that is Brownian motion. Observed under a microscope, pollen suspended in a glass of water moves erratically and incessantly, although the water appears to be still, and no work is done on pollen particles, to balance the energy dissipated by the viscosity of the fluid. This phenomenon was named after Robert Brown, the botanist who first tried to explain

it as a form of life, which seemed to animate pollen particles suspended in a fluid (Brown 1828). Among many other issues, Brownian motion constitutes the ultimate evidence of the existence of atoms and led to the determination of k in macroscopic experiments. Moreover, although it is not directly concerned with emergent properties, through Brownian motion we catch a glimpse of emergence, since the Brownian motion describes vividly the interplay of the different levels of reality at their borders (here molecules represent the microscopic level, hydrodynamics represents the macroscopic one, and pollen particles stay in between).

Why doesn't the motion of the pollen rapidly come to a halt? The equation for the velocity of one spherical particle of mass m and radius R, subjected to no other forces than that exerted by the viscosity η of the liquid,

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} + \frac{6\pi R\eta}{m}\mathbf{v} = 0, \tag{3.10}$$

predicts the exponentially decaying behaviour

$$\mathbf{v}(t) = \mathbf{v}(0)\exp(-t/\tau)\ , \quad \text{with} \quad \tau = \frac{m}{6\pi R\eta}, \tag{3.11}$$

where $\mathbf{v}(0)$ is the initial velocity, $\mathbf{v}(t)$ the velocity at a subsequent time t and τ is a characteristic time depending on the properties of both water and pollen. For pollen of radius $R \sim 10^{-4}$ m, one obtains $\tau \sim 10^{-4}$ s, which means that $\mathbf{v}(t)$ should practically vanish in a few milliseconds.

The observation made in the second half of the nineteenth century, that the velocity of pollen increases with temperature, while it decreases with the pollen size and with the fluid viscosity, suggested that the kinetic theory of gases could explain the phenomenon. However, it was unclear how to approach the problem, because the water molecules are so much lighter than pollen that they could hardly have any effect on the motion of the pollen particles. At the beginning of the twentieth century, Einstein and Smoluchowski proposed a theory, which Langevin simplified as follows: the motion of pollen is determined by two forces:

- the deterministic viscous force obtained from Stokes law,
- a stochastic force due to the collisions with water molecules, which bears no memory of events occurring at different times.

This implies that Eq. (3.10) should be modified as:

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} + \frac{6\pi R\eta}{m}\mathbf{v} = f_R(t), \tag{3.12}$$

where f_R is a random force representing the action of the water molecules on the pollen grains which, in accord with kinetic theory, is more energetic at higher temperatures. The randomness is reflected in the lack of correlations between the action of the water molecules at different instants in time and is justified by the vast separation of the temporal scales concerning microscopic impacts (of order $10^{-12} - 10^{-11}$ s)

and macroscopic viscous damping (of order $10^{-5} - 10^{-4}$ s). One then requires the average force acting on a pollen particle to vanish so that: the average work vanishes; there is no loss of energy in time; and pollen may then persist in its motion forever.

The result of this theory is the Einstein-Smoluchowski diffusion law,

$$\langle \mathbf{x}^2(t) \rangle \simeq 6Dt \ , \quad \text{where } D = \frac{kT}{6\pi\eta\mathscr{R}}, \tag{3.13}$$

in which $\mathbf{x}(t)$ is the displacement of a pollen grain at time t from its initial position, $\mathbf{x}(0)$, and T is the common temperature of water and pollen. The constant D is known as the *diffusion coefficient*.

Equation (3.13) turned out to be extremely important, since it connected easily measurable macroscopic quantities, such as $\langle \mathbf{x}^2(t) \rangle$, with Avogadro's number, which could at last be estimated.[12] Interestingly, Einstein had correctly anticipated that relations concerning fluctuations could be used to investigate the microscopic realm by means of macroscopic observations and wished, in particular, that Eq. (3.8) could one day be used to determine k (Einstein 1956).

The agreement between theory and experiments was demonstrated by Perrin only a few years later, a result which convinced practically everybody that atoms could indeed be "counted" and "measured", hence, that they had to exist. More recently, the fact that matter is made of atoms has been evidenced e.g. by scanning tunnelling microscopes. Nevertheless, the reality of atoms as constituents of matter seems to be more convincingly shown by Brownian motion than by the atomic landscapes produced by modern microscopes. The first, indeed, can be understood within a simple and clear theory, easily confirmed by direct observation. The second relies, instead, on images indirectly produced by a sophisticated technological process, whose interpretation rests rather elaborately on quantum mechanics. To prove (or disprove) the existence of atoms was indeed Einstein's purpose, as stated in the second paragraph of his paper, translated and reprinted in (Einstein 1956):

> If the movement discussed here can actually be observed (together with the laws relating to it that one would expect to find), then classical thermodynamics can no longer be looked upon as applicable with precision to bodies even of dimensions distinguishable in a microscope: an exact determination of actual atomic dimensions is then possible. On the other hand, had the prediction of this movement proven to be incorrect, a weighty argument would be provided against the molecular-kinetic conception of heat.

This means that qualitatively different representations of matter, like thermodynamics and kinetic theory, are required to describe observations which take place on the corresponding hugely different scales. Einstein's paper itself begins with a treatment of the apparently different problem of osmotic pressure—a masterly application of the molecular theory of matter— which clearly points out the question of relative scales. Amidother ingenious insights, Einstein observes that, according to the molecular kinetic theory of heat, a molecule dissolved in water is differentiated from a

[12] At that time, N_A was but a parameter of the atomic theory, whose value was unknown.

suspended body solely by its dimensions, and a number of suspended particles should produce the same osmotic pressure as the same number of molecules. To draw this conclusion, he assumes that the suspended particles perform an irregular movement, because of the molecular movement of the liquid molecules. Then the suspended particles will exert a pressure just like molecules in solution.

Such an idea is followed, for instance, in molecular dynamics simulations, which are employed equally to study the behaviour of common fluids, or of granular matter. In fact, given the limited number of "atoms" that can be presently simulated,[13] because of the limited computing power currently available, molecular dynamics is arguably closer to simulations of granular fluids than of standard fluids. If Einstein's thoughts were correct, experiments on granular matter shed light on mesoscopic phenomena, where relatively small numbers of particles are involved.

Observing Eq. (3.12), one finds that there are two different limiting situations, involving the mass of pollen particles. The first is the limit of large mass, which makes the effect of the molecular impacts negligible, compared to that of the fluid viscosity. The second is that of small pollen mass, which makes the effect of the molecular impacts dominant, with respect to the viscous forces. These two situations seem to delimit all possibilities, ranging from suspended particles of the same size as the fluid molecules (hence obeying microscopic reversible laws), performing an effectively random motion and experiencing no collective viscous damping, to macroscopic objects in a liquid, which are insensitive to molecular impacts and only feel the viscosity of the fluid.

Therefore, Brownian motion seems to lie at the border of the microscopic and the macroscopic worlds, revealing that this border is populated by a great variety of intermediate phenomena, which depend on the size of the objects of interest. Does this constitute one example of successful reduction of one level of description of matter to another?

Not really. Instead, the theory of Brownian motion unveils one level of description, the mesoscopic level, which is as hard to connect to the microscopic and the macroscopic levels. The ingenuity of the Einstein-Smoluchowsky theory lies in its ability to identify three separate scales concerning objects in thermodynamic equilibrium, and to daringly mingle them, by allowing microscopic and macroscopic forces act at once on pollen, and by using kinetic theory to determine the viscosity of the fluid. The phenomena which can be observed at the different scales, obviously coexist, but correspond to such widely separated scales that completely different kinds of description are required to understand them. This is what makes Brownian motion possible: the separation of scales. The mass of pollen is so much larger than that of liquid molecules, that the relaxation time τ of Eq. (3.11) is much larger than the molecular collision times and, at the same time, the mass of pollen is so much lighter than any macroscopic object, that the energy exchanged in molecules-pollen impacts suffices for thermal equilibrium to be established in times much shorter than τ.

[13] Up to the order of billions of particles.

To pretend that pollen agrees at once with both kinetic theory[14] and hydrodynamics[15] manifested Einstein's ingenious insight into the atomic hypothesis, and revealed an intricate texture of levels of observation, rather than the reduction of one level to another. Furthermore, in the theory of Brownian motion, macroscopic observations shed light on microscopic events, not the other way round.

The transition from one level of description to the other amounts to operating on the ratio $\delta = m_w/m$ of the masses of water molecules and of pollen grains. Mathematically, taking such a ratio to zero is far from a harmless operation since, technically, it is a singular limit, as evidenced, e.g. in (Primas 1998; Berry 1994). The singularity means that the limiting behaviour does not coincide with the behaviour obtained by setting $\delta = 0$ to begin with. In physics, this is often related to the fact that the limit one wants to consider is a sophisticated sequence of operations, which cannot be summarised in a single action.

As mentioned above, the hypothesis that water constitutes a heat bath in which pollen equilibrates requires that many interactions (collisions) take place between pollen and molecules in a sufficiently short time. If the mass of a water molecule is relatively very small, very many interactions are required to obtain the necessary energy exchange, and the equilibration time is very large. But, as long as δ does not vanish, Brownian motion takes place, although the observation time scales will increase with the mass of pollen. Therefore, the $\delta \to 0$ limit cannot be separated from the limit of large times. If δ identically vanishes, equilibration between suspended object and molecular motions will never be realised: however long long one waits, equilibrium will not be established and the observed phenomenon will differ in nature from Brownian Motion.

The singularity of the limit highlights the boundary between two different levels of description of a natural phenomenon, and foretells the emergence of new properties, preventing plain reduction of one description to the other.

As Brownian motion ultimately proves that matter is made of atoms, the macroscopic observable behaviours should *logically supervene*[16] the evolutions of atoms. This does not mean that the descriptions which have been developed to understand matter at different scales must reduce to one another; following Deutsch, this could be impossible and uninteresting. A building is certainly made of bricks, like skin is made of skin cells. But looking so closely at a building that the single bricks can be identified, like looking at a histologic specimen in a microscope, does not help to distinguish a cathedral from a factory, or a chicken from a human being. A higher level picture is needed for that and, yet, our understanding of the building or of the living organism would be much poorer if we did not know how bricks or

[14] Assuming that the pollen and the molecules have the same average kinetic energy.

[15] Assuming the validity of the Stokes law.

[16] Logical supervenience, in our framework, means that it is logically impossible to produce different macroscopic states with identical microscopic states. This requires the microstate to fully determine the properties of the macrostate.

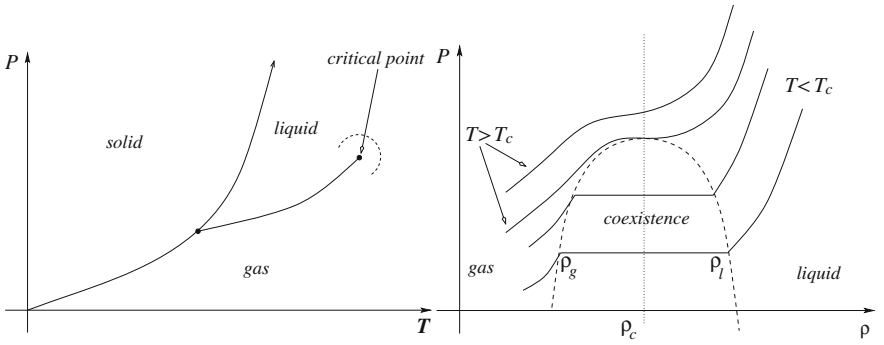

Fig. 3.1 Phase diagrams and critical point

cells have been assembled. Analogously, searching for the connections between the different descriptions of the structure of matter remains necessary to advance our understanding of the material world.

3.4 Critical Phenomena

Critical phenomena may appear a merely technical and involved chapter of SM. In reality, they are conceptually very important, because they conspicuously demonstrate how a variety of macroscopic phenomena do not depend on microscopic details.

Consider, for instance, the phenomena known as first-order phase transitions, in which the equilibrium state of a given macroscopic system is not uniquely specified by the usual thermodynamic quantities, such as the temperature and the pressure. The state is not uniquely specified becasuse two different phases characterised by different properties, e.g. liquid and solid or solid and gas, may equally exist at a phase transition and, typically, they coexist. Furthermore, these phase transitions are characterised by discontinuous changes of certain material properties, as in the case of ice turning liquid, whose density changes discontinuosly.

In the case of subtler transitions, such discontinuities do not occur, but the overall behaviour of the object of interest nevertheless changes drastically. These are called second order phase transitions and are related to the presence of *critical points* in the phase diagram for equilibrium states, e.g. the PT plane for systems whose equation of state takes the form $f(P, \rho, T) = 0$, where P is the pressure, ρ the density and T the temperature. In the PT plane of Fig. 3.1, one observes that the liquid-vapour co-existence line does not extend forever; it stops at a critical point, of critical thermodynamic coordinates P_c, ρ_c and T_c, around which the liquid phase turns continuously into gas and the distinction between the phases fades away. The $P\rho$ plane, at T close but less than T_c, shows that the liquid density ρ_l is larger than the gas density ρ_g, and that the difference $m = \rho_l - \rho_g$ (the order parameter) vanishes as

Table 3.1 Several phase transitions with critical point

Transition	Order parameter	Meaning
Liquid–gas	$\rho - \rho_c$	Difference of density and critical density
Ferromagnetic	m	Magnetisation
Fluid-superfluid	ψ	Condensate wave function

T tends to T_c. Furthermore, the isotherms become straighter and straighter (i.e. closer and closer to the ideal gas case) as T continues to grow above T_c. This reveals that their curved shape is due to the interactions among the components of the system, and the nature of these interactions close to the critical point produces a characteristic power-law behaviour:

$$m = \rho_l - \rho_g \sim \varepsilon^{\beta} \qquad \text{where } \varepsilon = \frac{T - T_c}{T_c} \tag{3.14}$$

The most striking feature of these phenomena is the existence of "critical" exponents, such as β in Eq. (3.14), which are largely independent of the details of the microscopic dynamics of the elementary constituents of the system of interest, hence are called "universal". This situation is common to all physical systems near their own critical points, and not just to the case of liquids and gases. Table 3.1 gives the order parameters relevant to a few quite different transitions.
Fundamental information on critical phenomena is provided by the correlation functions. For instance, let $m(\mathbf{r})$ be the magnetisation density about the point $\mathbf{r}$. The auto-correlation of this quantity with respect to the reference point $\mathbf{0}$ is defined by

$$C_T(\mathbf{r}, \mathbf{0}) = \langle m(\mathbf{r}) m(\mathbf{0}) \rangle - \langle m(\mathbf{r}) \rangle \langle m(\mathbf{0}) \rangle \tag{3.15}$$

The quantity C_T vanishes if the value of m at $\mathbf{r}$ has no relation with its value at $\mathbf{0}$, hence it measures how the "memory" of m persists in space. One well-established form of correlation function for isotropic fluctuations and $T > T_c$, is the following:

$$C_T(\mathbf{r}, \mathbf{0}) = \frac{e^{-r/\xi_T}}{r^{d-2+\eta}} \tag{3.16}$$

where r is the length of $\mathbf{r}$, the correlation length ξ_T represents the extent of the spatial memory, d is the spatial dimension and η is an exponent whose precise value is not particularly important here. Experimentally, one finds that ξ_T diverges like a power law, $\xi_T \sim \varepsilon^{-\nu}$, i.e. the whole system becomes strongly correlated, in the $T \to T_c$ limit. If one assumes that ξ_T is the only characteristic length around T_c, its divergence for $T \to T_c$ means that the system has no characteristic lengths at T_c and, therefore, must be invariant under scale transformations. In other words, the phenomenon appears to be the same at all scales.

Table 3.2 Critical exponents (β, ν, α, γ and δ) of magnetic systems, for several different qantities

Quantity	Behaviour
Zero-field magnetisation M	$(-\varepsilon)^{\beta}$
Correleation length ξ_T	$\varepsilon^{-\nu}$
Specific heat C_h	$\varepsilon^{-\alpha}$
Isothermal susceptibility K_T	$\varepsilon^{-\gamma}$
Critical isotherm H	$\lvert M\rvert^{\delta}\text{sign}(M)$

This suggests that, close to a critical point, the thermodynamic quantities may be expressed as a sum of a finite part and a singular part which either diverges, or has divergent derivatives, proportional to ε raised to the critical exponents. Table 3.2 lists some of the parameters involved in this phenomenon.

Why should such importance be placed on so limited a region of the thermodynamic phase diagram as that asymptotically close to a critical point? The fact is that the behaviour of the different observables characterised by the values of the critical exponents, are not particularly interesting of their own merits. Their importance lies rather in their universality: experiments show that a wide range of systems, with very different critical temperatures, share the same critical exponents. Furthermore, different exponents referring to different phenomena happen to be simply (linearly) related to each other. This is a very interesting fact: despite the very complicated microscopic interactions differing from case to case, the thermodynamic properties of different ferromagnetic systems, such as the temperature dependence of the specific heat and the susceptibility, are, remarkably, observed to be the same. Naturally, not all known critical phenomena are characterised by the same exponents, but they can be grouped into just a few "universality classes", within each of which the exponents do not vary. These classes depend only on the dimensionality and the symmetries of the systems of interest, and not on the details of the Hamiltonians. It is striking that very different kinds of systems, such as fluids and magnets, fall within the same universality class because their experimentally measured critical exponents are equal. For instance, the liquid-gas transition of CO_2 is characterised by $\beta \approx 0.34$, $\delta \approx 4.2$ and $\gamma \approx 1.37$, while the ferromagnetic transition of Nickel is characterised by $\beta \approx 0.33$, $\delta \approx 4.2$ and $\gamma \approx 1.35$ (Kadanoff et al. 1967).

Unfortunately, to obtain the phase diagram of a system of interacting particles, whose microscopic dynamics are described by a given Hamiltonian, is in general an exceedingly difficult task. Therefore, theorists have devised simplified and often highly idealised models, which allow explicit calculations to be performed. In particular, models known as lattice systems are used equally to approximate continuum systems (such as liquids) or to represent spins in magnetic materials.

Among the idealised models, the Ising ferromagnet consists of an array of spins which take the values $+1$ or -1 at the sites of a regular lattice.[17] When viewed on the

[17] The Ising model consists of a discrete set of positions in space representing the lattice of the atoms of a crystalline solid, in each of which the "spin" variable S_i takes either value 1 or -1. The spins are assumed to interact in pairs of nearest neighbours, with interaction energy that depends on whether the neighbours have same value or opposite values. In the presence of a "magnetic field",

scale of single spins, the magnetic moments on the lattice sites of a crystal and the subsequent ordering of the moments present complicated behaviours, which depend on many different parameters. But the thermodynamic (collective) critical properties of completely different physical systems, such as an Ising ferromagnet and a liquid at its critical point, show the same dependence on the temperature.

Which physical mechanism leads to this kind of universality? Given a ferromagnetic material and measuring its magnetisation, one finds that the macroscopic properties of each of its halves is the same as that of the whole, if the object is divided in two keeping the same temperature, magnetic field etc. One may repeat the process until something different eventually happens; the characteristic length scale at which the overall properties of the pieces of the magnet begin to differ markedly from those of the original object defines a correlation length. This is the length scale over which the fluctuations of the microscopic degrees of freedom are correlated. Now a ferromagnet may abruptly change its macroscopic behaviour if the external conditions vary, e.g. to reach a phase transition. As we have discussed, there are two ways in which the transition may occur, and which differ substantially, as far as the correlation lengths are concerned:

- the states separated by the line which (possibly) ends at the critical point coexist at the phase transition. In this case, a thermodynamic property changes discontinuously and we have a first-order phase transition, characterised by finite correlation lengths;
- the transition is continuous and correlation lengths grow without bounds: fluctuations become correlated over all distances and the whole system appears to have reached a unique, critical, phase. The two phases on either side of the transition line become indistinguishable as the critical point is approached. Therefore, as the correlation length diverges, the magnetisation goes smoothly to zero, and we have a second-order transition.[18]

(Footnote 17 continued)
the Hamiltonian of the system is given by:

$$H = -J \sum_{\langle i,j \rangle} S_i S_j - h \sum_i S_i$$

where J is the coupling between spins in different positions, h is the coupling with the magnetic field and $\langle i, j \rangle$ denotes summation over nearest neighbours only. If $J > 0$, the model is described as ferromagnetic; if $J < 0$, the model is anti-ferromagnetic. In the first case, spins tend to align, and the "material" to become magnetised. The equilibrium states of this system are described by the canonical ensemble. The Ising model is the first in which phase transitions could be explicitly proven to occur in the thermodynamic limit (cf. below), provided the dimensionality is higher than 1. The solution of the 1-dimensional model was obtained by Ising in 1925, while the much more difficult solution of the 2-dimensional case was derived by Onsager in 1944.

[18] In this case, the divergence of the correlation length suggests that near the critical point one must resort to a theory based on long-range collective fluctuations and on Hamiltonians or free energies constrained only by the fundamental symmetries of the system.

This leads to the conclusion that the details of the microscopic interactions are largely irrelevant for many questions of physical interest; in particular, they are not needed to obtain qualitative properties such as the scaling behaviours near the critical points. The same qualitative properties in the vicinity of a critical point are produced by a large class of different microscopic structures and dynamics. In other words, a limited number of universality classes, identified by a few fundamental properties, such as symmetries and dimensionality, contain the vast majority of possible behaviours. Indeed, the critical properties of the Ising phase transition in a one-component ferromagnet, which is a highly idealised model, is qualitatively the same as a liquid-gas transition. Similarly, in the jargon of statistical field theory, under certain circumstances a superconductor, with its complex order parameter, is in the same universality class as the Heisenberg ferromagnet (Kadanoff et al. 1967).[19]

Our description of these phenomena and of phase transitions relies, once again, on limiting operations which cannot be verified exactly by any physical system, but which are closely approximated, within our observation space and time scales, by macroscopic objects. The relevant limit, in this case, is the *thermodynamic limit*, i.e. a formal procedure in which the size of the system of interest and the number of its microscopic components grow together without bounds, in such a way that the number of particles per unit volume and the energy per unit volume remain constant. This represents an idealisation of a macroscopic system which, although necessarily finite, is very large on the scale of microscopic interactions. The thermodynamic limit is useful because it gets rid of boundaries and finite-size effects, which are mathematically problematic and practically irrelevant to the phenomena of interest. Indeed, neglecting to give the size and shape of the object of investigation, in reporting the results of a measurement, every experimentalist implicitly assumes that boundaries and finite-size effects are negligible, i.e. that the thermodynamic limit affords a good description of the object (Lebowitz 1999).

Interestingly enough, the details of the microscopic interactions and structure, wrapped up by the Hamiltonian, are necessary to determine T_C, something that constitutes quite a difficult problem. This was done by Onsager, for the two dimensional Ising model.[20] In such a case, one could conclude that the reductionistic programme actually succeeded. Unfortunately, this conclusion is thwarted by the fact that the partition function required for the calculation cannot be constructed, in general. Not even in the case of the three dimensional Ising model can the partition function be explicitly obtained. Furthermore, the Hamiltonian itself is, in general, unknown. On

[19] The Heisenberg ferromagnet has a structure similar to that of the Ising model, with Hamiltonian given by:

$$H = -J \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j - \sum_i \mathbf{S}_i \cdot \mathscr{H}$$

where each $\mathbf{S}_i$ is a vector of unit length in $\boldsymbol{R}^2$ or $\boldsymbol{R}^3$, the dot indicates the scalar product and $\mathscr{H}$ is an external magnetic field.

[20] Apart from the tremendous mathematical effort it required, the importance of Onsager's exact solution is that it first demonstrated that the formalism of SM can describe phase transitions and critical phenomena.

the other hand, the dimensionality and the symmetry properties of the Hamiltonian suffice to identify the universality classes, and that suffices to obtain good estimates of the critical exponents.

3.5 Discussion

In this chapter we swept through the history and the foundations of SM, whose prodromal clues date back to the mid-eighteenth century, and which imperiously sprang to life in the midnineteenth century. This history is characterised by extraordinary success, which advanced our understanding of the structure of matter and of our own representations of the world. The fact that atoms are more than a convenient mathematical tool in describing material objects, is one of the first and greatest achievements of SM, motivating the majority of the subsequent research in physics Emch and Liu (2002). The main ideas reported in this chapter can be summarised as follows.

- The world around us appears complex, especially in the sense of presenting ever more unexpected and intrinsically unpredictable facets to our eyes. This seems to be the result of interactions among the many constituents of almost any object of interest. The vastness of the number of elementary constituents does not in and of itself account for complex behaviour. Actually, if these numbers become exceedingly large, such as of the order of Avogadro's number, the collective behaviour typically becomes relatively predictable and independent of many details of the microscopic dynamics.
- This suggests the idea that matter could be made of very many simple constituents, interacting with each other in equally simple fashions: atoms. Daniel Bernoulli revived this ancient view, in order to turn it into a scientific theory; the works of Maxwell and Boltzmann made it quantitatively predictive.
- To connect the properties of the microscopic world to those of thermodynamics, atoms were boldly assumed to be minute particles obeying the deterministic laws of classical mechanics. This was motivated by the extremely successful applications of classical mechanics which, however, considered totally different realms, such as astrophysics. Thermodynamics was then supposed to emerge solely from the dynamical properties of atoms.
- This supposition amounts to an instance of heterogeneous reduction, which requires a bridge law. The appropriate bridge law is given by Boltzmann's relation (3.5), which connects entropy and phase space volumes. Through the entropy representation of thermodynamics, all thermodynamic observables are then connected to microscopic dynamical quantities. In philosophical language it is probably fair to say that equilibrium thermodynamics weakly emerges from microscopic level.
- The separation of scales which characterises, in both space and time, the microscopic, the mesoscopic and the macroscopic realms, is the main ingredient in the treatment of phenomena such as Brownian motion. Without this perspective, that

phenomenon, which demonstrates the existence of atoms, would be too complicated to investigate.

- Taking classical dynamics seriously, mathematical problems arise: ergodicity becomes a necessary ingredient of the theory, but the straightforward mathematical formulation of ergodicity is too strong a condition to be satisfied by any system of physical interest. Furthermore, ergodicity appears to be a property that pertains as much to systems comprising few degrees of freedom as comprising very many. In fact, in systems with many degrees of freedom it completely misses the physical time scales. Khinchin explained how this problem may be overcome, replacing the strict notion of ergodicity with one which capitalises on the large number of degrees of freedom and the reduced number of relevant observables.
- Critical phenomena eventually demostrated that the study of the structure of matter may be legitimately pursued by means of simple, properly conceived models. The large numbers typical of SM and the limited set of relevant observables allows many different phenomena to be collected in a small number of universality classes, which are not sensitive to the details of the real dynamics.

The role of mathematical modelling and its effectiveness in describing the world around us is another example of what SM helped clarify. Likewise, questions posed by SM, such as that of irreversibility (cf. the next chapter), clarify the idea of emergence, which has been popularised by the (somewhat pretentious) all-embracing sciences of complexity. SM clearly shows that emergence is a *borderline* product of the singularity of limits that must be taken in order to jump from one level of description to another. This explains why strict reduction should not be expected, except in very special circumstances, and why it is not even necessary. Matter is doubtless made of incredible numbers of interacting microscopic elements, but the theories of the different levels of observation can only be required not to be incompatible.

References

Berry, M.V.: Asymptotics, singularities and the reduction of theories. In: Logic, Methodology and Philosophy of Science IX, p. 597 (1994)

Boscovich, R.J.: A Theory of Natural Philosophy. MIT Press, Cambridge (1966)

Brown, R.: A brief account of microscopical observations made in the months of June, July, and August, 1827 on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. Philos. Mag. **4**, 161 (1828)

Brush, S.G.: The Kind of Motion We Call Heat. North Holland, Amsterdam (1986)

Carnot, S.: Reflections on the Motive Power of Fire and Other Papers on the Second Law of Thermodynamics. Robert Drummond, New York (1897)

Cencini, M., Cecconi, F., Vulpiani, A.: Chaos: From Simple Models to Complex Systems. World Scientific, Singapore (2009)

Cercignani, C.: Ludwig Boltzmann: The Man Who Trusted Atoms. Oxford University Press, Oxford (2006)

Deutsch, D.: The Fabric of Reality. Penguin Books, London (1997)

Einstein, A.: Investigations on the Theory of the Brownian Movement. Dover Publications, New York (1956)

Emch, G.G., Liu, C.: The Logic of Thermostatical Physics. Springer, Berlin (2002)
Fermi, E.: Thermodynamics. Dover Publications, New York (1956)
Feynman, R.P., Leighton, R.B., Sands, M.: The Feynman Lectures on Physics. Addison-Wesley, Reading (1964)
Kadanoff, L.P., Götze, W., Hamblen, D., Hecht, R., Lewis, E.A.S., Palciauskas, V.V., Rayl, M., Swift, J., Aspnes, D., Kane, J.: Static phenomena near critical points: theory and experiment. Rev. Mod. Phys. **390**(2), 395 (1967)
Khinchin, A.I.: Mathematical Foundations of Statistical Mechanics. Dover Publications, New York (1949)
Laplace, P.S.: Essai philosophique sur les probabilités. H. Remy (1829)
Lebowitz, J.L.: Boltzmann's entropy and time's arrow. Phys. Today **46**, 32 (1993)
Lebowitz, J.L.: Statistical mechanics: a selective review of two central issues. Rev. Mod. Phys. **71**, S346 (1999)
Mazur, P., van der Linden, J.: Asymptotic form of the structure function for real systems. J. Math. Phys. **4**, 22–271 (1963)
Nagel, E.: The Structure of Science. Hackett Publishing Company, Indianapolis (1979)
Poincaré, H.: Science and Method. Thomas Nelson and Sons, London (1914)
Primas, H.: Emergence in exact natural science. Acta Polytech. Scand. **91**, 83 (1998)

Chapter 4
From Microscopic Reversibility to Macroscopic Irreversibility

Nothing is more practical than a good theory.

L. Boltzmann

4.1 The Problem of Irreversibility

The idea that natural phenomena proceed in a well-defined temporal direction, and therefore that the past is clearly distinguishable from the future, is based on indisputable empirical evidence. In everyday life, over and over we experience phenomena clearly indicating that natural processes are intrinsically irreversible. It suffices to think about the dispersion of ink in a glass of water, the melting of a sugar lump in coffee, or the shattering of a plate on the floor.

Anyone who has filmed the breaking of a glass, the lighting of a match, or a diver jumping into a swimming pool, knows that rewinding the film produces something that looks impossible and a bit ridiculous: the pieces of glass recombine, the ashes and the flame disappear while the match is restored, the splashing water is restored to form a flat water surface, while the diver returns to the diving board.

One might think that some sophisticated procedure could actually do the apparently impossible. For instance, could one unscramble an egg? Someone would say: "*Yes: simply feed the scrambled egg to a hen!*" The idea is that the hen is a very sophisticated machine which can produce an unscrambled egg, through complex biophysical processes. However, this does not solve the problem. In the first place we do not get back the original unscrambled egg, since part of it is lost in the digestion process. But, leaving aside details irrelevant to someone who simply wanted a boiled egg instead of a scrambled one, the modifications undergone by the environment that allow the hen to produce a new egg cannot be undone. Indeed, they come at a much higher environmental cost than the egg can afford. One way or another, the new situation never matches the previous one.

S. Chibbaro et al., *Reductionism, Emergence and Levels of Reality*,
DOI: 10.1007/978-3-319-06361-4_4,

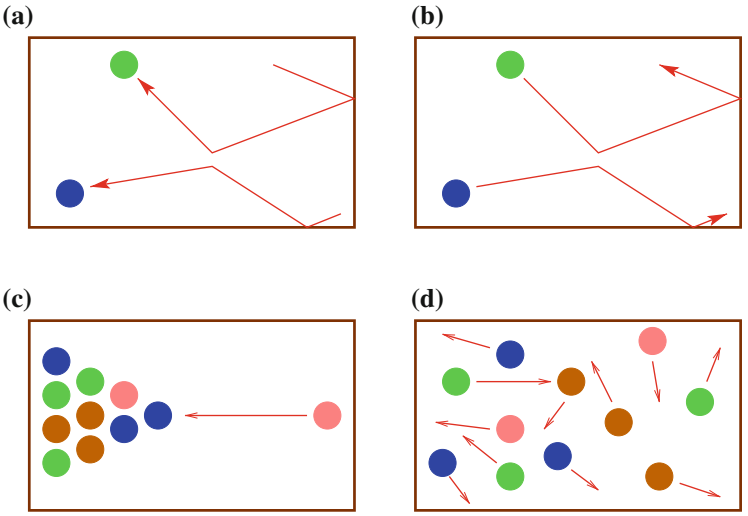

Fig. 4.1 Looking at time evolutions of two interacting objects, such as those represented by panels **a** and **b**, one cannot tell whether the film is played normally or is rewinding. In the case of even moderately many particles, as in panels **c** and **d**, there is no doubt that the situation on the left precedes the one on the right

In contrast, the reverse sequence of images of phenomena concerning a small number of objects, such as the oscillations of a pendulum or the collisions between elastic spheres, does not appear unnatural, because the forward sequence looks exactly the same. This is clearly illustrated by pictures of balls moving on a billiard table, cf. Fig. 4.1. In the case of two interacting balls, one cannot tell whether picture (a) or picture (b) represents the rewinding of the film. But, pictures (c) and (d) leave no doubt about their chronological order, and the situation becomes clearer if the number of particles increases.

Macroscopic phenomena, which involve enormous numbers of interacting particles, seem to be clearly distinguishable from microscopic phenomena, where one considers only small sets of particles. Yet, the particles are of the same kind and follow the same dynamical rules in both cases. For instance, the evolution laws describing phenomena such as heat conduction and diffusion have an intrinsically irreversible structure, and the second law of thermodynamics states the existence of a well-defined arrow of time: heat flows spontaneously only from hot bodies to cold ones; substances in solution always diffuse from regions of high concentration to regions of low concentration.

By comparison, phenomena such as those concerning the simple pendulum or the motion of planets seem to make no distinction between forward and reverse motions. But even a small part of a macroscopic system does not appear to behave irreversibly: the motion of one or two molecules of a gas does not reveal the direction of time, as in panels (a) and (b) of Fig. 4.1.

These remarks pose an evident problem for mechanistic interpretations of the world, which assume that the ultimate constituents of matter are atoms that interact according to some kind of mechanical laws. Indeed, both classical mechanics and quantum mechanics, which are supposed to describe the motions of the elementary constituents of matter, are time reversal invariant, i.e. they allow the reverse evolution of every allowed forward evolution; it is just a matter of starting from the proper initial conditions. The reversible structure of the known mechanics invites a detailed discussion of the dichotomy between irreversible macroscopic phenomena and fundamental mechanical laws, which are thought to apply equally to large or to small assemblies of particles.

In this chapter we shall discuss this issue and its connections with reductionism in some detail. On this topic, the literature is often confusing. For example, at times it is claimed that Boltzmann derived macroscopic irreversibility from reversible microscopic dynamics, hence that Boltzmann's work can be interpreted as a paradigmatic case of reductionism. We will see, instead, that irreversibility can be seen as an emergent property in the macroscopic limit. This is a property that does not depend only upon the microscopic dynamics, but requires additional conditions: a large number of degrees of freedom and appropriate initial conditions.

This is somewhat ironic: the research program delineated by Boltzmann, which could originally have seemed to be reductionist since he set out to explain the second law of thermodynamics from the laws of atomistic dynamics, ended up recognising the need of additional ingredients, which are extraneous to the microscopic mechanical laws. This is an example, in our opinion, of the validity of reductionism as a research program, if accompanied by non-dogmatic visions. Indeed, while leading to the failure of literal reductionism (e.g. of the reduction of irreversibility to mechanics), reductionism succeeds as a methodology which gives birth to a new independent discipline (e.g. to statistical mechanics).

For simplicity, we only discuss the case of classical physics based on Newton's equations of motion. Newton's second law specifies the equations governing the evolution of a system of N particles, of masses $m_1, m_2, \ldots, m_N$, that interact with each other:

$$m_i \frac{d^2 \mathbf{x}_i}{dt^2} = \mathbf{F}_i(\mathbf{x}_1, \ldots, \mathbf{x}_N) = -\frac{\partial U(\mathbf{x}_1, \ldots, \mathbf{x}_N)}{\partial \mathbf{x}_i}, \tag{4.1}$$

where $U(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ is the potential energy and $\mathbf{F}_i$ is the force acting on the i-th particle. Let us evolve the system according to Eq. (4.1) from an initial condition $(\mathbf{x}_1(0), \ldots, \mathbf{x}_N(0); \mathbf{v}_1(0), \ldots, \mathbf{v}_N(0))$ up to a specific time $t > 0$. This is equivalent to playing the film forward in time. At the instant t, time is "reversed" by performing the following operation: the positions remain equal to those attained at time t, $\mathbf{x}_1(t), \ldots, \mathbf{x}_N(t)$, while the velocities are replaced by their opposites, i.e. the quantities $\mathbf{v}_1(t), \ldots, \mathbf{v}_N(t)$ are replaced by $-\mathbf{v}_1(t), \ldots, -\mathbf{v}_N(t)$. The system then evolves in time from this new initial condition. The result is that the system traces back in space its previous trajectory as it evolves forward in time. In a sense, this the is mathematical analogy of playing the film backwards. Since Newton's equations are invariant with respect to the time reversal transformation

$$\mathbf{x}_i \to \mathbf{x}_i \ , \quad \mathbf{v}_i \to -\mathbf{v}_i \ , \quad t \to -t \ , \tag{4.2}$$

the direct and reverse evolutions of the forward and backward trajectories are equally possible: the system traces back its history and a time t after the "reversal" it returns to the same initial position, but with velocities reversed.

Given its importance, especially in contexts other than those of physics, the issue of irreversibility has been at the heart of impassioned controversies between scientists and philosophers for a long time. The debate on the arrow of time has often been heated, and many, even diametrically opposed resolutions have been proposed. On the one hand, it is hard to give up the reversibility of fundamental laws, but a few—including Einstein—have thought that the distinction between past, present and future is just an illusion, however tenacious. It is difficult to abandon the intuitive perception of a reality which develops and evolves; in particular, biological, geological, and astrophysical processes suggest that irreversibility is not purely subjective. Who would ever dare say that Hiroshima was an illusion?

4.1.1 Boltzmann and Irreversibility

According to many (including the authors of this book), Boltzmann made the first major contribution to reconciling the irreversibility observed at the macroscopic level with the reversibility of evolution laws governing the microscopic level. His H-theorem and the subsequent developments constitute an explanation which originated from the debate generated by various criticisms to Boltzmann's work (Cercignani 2006). Under suitable assumptions which we will analyze later, Boltzmann derived the evolution law known as the *Boltzmann equation*, for a gas of molecules treated as elastic spheres. This equation determines the form of the distribution function $f(\mathbf{x}, \mathbf{v}, t)$, i.e. the probability density of finding a molecule with velocity around $\mathbf{v}$ and position around $\mathbf{x}$, at time t. This equation reads:

$$\frac{\partial f}{\partial t} + (\mathbf{v} \cdot \nabla)\, f + \left(\frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}}\right) f = \int \left(f' f_1' - f f_1\right) |\mathbf{v} - \mathbf{v}_1| \frac{d\sigma}{d\Omega} d\mathbf{v}_1 d\Omega \tag{4.3}$$

where $\mathbf{F}$ is the external force acting on the single particle at position $\mathbf{x}$, and $\nabla_{\mathbf{v}}$ is the gradient with respect to the components of the velocity. The integral in the right hand side of the equation takes into account the variation of f produced by the collisions between two particles, where $\mathbf{v}$ and $\mathbf{v}_1$ are the two velocities before the interaction and f and f_1 the corresponding distributions, while the same quantities after the interaction are denoted by the the prime. Furthermore, $d\sigma/d\Omega$ is the differential cross section of the collision process. The validity of Eq. (4.3) implies the validity of the so-called H-theorem, which states that the quantity

$$H(t) = \int f(\mathbf{x}, \mathbf{v}, t) \ln f(\mathbf{x}, \mathbf{v}, t) d\mathbf{x} d\mathbf{v}$$

decreases monotonically in time, until it reaches its minimum value. More precisely

$$\frac{\mathrm{d}H}{\mathrm{d}t} \leq 0 \tag{4.4}$$

where equality holds only when f becomes the Maxwell-Boltzmann distribution:

$$f_{MB}(\mathbf{x}, \mathbf{v}) \propto \exp\left(-\frac{v^2}{2mkT}\right).$$

Here k is Boltzmann's constant and T the thermodynamic temperature of the gas. Noting that the entropy S is proportional to $-H$, one concludes that the entropy never decreases and reaches its maximum for the Maxwell-Boltzmann distribution (Cercignani 2006). Therefore, one may think that the H-theorem provides a "proof" of the second law, one of the fundamental laws of macroscopic physics, in terms of kinetic theory.

4.1.1.1 Paradoxes (Apparently) Against Boltzmann

The quotation marks around *proof*, above, indicate that things are not so simple. The first solution to the problem of irreversibility proposed by Boltzmann clashes with the recurrence paradox, formulated by Zermelo, and with the reversibility paradox usually attributed to Loschmidt (Cercignani 2006). Zermelo noted that the H-theorem is in disagreement with Poincaré's recurrence theorem, according to which the states of mechanical systems whose motion occurs in a bounded region of phase space,[1] return as close as one wants to the initial state. In other words, given a distance $R > 0$ between points in the phase space, there is a finite time $T_R > 0$ at which the state is closer than R to the initial conditions. Because the distribution function f, and therefore H, depend smoothly on the position and velocity of molecules, H must necessarily assume a value close to its initial value when the system returns close to H(0). Therefore, any growth of the entropy would be followed by a decrease. Zermelo proposed this paradox in 1896, but Poincaré (1893) had published first. In a popular essay, he stated:

> An easily established theorem states that a bounded world, subjected to the laws of mechanics, will pass by its initial state over and over. On the contrary, the known experimental laws (if one takes them as universally valid, and pushes them to their ultimate consequences), imply that the universe tends towards a final state from which it will never escape. In this final state, which is a kind of death, all bodies will be at rest with same temperature.

Poincaré also anticipated Boltzmann's reply to Zermelo's objection. Indeed, he stated that:

> the world will remain in that state only for an enormously long period of time, which will be correspondingly longer for larger numbers of molecules.

[1] Which means that the dynamics remain bounded both in position and velocity.

The paradox of reversibility may be formulated as follows. Let H decrease from time 0 to time t. At time t, reverse the velocities and start a new evolution with the initial condition thus generated. Then, due to the symmetry of the equations of motion under time reversal, the system should trace back its history and one finds that H increases for a time not shorter than t. These difficulties concerning the H-theorem are not related to the mathematical derivation of (4.4) from the Boltzmann equation, which is a simple excercise. The difficulty is instead hidden in the derivation of the Boltzmann equation itself, as will be discussed below. Boltzmann himself provided the first answers to the problems raised by the two paradoxes, which were later proven to be substantially correct. In regard to the recurrence paradox, he noticed that Poincaré's recurrence time is enormously long in macroscopic systems: so long that recurrences will never be observed in the lifetime of our universe. For example, given a cubic centimetre of gas at room pressure and temperature, recurrence with an accuracy of 10^{-9} m in positions and 1 m/s in speed could take as long as $10^{10^{19}}$ years, which is an extremely long time, even compared to the age of the universe.

The paradox of reversibility arises, instead, only if the particles' velocities are exactly reversed. This amounts to such an exceedingly delicate preparation of the system of interest that is impossible to achieve. Indeed, preparing the system for a given experiment means restricting the values of a few macroscopic observables, accessible to our manipulations. Consequently, very many different microscopic states, and not necessarily the desired one, are obtained when preparing the same macrostate. Producing a reversal of the velocities of all particles would require a direct action on the microscopic state, which is impossible to do. Therefore, Boltzmann answered that the validity of the H-theorem rests on the somewhat simpler fact that there are many more microscopic configurations which lead to a decrease of H, than configurations leading to its increase.

This reasoning may sound vague, even contradictory, and not entirely satisfying. However, Ehrenfest and Ehrenfest (2002) and Kac (1957) devised simplified probabilistic models through which it is possible to show that Boltzmann's intuition was substantially correct and, indeed, mathematically rigorous in the limit of infinitely many particles.

4.1.2 Different Ideas About Irreversibility

Boltzmann's deep and subtle argumentations have not been universally accepted. There continue to coexist a variety of views regarding the very origin of irreversibility (Guerra 1980), which we summarise as follows:

(a) Irreversibility is a basic law of nature and should be taken into account in addition to Newton's equations (or to the Schrödinger equation in the quantum case). This point of view finds support in the existence of the temporal asymmetry observed in the decay of the neutral K meson (Cronin 1981). Many are quite sceptical about the possibility that the origin of macroscopic irreversibility is

determined by such a phenomenon, which is confined to the realm of high energy physics.

(b) Irreversibility arises from the fact that real systems are open, i.e. that they interact with their environment. Consequently, they are appropriately described by stochastic processes, which have an irreversible nature.

(c) The basic ingredients of irreversibility are deterministic chaos and uncertainties in the initial data. In chaotic systems, uncertainties are amplified very rapidly, so that correlations with the initial state are readily lost, and returning to the initial situation is impossible.

(d) The origin of irreversibility resides in the measurement procedure which, at the quantum mechanical level, leads to the collapse of the wave function.

(e) The key to irreversibility is the very large number of particles that are involved.

There is also another point of view, which we could classify as pragmatic and diametrically opposed to standpoint (a). It implies that the mechanisms underlying macroscopic irreversibility should be ignored, because irreversibility is a macroscopic fact which requires no deeper or more fundamental explanation. Irreversibility should then merely be formalised within the realm of thermodynamics (Lieb and Yngvason 2000).

It is impossible to present in a few pages a detailed discussion of all these different opinions. Nonetheless, some remarks are in order. Standpoint (a) is the most "fundamentalist": the problem of irreversibility does not exist since it is inherent in the microscopic laws at the level of elementary particles. Standpoint (b) refuses to address the problem: individual realisations of a stochastic system can be reversible. Standpoint (c) is common to a very broad spectrum of viewpoints that range from Prigogine's position to subjectivist views. In particular, Prigogine's positions are difficult to tackle, because they changed in time. For example in (Nicolis and Prigogine 1977), it is argued that irreversibility is not in contradiction with the laws of dynamics, and that it follows from these laws in sufficiently complex dynamics. Some years later, this position is abandoned, as demonstrated by Prigogine and Stengers (1979):

> we must accept the evidence of facts [...] irreversibility cannot be deduced from the dynamics, it must therefore be introduced by hand

Prigogine also placed great emphasis on the technical and conceptual role of deterministic chaos in the description of nature:

> the notion of chaos leads us to rethink the notion of "law of nature" [...] For chaotic systems, trajectories are eliminated from the probabilistic description [...] The statistical description is irreducible (Prigogine 1994).

In response to an article by Lebowitz (1993), who supported the "orthodox" view inspired by Boltzmann, a disciple of Prigogine argues that irreversibility is not to be found at the level of trajectories, therefore the arrow of time is not due to any phenomenological approximation but is an intrinsic property of a class of unstable dynamical systems (Barnum et al. 1994).

By contrast, we shall argue that the role of chaos is not particularly important.

Standpoint (d) is, in our opinion, a downward spiral that shifts the problem to even more controversial grounds than those of irreversibility. Our personal preference goes to standpoint (e), which is in line with Boltzmann's original idea.

4.2 Irreversibility and Emergence

In the framework of the Boltzmann equation, irreversibility appears in terms of a probability distribution defined on the μ-space, which is the space of positions and velocities of a single particle. Since probabilities have often been considered as measures of our ignorance on a given phenomenon, or have been introduced to circumvent the impossibility of exact descriptions, one may be led to the conclusion that the irreversible behaviour of macroscopic bodies is a subjective property related to our imprecise knowledge of a system's microstate, and not a genuine physical property. We regard this as an erroneous conclusion; irreversibility is a matter of fact and the probability density introduced by Boltzmann is, in reality, the mathematical expression of a physical property of the system of interest.

There has been a misleading and often heated debate which has generated confusion over the relevance of chaos in science. It is therefore appropriate to discuss briefly the relevance of chaos to the problem of irreversibility (Bricmont 1995). It is nowadays clearly established that even systems with few degrees of freedom may behave, in many respects, like stochastic processes, but this should be properly eludicated. If we consider a chaotic system and an ensemble of initial conditions,[2] then under fairly general conditions, the probability density evolves rapidly towards an invariant distribution (i.e. a probability distribution which does not change with time): say $\rho_t(\mathbf{x}) \to \rho_{inv}(\mathbf{x})$ when $t \to \infty$. However, despite some folklore, such a result has no relevance for the problem of irreversibility. For instance, the time-reversed evolution of a chaotic system with few degrees of freedom, such as that concerning the baker map often considered by Prigogine, does not reveal any special features: it does not at all look more unnatural than its forward evolution. Unnatural behaviour becomes evident only if one rewinds the motion of a set of many trajectories at once or, equivalently, if one rewinds the evolution of a probability density $\rho_t(\mathbf{x})$. But this has nothing to do with the irreversibility of real systems, which is experienced by *single systems* and corresponds to single trajectories. A given gas, which is represented by a single point $\mathbf{x}$ in the phase space, irreversibly evolves till it is uniformly distributed in its container. Nevertheless, the trajectory drawn by $\mathbf{x}$ in phase space never spreads to become a cloud that uniformly covers the phase space. Moreover, the uniform distribution of the gas in the container occurs in times so short that the trajectory of $\mathbf{x}$ may only explore a minuscule fraction of the phase space.

Therefore, the fact that the Boltzmann equation describes the evolution of a probability distribution should not be misunderstood: $f(\mathbf{x}, \mathbf{v}, t)$ does not represent

[2] Which is described by a probability density $\rho_0(\mathbf{x})$ localised around a certain point $\mathbf{x}^c$: this can indeed reflect our imprecise knowledge of the initial state of the system.

an ensemble of systems; it is, instead, the distribution of the density of particles of a single system made of very many particles.

Deterministic chaos nevertheless plays a role, in light of which Boltzmann's ideas can be proven to be substantially correct. From the sensitive dependence on initial conditions, which characterises chaotic dynamics (Bricmont 1995; Cencini et al. 2009), one can understand how the reversibility paradox is practically irrelevant, even for systems of moderately many degrees of freedom. Indeed, performing an exact reversal of the velocities of all particles of the system, the subsequent evolution makes H increase for at least as long as it had decreased. However, in the case in which the reversal of velocities is affected by any small deviation from the exact reversal, the exponential growth of perturbations makes the trajectory rapidly separate from the truly reversed one. The consequence is that H grows for only a very short time, after which it decreases again and, indeed, even the exactly reversed trajectory will eventually lead to the decrease of H. This has been clearly illustrated by various numerical experiments. In particular, Fig. 4.2 shows a series of time inversions with a small error, in a model system consisting of disks moving in two dimensions and interacting only through elastic collisions. The rapid return to decreasing H is favoured by an increase in the number of particles; therefore the H-theorem is very robust in macroscopic systems, which are made of very many particles.

By contrast, chaos plays no role in the recurrence paradox. Poincaré's recurrence time T_R grows exponentially with the number of degrees of freedom N, independently of the presence of chaos: $T_R \sim \tau C^N$ where τ is a characteristic time and $C > 1$. Therefore, even with moderate values of N, e.g. $N \sim 10 - 100$, T_R is enormous and, actually, longer than any experimentally relevant time. Boltzmann himself, as recalled, reached this conclusions on the grounds of probabilistic considerations.

Why, then, does the evolution of $f(\mathbf{x}, \mathbf{v}, t)$ described by the Boltzmann equation result in a monotonic decrease of H? Which feature of the Boltzmann equation yields this physically relevant, but mechanically unwarrantable result? A detailed analysis (Cercignani 2006) shows that the derivation of the Boltzmann equation is based on the hypothesis of molecular chaos, or the *Stosszahlansatz* and on the $N \to \infty$ limit. In particular, the *Stosszahlansatz* concerns the collisions between two molecules and states that particles are uncorrelated *before* the collision. This assumption, which allows after-collision states to be described in terms of pre-collision states, rather than the latter in terms of the former, cannot be justified by dynamic considerations and is, indeed, of a probabilistic nature.[3] As a consequence, the Boltzmann equation and the H-theorem cannot be considered valid in absolute terms, but only in the statistical sense according to which Eq. (4.3) provides the most likely behaviour. In any finite system, $H(t)$ does not decrease monotonically (unlike the prediction of the H-theorem) but fluctuates irregularly around the monotonic behaviour implied by the H-theorem, as shown in Fig. 4.3. Even the monotonic trend does not last forever, but it persists well beyond any physically relevant time. At the time t_i, the molecular chaos hypothesis holds if $H(t_i)$ is a local maximum (Fig. 4.4). Hence Boltzmann's

[3] As a matter of fact, the opposite assumption is equally viable dynamically, and leads to the so-called anti-Boltzmann equation.

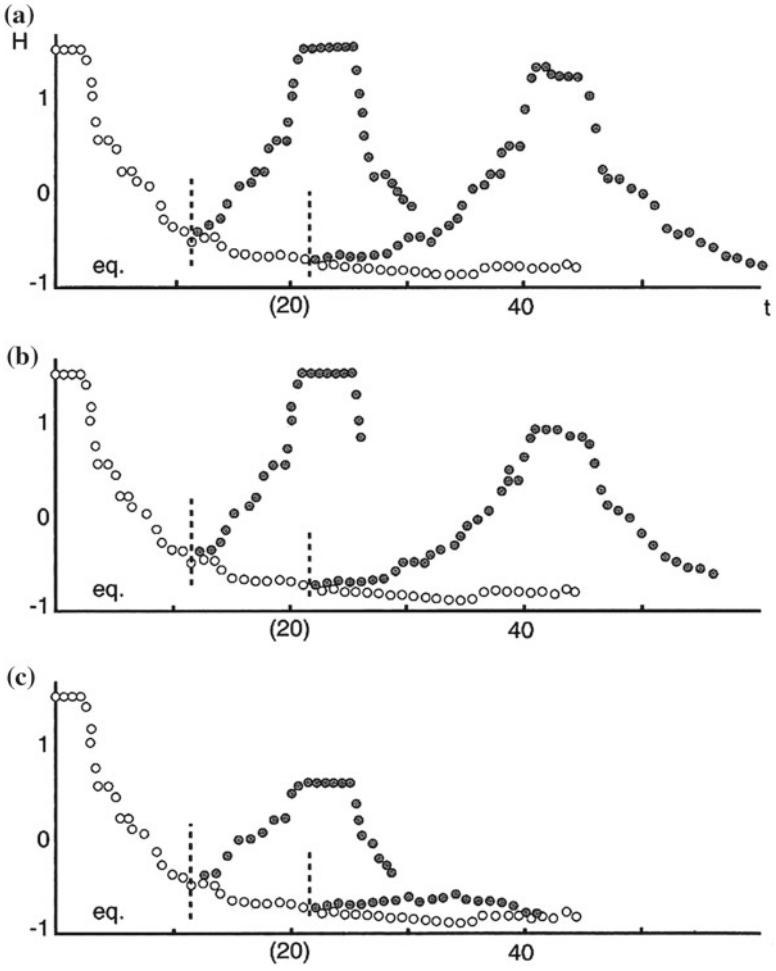

Fig. 4.2 $H(t)$ versus t in a system of 100 elastic disks, with periodic boundary conditions. The *open circles* indicate the forward evolution, the *black circles* show the evolution of the configurations with a time reversal after 50 and 100 collisions. In the time reversal operation there is an error of 10^{-8} in (**a**), 10^{-5} in (**b**) and 10^{-2} in (**c**). This picture shows the computer experiment in (Orban and Bellemans 1967)

assumption, in a sense, requires the functional $H(t)$ to be made only of peaks. The fact is that the peaks in the fluctuating curve approach each other, and the smooth monotonic curve, as N grows. In the $N \to \infty$ limit, T_R also tends to infinity and the real behaviour of H tends to coincide at all finite times with its approximation given by the H-theorem.

Simple probabilistic models illustrate the validity of Boltzmann's hypothesis, despite the formal correctness of the recurrence and reversibility objections. In

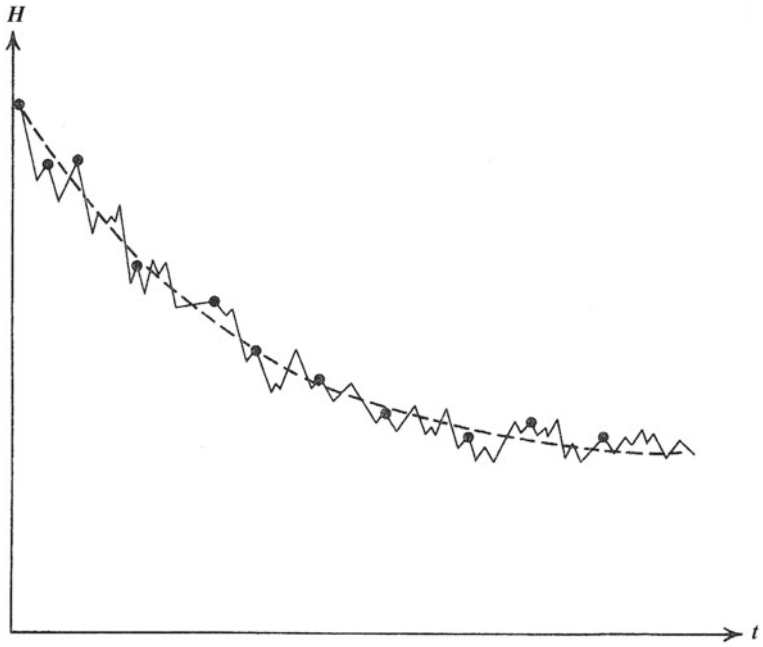

Fig. 4.3 The typical behaviour of $H(t)$ versus t in a system far from thermodynamic equilibrium. The *dashed line* gives the prediction of the H-theorem. The *black dots* are the points at which the hypothesis of molecular chaos is verified

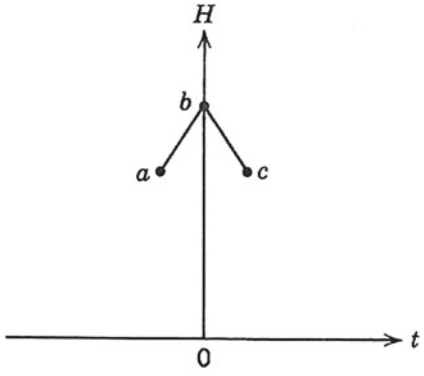

Fig. 4.4 $H(t)$ has a local peak when molecular chaos occurs

particular, consider the model of fleas on two dogs, introduced by Ehrenfest and Ehrenfest (2002), which is quite enlightening, despite its simplicity. Let $2N$ be the number of fleas hosted by two dogs, A and B. Every second, a flea, chosen at random, jumps from one dog to the other. For instance, the fleas could be numbered and every second a number could be drawn from an urn containing $2N$ numbered balls. After the flea has jumped, its ball is placed back in the urn. Let n_t be the number of fleas on dog A at time t, and let the initial condition n_0 be close to $2N$, so that almost

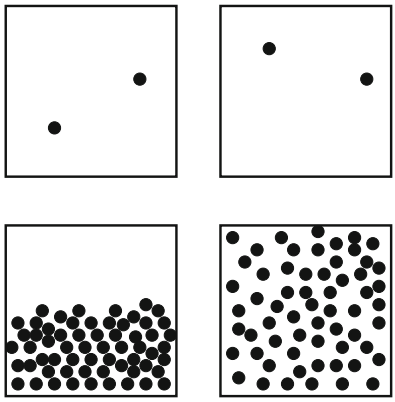

Fig. 4.5 Two successive snapshots of a cell of volume δV containing a few particles (*top row*), and of the same cell with many particles (*bottom row*). If the number of particles is small, there is no notion of relaxation to equilibrium in time. Conversely, if the number of particles is large and they interact, they tend to uniformly occupy δV, increasing their "*disorder*"

all fleas are initially on dog A. One expects n_t to decreases in time, until it settles around N, with fluctuations of the order of $\sqrt{N}$.

It is interesting and easy to show (Kac 1957) that configurations with

$$n_{t-1} = m - 1, \quad n_t = m, \quad n_{t+1} = m - 1, \quad 0 \leq m \leq N$$

which verify the molecular chaos hypothesis (see Fig. 4.5), dominate over all others when $N >> 1$ and the system is very far from equilibrium, i.e. for $|m - N| \gg \sqrt{N}$. This provides an interpretation and an explanation of Boltzmann's assumption, that could look paradoxical from certain points of view. In practice, almost all points of the curve H are local maxima, if the number of fleas is very large.

This result for the dog-flea model suggests that the Boltzmann equation (hence the H-theorem) may be proven to hold in some appropriate limit.

Starting with the fundamental work of Grad of 1948, Boltzmann's deep physical intuition was eventually formulated rigorously and proven to be valid. Among the many scientists who have contributed significantly to this result, one may recall Illner, Lanford, Shinbrot, DiPerna, Lions, Pulvirenti and Cercignani. The essence of their work may be summarised as follows (Cercignani et al. 1994).

Let the molecules be represented by hard spheres of diameter σ. Let their number per unit volume, N say, tend to infinity while σ tends 0, in such a way that $N\sigma^2 \to$ constant. Then, the probability distribution $f(\mathbf{x}, \mathbf{v}, t)$ obeys the Boltzmann equation and the H-theorem holds, if the initial condition fulfils the hypothesis of molecular chaos. In addition, the $N \to \infty$ limit implies that molecular chaos is verified with probability one. The agreement of this rigorous result with Boltzmann's original views is impressive (Boltzmann 1974):

> Since in the differential equations of mechanics themselves there is absolutely nothing analogous to the second law of thermodynamics the latter can be mechanically represented only by means of assumptions regarding initial conditions.

In our opinion, Boltzmann's stroke of genius is not the reduction of thermodynamics to statistical mechanics, but having understood that irreversibility cannot be deduced solely from the laws of mechanics. He deduced the singular nature of the limit[4] leading to the emergence of irreversibility, and pointed out the two fundamental essential ingredients:

(a) the large number of particles (atoms or molecules), hence the great disparity between microscopic and macroscopic scales;
(b) the appropriate initial conditions (those which verify molecular chaos). Today we can identify a third element, somehow implicit in b):
(c) a probabilistic fact: not all microscopic states evolve in the irreversible fashion predicted by the macroscopic theories, but only the vast majority. In macroscopic systems containing an enormous number of particles, "vast majority" means practically all microstates, hence irreversibility is a universal fact.

In our opinion, irreversibility is closely related to the coexistence of different processes occurring on different space and time scales within a given phenomenon. Therefore, different levels of description can be adopted to investigate that phenomenon. As noted by Duhem (1903): observing the single molecules of a fluid, one would observe a tumultuous agitation, both in the case of a fluid in a laminar regime, which looks very regular at the macroscopic level, and in a turbulent fluid, which is characterised instead by an irregular behaviour also at the macroscopic level. If one spills perfume in one corner of a room, the perfume molecules will rapidly spread all over the room. Now let us film the molecules and then look at the movie while it rewinds. An unnatural phenomenon is observed: all the molecules which had initially spread to reach all parts of the room reunite in the original corner. On the contrary, the motion of a single molecule does not reveal any abnormal behaviour. Similarly, nothing strange is apparent in the backward motion of the molecules contained in a small volume. The impression of unnatural behaviour strikes us only when a large number of molecules or a macroscopic volume is observed.

Obviously, the state of the system does not depend on our viewpoint: it is objectively what it is. But in studying its properties, we may choose the thermodynamic perspective, which emerges at the scale of the collective or macroscopic quantities, or we may choose to focus on a few microscopic constituents of the system. The result of the observations will of course be different. Analogously, looking closely at the pores in the marble of a cathedral's facade, one gets a rather different perspective on the cathedral than from observing the facade at some distance. The properties of the cathedral are objectively what they are, independent of the

[4] At any finite N, the evolution is reversible. Therefore, from a conceptual, qualitative point of view, the case with $0 < \frac{1}{N} \ll 1$ differs from that with $\frac{1}{N} = 0$.

observer's point of view, but apparently contrasting features[5] coexist within a unique reality.

From the point of view of the objective state of a macroscopic system, the emergence of different levels of description requires that the number of molecules N be really very large, and that molecules interact with each other, although the fine details of these interactions are not so important. Under these conditions, relaxation processes, and therefore irreversibility (as illustrated by Fig. 4.5) may take place.

This fact appears not to be understood by everyone. For example, Prigogine and his school have a very different opinion, as revealed by the following statement:

> Irreversibility is either true on all levels or on none: it cannot emerge as if out of nothing, on going from one level to another (Prigogine and Stengers 1984).

Such a point of view is an explicit type of reductionism, since it clearly denies that new phenomena may emerge, when switching from one level of observation to another. In contrast, we see irreversibility emerge from the $1/N \to 0$ singular limit, as posited by Boltzmann and later demonstrated in a rigorous way. Irreversibility does emerge in the passage from microscopic to macroscopic descriptions (Primas 1998), like the unique character of a cathedral emerges when moving the observation point from very close to a proper distance. This kind of emergence is quite similar to the emergence of chemistry from quantum mechanics, which will be discussed in Chap. 6.

4.3 From Microscopic to Macroscopic Equations

4.3.1 Continuous Media and Thermodynamics

Macroscopic objects appear continuous to our senses, even though they are made of molecules, which are small separate lumps of matter. Macroscopic objects are therefore conveniently described in terms of *fields*, i.e. of functions ϕ which associate a real number $\phi(\mathbf{x})$ to all points $\mathbf{x}$ in the space V occupied by the object of interest. The most common fields are the temperature T, the mass density ρ, the pressure p and the velocity field $\mathbf{u}$. These fields emerge from a coarse-graining in real space, which identifies the value of the given field at point $\mathbf{x}$ with an average computed over a small volume δV containing $\mathbf{x}$. The volume δV has to be small enough to be perceived as one point by our measuring tools, but large enough to contain very many molecules, cf. Fig. 4.6. Averaging over cells that are small compared to the resolution of our measuring instruments, but that contain enough molecules for the granularity of matter not to be apparent, produces quantities which are meaningful at the macroscopic level. If the state of the system allows us to replace the detailed

[5] For instance, from the point of view of the pores in the marble, all cathedrals look the same, and yet they all look different from a certain distance.

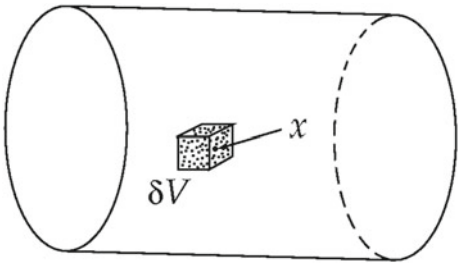

Fig. 4.6 A point **x** in the volument element δV, containing $N_{\delta V}$ molecules. If δV is smaller and $N_{\delta V}$ larger, the granularity of matter is even less evident, and the object can be seen as continuous. As small and large are relative concepts, they must be defined relative to the resolution of our observations

microscopic evolution with such space-time averages, hence to introduce the thermodynamic fields, we have a state of *local thermodynamic equilibrium*. This, in turn, depends on the dynamics of the components of the object of interest, and on the resolution of the observations. In particular, one may only speak of local thermodynamic equilibrium if the following conditions are verified:

(a) there are three vastly separated temporal and spacial scales: the microscopic, the mesoscopic and the macroscopic scales:

$$\ell \ll \delta L \ll L; \quad \tau \ll \delta t \ll t; \tag{4.5}$$

(b) the very many atoms (or molecules) within a mesoscopic cube of side δL reach a uniform distribution in a time shorter than the mesoscopic time δt;
(c) observations concern space and time scales much larger than δL and δt.

In kinetic theory, ℓ is the mean free distance travelled by a molecule between two collisions with other molecules and τ is the corresponding mean free time, while L and t are the macroscopic scales. If micro- and macro-scales are sufficiently separated, there is room for an intermediate scale, vastly separated from the other two. The concepts of thermodynamics emerge in phenomena observed on the macroscopic scale that satisfy conditions (a), (b) and (c).

There are, of course, macroscopic events that violate the condition of local thermodynamic equilibrium. For instance, shock waves evolve so rapidly and produce such strong inhomogeneities that relaxation to uniformity takes longer than mesosocopic times. There are also phenomena, such as the protein-folding and many other e.g. biophysical processes, which are characterised by too many different scales, which in turn cannot be sufficiently widely separated from each other.

4.3.1.1 Hydrodynamics

In the most common situations of our daily lives, fluids can be treated as continuous systems in local thermodynamic equilibrium, even if they flow from one side of a pipe to the other, or are pushed around the globe by pressure and temperature gradients, to form the gigantic swirls of tornadoes and hurricanes. Under general assumptions,

one obtains descriptions of their evolution, like the *Navier-Stokes equation* for the velocity field $\mathbf{u}$:

$$\nabla \cdot \mathbf{u} = 0 \tag{4.6}$$

$$\partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\frac{1}{\rho}\nabla p + \nu \Delta \mathbf{u} + \mathbf{f}\,, \tag{4.7}$$

where $\mathbf{f}$ is an external force (per unit volume), ρ the (constant) density and ν the viscosity of the liquid.

The Navier-Stokes equation is universally recognised as exact in practice, for the description of fluids in the most diverse situations. Could one obtain a similarly useful description in terms of the motion of the molecules which constitute a fluid?

As a matter of fact, the Boltzmann equation, which we have discussed from the historical and conceptual standpoint, is important also for the derivation of macroscopic hydrodynamic laws, from a microscopic perspective.

Let us formulate the problem. The equations for the evolution of fluids had been originally obtained (in the inviscid limit) by Euler, and then in more complete form by Navier and Stokes in the nineteenth century. These authors used a continuum description whose main ingredients were the conservation laws for mass and momentum. These partial differential equations had been derived with no reference to the underlying atomic dynamics,[6] using just Newton's laws and the assumption of smoothness for the hydrodynamic field.

4.3.2 In Boltzmann's Footsteps

If one believes in the existence of atoms, and assumes that these evolve according to the laws of mechanics, as Boltzmann did, it is natural to try to derive the macroscopic equations of evolution from the microscopic ones. We shall see how the success of this ambitious program has been based on the fundamental roles of the separation of microscopic and macroscopic length and time scales, and of the huge number of particles. In the passage from the atomic to the hydrodynamic scale one has to pass through different levels of description:

I. Adopting a statistical approach, the microscopic level is described by the Liouville equation, which governs the evolution of the probability density in the phase space Γ (whose dimension is $6N$). Such a description includes all the statistical information available on the system and is far too detailed, for almost any purpose.
II. The Boltzmann equation concerns a microscopic level, whose reference space is the μ-space of the single particle description. The dimensionality of the μ-space is 6.

[6] At that time the existence of atoms was still an open problem.

III. At the macroscopic level, the hydrodynamic description is based on continuous fields (of velocity, pressure, temperature etc.) which evolve according to partial differential equations, such as the Navier-Stokes or Boussinesq equations.

Some problems are best tackled in μ-space, at a scale larger than that used in the derivation of the Boltzmann equation. This scale is still part of the microscopic (or mesoscopic) level.[7]

In the previous sections, we have already discussed the basic aspects of the transition from level (I) to level (II), demonstrating that a rigorous derivation of the Boltzmann equation from the microscopic dynamics requires N to grow without bounds, while the mass m and the size σ of each particle tends to zero. More precisely, the Boltzmann equation holds under the following conditions:

$$N \to \infty, \quad m \to 0, \quad \sigma \to 0,$$

in such a way that

$$Nm \to M = const., \quad N\sigma^3 \to 0, \quad N\sigma^2 \to \ell_m^{-1}.$$

The physical meaning of these conditions, which amount to the so-called Boltzmann-Grad limit, is rather transparent: M is the total mass of the gas, $N\sigma^3$ is the total volume occupied by the molecules, which must be very small, and $1/N\sigma^2$ is the mean free path ℓ_m.

4.3.2.1 The Hilbert Contraction: A Real Reductionist Procedure?

It is well established that the path (I) $\to$ (II) $\to$ (III) can be followed to obtain the hydrodynamic equations, from a bottom up approach, Cercignani et al. (1994), Gorban et al. (2004). At first glance, the success of such a program may appear as a triumph of the reductionist point of view. In the following we show that, unfortunately, the steps from (I) to (III) are not straightforward, and hence that the reduction of hydrodynamics to microscopic dynamics is not so simple. Actually, the success of the derivation of the Navier-Stokes equation from kinetic theory can only be justified after the fact. This is revealed through the so called "Hilbert contraction" which plays a basic role in the treatment of the Boltzmann equation.[8]

Indeed, this derivation requires a manipulation of the Boltzmann equation, which is a rather complicated problem. Hilbert's clever idea is based on profound physical

[7] For instance the statistical features of a colloidal particle, which is much bigger than the single molecules, is described by a Fokker-Planck equation which takes into account a possible external potential and the interactions between the colloidal particle and the molecules.

[8] Hilbert's ideas had been further developed by Chapman and Enskog who introduced an efficient, although not fully rigorous method of computing transport coefficients (such as the viscosity) in terms of the microscopic interactions among the molecules(Chapman and Cowling 1970; Cercignani et al. 1994).

intuition, and may be summarised as follows. Hydrodynamics concerns time and length scales which are much larger than those of the molecular level; therefore describing hydrodynamic properties in terms of the one-body density distribution $f(\mathbf{x}, \mathbf{v}, t)$ of the Boltzmann equation is hopeless, since the amount of information provided by $f(\mathbf{x}, \mathbf{v}, t)$ is too large. The separation of scales, however, may justify the mathematical procedure known as multiscale analysis, whose main ingredients are the following:

(a) The gas reaches local equilibrium in a short time: the time taken by a particle to undergo a few collisions. Therefore, $f(\mathbf{x}, \mathbf{v}, t)$ can be approximated by the Maxwell-Boltzmann distribution $f_{MB}(\mathbf{x}, \mathbf{v}, t)$, with a space-dependent temperature.

(b) $f(\mathbf{x}, \mathbf{v}, t)$ is expressed by a perturbation expansion, in which the Knudsen number (the ratio between the mean free path ℓ_m and the typical macroscopic length L) is the small parameter ε:

$$f(\mathbf{x}, \mathbf{v}, t) = f_{MB}\Big(1 + \varepsilon f^{(1)} + \varepsilon^2 f^{(2)} + \cdots\Big).$$

(c) Because of the assumed local equilibrium, the distribution function $f(\mathbf{x}, \mathbf{v}, t)$ depends on $\mathbf{x}$, $\mathbf{v}$ and t only through the macroscopic fields, i.e. through the density ρ, the hydrodynamic velocity $\mathbf{u}$ and the temperature T. One may then write:

$$f(\mathbf{x}, \mathbf{v}, t) = F\Big(\rho(\mathbf{x}, t), \mathbf{u}(\mathbf{x}, t), T(\mathbf{x}, t)\Big),$$

where

$$\rho(\mathbf{x}, t) = \int f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v}, \quad \mathbf{u}(\mathbf{x}, t) = \frac{1}{\rho(\mathbf{x}, t)} \int \mathbf{v} f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v},$$
$$T(\mathbf{x}, t) = \frac{1}{\rho(\mathbf{x}, t)} \int \frac{1}{2}\Big(\mathbf{v} - \mathbf{u}(\mathbf{x}, t)\Big)^2 f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v}.$$

This contraction mechanism drastically reduces the number of degrees of freedom involved in the problem.

(d) The multiscale method leads to the Navier-Stokes equations and provides an explicit way of computing the transport coefficient, such as the viscosity of the fluid (Gorban et al. 2004).

The two conceptual aspects of this successful procedure are that:

1. the time scales of the microscopic and macroscopic realms are assumed to be widely separated (therefore, local equilibrium is realised in a time which is very short from the macroscopic point of view): and
2. the one-body distribution function $f(\mathbf{x}, \mathbf{v}, t)$ is assumed to depend only on the macroscopic fields ρ, $\mathbf{u}$ and T. Note that this is an assumption on the existence of the macroscopic behaviour; it is not proven from the microscopic point of view, except in some simplified models (Saint-Raymond 2009; Villani 2002).

This derivation of hydrodynamics follows a sort of bootstrap procedure.

4.3.3 The Emergence of a Dissipative Phenomenon: Friction in a Reversible World

In Chap. 3, we discussed the motion of a colloidal particle in a fluid. In particular, Eq. (3.10) contains the Stokes force due to friction between the particle and the fluid, $\mathbf{F_S} = -\mu\mathbf{V}$, where the friction coefficient μ is determined by macroscopic quantities: $\mu = 6\pi\eta R$ (R is the radius of the colloidal particle and η is the viscosity of the fluid). Such a force, obtained from a phenomenological theory (hydrodynamics), has a non-conservative character and induces a decrease in the particle energy. It is therefore natural to ask whether this friction can be obtained from the conservative and time-reversible dynamics by which we describe the microscopic constituents of the fluid.

In the case of dilute gases, one may follow an approach due to Smoluchowski, which is rather simple and leads to an explicit expression for μ (Cecconi et al. 2007). Consider a two-dimensional square domain of side L, and impose periodic boundary conditions on the dynamics of $N \gg 1$ particles which constitute a gas contained in this domain. Let m be the mass and r the radius of each fluid particle, while M and R are the mass and radius, respectively, of a colloidal particle.

If the velocity of the colloidal particle does not vanish, $\mathbf{V} \neq 0$, the collisions with the fluid particles are less frequent in the direction of the motion than in the opposite direction. For instance, if $\mathbf{V} = (v, 0)$ with $v > 0$ the collisions due to the particles coming from the right are more frequent than those due to particles coming from the left. Such an imbalance gives rise to a force opposing the motion, which can be explicitly computed under the following assumptions:

1. the gas is dilute, i.e. $r^2 \ll L^2/N^2$;
2. the velocities of the particles are distributed according to the Maxwell-Boltzmann distribution at temperature T;
3. there is a very wide separation between the scales of the colloidal particle and those of the gas particles, i.e. $M \gg m$ and $R \gg r$;
4. the speed of the colloidal particle is small compared with the speed $v_T = \sqrt{kT/M}$ of thermal agitation;

The result is[9]:

$$\mathrm{M}\frac{d\mathbf{V}}{dt} = -\mu\mathbf{V}, \tag{4.8}$$

where $\mu = 2\rho R\sqrt{2\pi mkT}$, and $\rho = N/L^2$ is the number density. Note that the above result, which holds for gases, differs from that expressed by Stokes' law for liquids, see Sect. 3.3. Equation (4.8) now yields $\tau = M/\mu$, for the relaxation time τ.

[9] We write just the systematic deterministic part. The complete Langevin equation also contains a noise term (as in Sect. 3.3).

The irreversible nature of the friction law Eq. (4.8) appears to contrast with the reversible character of the dynamics of the system made up of the N fluid particles plus the colloidal particle. However, we should note that Eq. (4.8) refers to the behaviour of a single particle (that is, a projection of the global dynamics). The global dynamics does indeed obey Poincaré's recurrence theorem (Poincaré 1890), but the recurrence time is extremely long: $T_R \sim \tau_0 C^N$, where τ_0 is a characteristic time of the system and $C > 1$. Therefore, one finds $\tau \ll T_R$. In other words, the true reversible nature of the dynamics may be revealed only after an enormous time; a time much larger, for instance, than the age of the universe. On the scale of our observations the colloidal particle follows precisely the dissipative Eq. (4.8).

4.3.4 Hydrodynamics in an Artificial World: Cellular Automata

Let us address the key points that determine the emergence of the macroscopic level of description from the microscopic level. In particular, let us ask whether the details of the microscopic dynamics, such as the potential $U(r)$ of the molecular interactions, plays any relevant role in the properties of the Navier-Stokes equations. First, we note that the form of these equations does not explicitly depend on $U(r)$. Indeed, $U(r)$ is only relevant for the numerical value of the transport coefficients, which appears as a phenomenological constant in the Navier-Stokes equations. These constants must be experimentally determined or must be computed e.g. from kinetic theory. Therefore, one may hope that the macroscopic description of a fluid remains the same, even if the details of the microscopic dynamics change.

This issue is not only conceptually interesting, but has practical relevance, because it concerns the numerical approach to the hydrodynamic equations. The fact is that the Navier-Stokes equations are quite difficult to treat, because of their nonlinear terms, and only in a few special cases they can be solved analytically. Therefore, a numerical approach is unavoidable. One common way to tackle the problem is to discretise the equations in both space and time, so that they can be loaded into a computer (cf. next section for an explicit example).

Let us briefly mention an interesting alternative approach, in which the deterministic microscopic dynamics of the fluid molecules (which obey Newton's equations) is replaced by a discrete probabilistic model defined on a lattice and known as a cellular automaton (Frisch et al. 1986; Rothman and Zaleski 2004). For instance, one may consider a two-dimensional hexagonal lattice. The particles of this artificial world have only a discrete set of velocities and hop from one site of the lattice to a nearest-neighbour site. If two or more particles are at the same site, a collision changes their velocities. The rules governing the collisions are devised to satisfy the minimal physical requirements, namely the conservation laws. Such a simple model possesses remarkable properties. In particular, performing an average over a spatial region much larger than the lattice spacing, one obtains a kind of macro-dynamics which represents the hydrodynamics of that artificial world. Remarkably, by adopting various collision rules, the automata provide an efficient simulation technique,

suitable for describing a wide range of problems ranging from phase separation to interfaces and multiphase flows.[10]

Although the cellular-automata description for fluids presupposes an artificial world, from a technical point of view this approach works, because one may repeat, step by step, the bottom-up procedure that leads from the microscopic dynamics (in this case, the probabilistic collision rules) to the one-body description (the Boltzmann equation). The Hilbert procedure or the Chapman-Enskog approach, applied to the class of cellular automata introduced to simulate fluids, then leads to the correct macroscopic hydrodynamics.

4.4 From Atoms to Cold Fronts: A Random Walk Through Hydrodynamics and Meteorology

Despite being a champion of extreme reductionism, Weinberg admits the distinction between reductionism as a "view of the nature" and reductionism as a "program for scientific research". He states:

> Reductionism may or may not be a good guide for a program of weather forecasting, but it provides the necessary insight that there are not autonomous laws of weather forecasting that are logically independent of the principles of physics. Whether or not it helps the meteorologist to keep it in mind, cold fronts are the way they are because of the properties of air and water vapor and so on which in turn are the way they are because of the principles of chemistry and physics. We don't know the final laws of nature, but we know that they are not expressed in terms of cold fronts or thunderstorms.

An overview of the procedures actually followed in weather forecasting clarifies the importance of effective equations much more than many abstract discussions, for systems with a multiscale character.

The first modern steps in weather forecasting are due to Richardson (1922)[11] who, in his visionary work, introduced many of the ideas on which modern meteorology is based. His approach was, to a certain extent, in line with genuine reductionism, and may be summarised as follows: the atmosphere evolves according to the hydrodynamic equations for the velocity u, the density ρ, the pressure p and the temperature T, and according to thermodynamics, which links ρ, T and p together. Therefore, future weather can be predicted, in principle at least, by solving the proper par-

[10] In same cases, as in the presence of complicated spatial geometry, the cellular automata approach has practical advantages over the usual standard numerical methods.

[11] During the first world war, when he was an ambulance driver on the French front, Richardson wrote the first draft of his book on numerical weather forecasting. His manuscript, lost during the battle of Champagne (April 1917), was fortunately recovered under a heap of coal, several months later. Besides meteorology, Richardson wrote seminal works to numerical analysis, turbulence, modelling in psychology and to the discovery of some aspects of fractals. The first to pose the question about the length of the coast of Britain was not Mandelbrot, as commonly believed, but Richardson.

tial differential equations, with initial conditions given by the present state of the atmosphere.

Of course these equations cannot be solved with pen and paper: numerical integration is mandatory and Richardson himself worked hard in that direction. In his pre-computer era, computations had to be done by hand in any case, except for minimal help from rudimentary computing machines. After very long, patient and painful work, Richardson obtained a rather limited result. He then correctly observed that *the calculation is complicated because the atmosphere is complicated*. Nevertheless, he remained moderately optimistic about weather forecasts, thinking that:

> perhaps some day in the dim future it will be possible to advance the computations faster than the weather advances. [...] But that is a dream.

The idea was correct, but for its practical implementation another ingredient had to be introduced. Indeed, the numerical integration of the hydrodynamic equations remains a hard task, even for modern computers. The reason is that so-called numerical instabilities force simulations to be performed with very small integration steps Δt.[12] This negative result is not merely a technical issue: it is rooted in the physics of fluids and, precisely, in the presence of phenomena such as gravity and sound waves, whose characteristic times are very short compared to those of the atmospheric phenomena of interest. Therefore, brute force integration of hydrodynamic and thermodynamic equations is doomed to failure, as anticipated by expert meteorologists in the early days of numerical forecasts (Haurwitz 1941). Significant progress had to wait for an informed distinction between relevant and irrelevant aspects of the meteorologic phenomena. For instance, processes occurring on the scale of seconds or of centimetres are presumably of scarce interest for weather forecasting, and can be neglected. Then, if certain features of the phenomenon can be neglected without affecting the result of interest, the simulations are considerably simplified and shortened.

---

[12] In the numerical integration of partial differential equations, one has to discretise space, with a grid of given size Δx, and time, with an integration step Δt. The numerical results then converge to the solution of the partial differential equation only if the ratio $\Delta t/\Delta x$ is smaller than a number C, which depends on the equation and, typically, also on the initial conditions. This bound is called the "Courant condition" (Courant et al. 1928; Press et al. 1986). As an example, consider the following partial differential equation:

$$\frac{\partial u(x,t)}{\partial t} = -v\frac{\partial u(x,t)}{\partial x}$$

where v is a constant and $-\infty < x < \infty$. A simple finite-difference approximation of this equation yields

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = -v\frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x}$$

where u_j^n is the value taken by u in the point $j\Delta x$ at time $n\Delta t$: $u_j^n = u(j\Delta x, n\Delta t)$. In this case the Courant condition for the stability of the algorithm reads

$$\frac{\Delta t}{\Delta x} \leq C = \frac{1}{|v|}.$$

Here, C does not depend on the initial condition, because the equation is linear.

This approach was pursued by Charney and his colleagues, in the 1940–50s, through the *Meteorological Project*.[13] They noted that the set of equations originally proposed by Richardson contained too much detail, including meteorologically insignificant high-frequency waves, and that an effective equation which filtered out such irrelevant variables was to be preferred. Separating the meteorologically significant part of the phenomenon from the insignificant brought a clear practical advantage: the numerical instabilities became less severe, hence relatively large integration time steps Δt could be used, at last making the numerical computations satisfactorily efficient.[14] Even more importantly, the effective equations for the slow dynamics capture the essence of the phenomena of interest, which could otherwise be hidden in too detailed a description, as in the case of the complete set of original equations. As noted by Charney himself, it is remarkable that the forecasts obtained from the first effective model proposed in the Meteorological Project[15] *were as good as they were*, and in any case nobody had anticipated the enormous practical interest they would later receive.[16] The philosophy of the Meteorological Project, adopted in successive studies, was to create a hierarchy of increasing complex models, each successive model built from the analysis of the previous one.

---

[13] The Meteorological Project, developed at the Institute for Advanced Study, in Princeton, involved scientists from different fields, including leading mathematicians such as J. von Neumann, experts of meteorology, engineers and computer programmers. This project began to realise Richardson's dream. Three fundamental issues were tackled by the project: technology, which led to the design of the first modern computer, ENIAC; numerical methods for the integration of partial differential equations; and the introduction of effective equations for meteorology.

[14] A simple example which illustrates how an effective equation for large scale behaviour can be obtained is the following, (Frisch 1995; E and Engquist 2003): consider the diffusion equation in one spatial dimension:

$$\frac{\partial}{\partial t}\theta = \frac{\partial}{\partial x}\left(D(x, x/\varepsilon)\frac{\partial \theta}{\partial x}\right)$$

where the coefficient $D(x, \frac{x}{\varepsilon})$ contains two scales: a scale of $O(\varepsilon)$ and a scale of $O(1)$. For instance, $D(x, y)$ could be periodic in y with period 1. The above system describes physical processes such as heat conduction in a composite material. The aim is to write an effective diffusion equation valid at long times and scales much larger than ε, i.e. an equation of the form:

$$\frac{\partial}{\partial t}\theta = \frac{\partial}{\partial x}\left(D^E(x)\frac{\partial}{\partial x}\theta\right),$$

where $D^E(x)$ must be obtained in terms of $D(x, y)$. The result is simply given by the harmonic average, computed over the variable y and expressed by:

$$D^E(x) = \frac{1}{\left\langle \frac{1}{D(x,y)} \right\rangle_y} = \left[\int_0^1 \frac{dy}{D(x, y)}\right]^{-1}.$$

[15] These equations are called quasigeostrophic; their simplest instance is a barotropic equation, in which the pressure depends only on the horizontal location.

[16] This equation had already been used by Rossby to study atmospheric waves, but before the results of Charney and coworkers, nobody seriously believed that the model could have produced quantitatively accurate predictions.

It is of paramount importance to understand that the models in the hierarchy were not mere approximations of the original set of equations, obtained from a systematic strategy based on fundamental principles. On the contrary, they were obtained from a subtle mixture of hypotheses, theory and observations.

This brief history of weather forecasting, from Richardson to modern times, shows that knowledge of the ultimate laws governing the behaviour of the atmosphere, in its tiniest detail, is uninteresting. Responding to Weinberg, one may then say that the only feasible, useful approach rests on the derivation of effective equations that somehow provide a correct "*mathematical description of cold fronts*".

As suggested by Charney himself in the Meteorological Project, a computer can be used to crunch numbers rapidly, but also as an *inductive machine*, putting physical hypotheses and models to the test. We will return to the role of computers and models in the last chapter.

4.5 Concluding Remarks

Let us make some general comments on the path leading from a reversible microscopic description to irreversible equations for macroscopic phenomena.

First, we want to stress our opinion[17] that Boltzmann substantially understood the essence of the mechanisms underlying the second law of thermodynamics. After the long process of clarification, which started with Boltzmann himself, and culminated with the work of Grad (1949) and the precise formulation of Lanford (1981) and others (Cercignani et al. 1994), it is now clear that the basic technical ingredients are:

(a) to take the $N \to \infty$ limit and to rescale the size of the molecules in an appropriate way (which is the Boltzmann-Grad limit);
(b) to select appropriate initial conditions for the hypothesis of molecular chaos to hold, and to prove that such initial conditions approach probability one in the $N \to \infty$ limit.

With these assumptions, the temporal fluctuations of $H(t)$ can be eliminated and, therefore, the objections raised by Loschmidt and Zermelo overcome. It is thus surprising that a large part of the literature, including popular books (Prigogine and Stengers 1979), still maintains that Boltzmann's explanation of irreversibility must be revised in the light of modern approaches, mainly based on the idea of deterministic chaos.[18]

The relationship between kinetic theory and the Navier-Stokes equations thus appears similar to the "reduction" of Brownian motion in terms of statistical mechanics. As noted in Chap. 3, Brownian motion brings together the macroscopic and the

[17] and surely our point of view is shared by many physicists.

[18] For example, Hoover (1999) claims that *Our exploration of time reversibility from the perspective of computer simulation and chaos has provided us with insights into the breaking of the time symmetry which were not available to Boltzmann or Gibbs [...] Simulations have clarified the formation and significance of time-reversible ergodic multifractal phase-space structures*.

microscopic levels of description of a given phenomenon, through certain assumptions. Similarly, the treatment of the Boltzmann equation, which lives in μ-space, required assumptions on the macroscopic level. Therefore, the path linking the molecular dynamics to our descriptions of the macroscopic world is not straightforward.

The hydrodynamic equations do not depend too sensitively on the details of the microscopic dynamics, which may then be drastically modified, as in the case of probabilistic cellular automata, without modifying the macroscopic behaviour. The really important properties of the microscopic dynamics reduce to conservation laws and symmetries, which are also the ingredients relevant to critical phenomena and their universality classes.

Once established, the hydrodynamic equations should be solved, but their nonlinear character makes analytical studies possible only in a few special cases. Numerical approaches are mandatory in most situations of interest. However, the problem of weather forecasting remains far from trivial, even with the aid of modern computers. The main difficulty is related to the numerical instabilities which arise because of the diversity of scales involved. In particular, the atmosphere is affected by gravity and sound waves, whose characteristic times are very short compared to those of the forecasts.

The approach adopted by meteorologists (and arguably the only feasible approach) is to devise effective equations for the slow dynamics, which capture the features relevant for forecasting, while neglecting the irrelevant ones. This is not done by approximating the original set of equations and following a systematic strategy outlined by fundamental physical principles. On the contrary, the goal is achieved through a subtle combination of hypotheses, theory and observations.

In summary, the steps from the microscopic level of description to the macroscopic one, and from the macroscopic level of description to the effective equations, go through the following stages (where we have highlighted the relevant spaces and equations):

(A) Microscopic level, Γ-space description (Liouville equation)
(B) μ-microscopic level, μ-space description at small scale (Boltzmann equation)
(C) Mesoscopic level at large scale (Fokker-Planck equation for colloidal particles)
(D) Macroscopic level, hydrodynamic description (Navier-Stokes equation, Fourier Law, . . .)
(E) Macroscopic Level, effective description (geostrophic equations)

The logical jumps required to climb from one level of description to another are performed via coarse-graining and/or projection procedures, both of which entail a loss of information. Furthermore, it is to be stressed that these steps are only possible if the scales characterising the different levels are sufficiently far apart. In that case, indeed, the processes of one of the levels of description contain too much detail to be of interest at a coarser level. Like the pores of the marble are irrelevant to appreciate the beauty of a cathedral, or the grains of color of single frames are irrelevant to understand a film, yet cathedrals are made of porous marble and films of frames, so the dynamics of single molecules do not give a picture of irreversibility. When scales

are sufficiently far apart, the jumps from one level of description to a less detailed one take the form of mathematically rather delicate singular limits.

Irreversibility is a striking example of genuine emergence, in a strong sense at epistemic level at least, since it affects very visibly all of us.

References

Barnum, H., Caves, C.M., Fuchs, C., Schack, R., Driebe, D.J., Hoover, W.G., Posch, H., Holian, B.L., Peierls, R., Lebowitz, J.L.: Is Boltzmann entropy time's arrow's archer? Phys. Today **47**, 11 (1994)

Boltzmann, L.: Theoretical Physics and Philosophical Problems: Selected Writings. Springer, New York (1974)

Bricmont, J.: Science of chaos or chaos in science? Ann. N.Y. Acad. Sci. **775**(1), 131 (1995)

Cecconi, F., Cencini, M., Vulpiani, A.: Transport properties of chaotic and non-chaotic many particle systems. J. Stat. Mech. Theory Exp. P12001 (2007)

Cencini, M., Cecconi, F., Vulpiani, A.: Chaos: From Simple Models to Complex Systems. World Scientific, Singapore (2009)

Cercignani, C., Illner, R., Pulvirenti, M.: The Mathematical Theory of Dilute Gases. Springer, New York (1994)

Cercignani, C.: Ludwig Boltzmann: The Man Who Trusted Atoms, vol. Oxford. Oxford University Press, USA (2006)

Chapman, S., Cowling, T.G.: The mathematical theory of non-uniform gases: an account of the kinetic theory of viscosity, thermal conduction and diffusion in gases. Cambridge University Press, Cambridge (1970)

Courant, R., Friedrichs, K., Lewy, H.: Uber die partiellen differenzengleichungen der mathematischen physik. Math. Ann. **100**(1), 32 (1928)

Cronin, J.W.: CP symmetry violation the search for its origin. Rev. Mod. Phys. **53**(3), 373 (1981)

Duhem, P.: L'évolution de la mécanique. Revue générale des sciences pures et appliques **XIV** (1903)

Ehrenfest, P., Ehrenfest, T.: The Conceptual Foundations of the Statistical Approach in Mechanics. Dover Publications, New York (2002)

E, W., Engquist, B.: Multiscale modeling and computation. Not. AMS **50**, 1062 (2003)

Frisch, U., Hasslacher, B., Pomeau, Y.: Lattice-gas automata for the Navier-Stokes equation. Phys. Rev. Lett. **56**(14), 1505 (1986)

Frisch, U.: Turbulence. Cambridge University Press, Cambridge (1995)

Gorban, A.N., Karlin, I.V., Zinovyev, A.Y.: Constructive methods of invariant manifolds for kinetic problems. Phys. Rep. **396**(4), 197 (2004)

Grad, H.: On the kinetic theory of rarefied gases. Commun. Pure Appl. Math. **2**(4), 331 (1949)

Guerra, F.: Reversibilità/irreversibilità. Enciclopedia Einaudi Torino **XI**, 1067 (1980)

Haurwitz, B.: Dynamic Meteorology. McGraw-Hill, New York (1941)

Hoover, W.G.: Time Reversibility, Computer Simulation, and Chaos. World Scientific, Singapore (1999)

Kac, M.: Probability and Related Topics in Physical Sciences. American Mathematical Society, Providence (1957)

Lanford, O.E.: The hard sphere gas in the Boltzmann-Grad limit. Phys. A: Stat. Theoret. Phys. **106**(1–2), 70 (1981)

Lebowitz, J.L.: Boltzmann's entropy and time's arrow. Phys. Today **46**, 32 (1993)

Lieb, E.H., Yngvason, J.: A fresh look at entropy and the second law of thermodynamics. Phys. Today **53** (2000)

Nicolis, G., Prigogine, I.: Self-organization in Nonequilibrium Systems. Wiley, New York (1977)

Orban, J., Bellemans, A.: Velocity-inversion and irreversibility in a dilute gas of hard disks. Phys. Lett. A **24**, 620 (1967)
Poincaré, H.: Sur le problme des trois corps et les quations de la dynamique. Acta Math. **13**, 1 (1890)
Poincaré, H.: Le mecanisme et l'experience. Revue de Métaphysique et de Morale **1**(6), 534 (1893)
Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flanner, B.P.: Numerical Recipes: The Art of Scientific Computing. Cambridge University Press, Cambridge (1986)
Prigogine, I., Stengers, I.: La Nouvelle Alliance. Gallimard, Paris (1979)
Prigogine, I., Stengers, I.: Order out of Chaos. Bantam Books, Toronto (1984)
Prigogine, I.: Les lois du chaos. Flammarion, Paris (1994)
Primas, H.: Emergence in exact natural science. Acta Polytech. Scand. **91**, 83 (1998)
Richardson, L.F.: Weather Prediction by Numerical Methods. Cambridge University Press, Cambridge (1922)
Rothman, D.H., Zaleski, S.: Lattice-gas cellular automata: simple models of complex hydrodynamics. Cambridge University Press, Cambridge (2004)
Saint-Raymond, L.: Hydrodynamic limits of the Boltzmann equation. Springer, New York (2009)
Villani, C.: A review of mathematical topics in collisional kinetic theory. Handb. Math. Fluid Dyn. **1**, 71 (2002)

Chapter 5
Determinism, Chaos and Reductionism

Everything that is necessary is also easy. You just have to accept it.

F. Durrenmatt

5.1 General Remarks on Determinism

The term *determinism* has often been used in fields other than physics, such as psychology and sociology, causing some bewilderment. For instance, Popper (1992) feared a strictly deterministic world as a nightmare, because it would have meant that our universe is like a big robot, in which we merely play the role of small cogwheels or, at best, of sub-automata.[1] To avoid such misunderstandings, and because, at times, determinism has been improperly associated with reductionism,[2] we are going to briefly review the correct notion of determinism, which is used in physics. For a brilliant and exhaustive discussion on this subject, we refer the reader to Kojeve (1990).

One can readily acknowledge that a completely indeterminate world, whose phenomena obey no rules, would present totally uncorrelated facts and sequences of events, and we would have no chance of ever understanding it. For this reason, words such as *disorder* and *chaos* sound rather pejorative or disruptive to the ideal natural order usually associated with the idea of the "cosmos" since the beginning of Greek philosophy (Thuillier 1991). The historical development of science could be seen as the struggle against disorder, in an attempt to find regularity in phenomena that appear to be irregularly changing. This struggle reached its apogee with the mecha-

[1] This parallels with the totalitarian views is expressed in a paradigmatic way by two classical books of the period (Bauman 2000) by Orwell (1949) and Huxley (1932).

[2] For instance, Popper (1992) argued that a determinist must be a reductionist, although a reductionist is not necessarily a determinist; while others identify reductionism with determinism.

S. Chibbaro et al., *Reductionism, Emergence and Levels of Reality*,
DOI: 10.1007/978-3-319-06361-4_5, © Springer International Publishing Switzerland 2014

nistic determinism shaped in the Enlightenment era, by the belief that the world is a sort of cosmic clock and thus completely predictable.

Our opinion is that the most interesting question is not the determinism of the laws of physics— something that, as we shall see, seems destined to remain unanswered, even neglecting the problems posed by quantum mechanics —but rather the discovery of apparently simple deterministic systems that behave in very irregular fashions. Furthermore, the theory of deterministic chaos has shown that complex behaviours are not the exclusive prerogative of systems of many interacting components, but may also be found in the dynamics of simple systems with few degrees of freedom. Because of this, scientists eventually abandoned the idea that the observed complexity of the world necessarily arises from the cooperation of many elementary building blocks.

But how is the theory of deterministic chaos relevant, in a book on the reduction of theories and the role of singular limits? First of all, recall the important message implied by Lorenz's work on simplified deterministic models of the atmosphere's dynamics (Lorenz 1963): the postulated "elementary building block" is often not elementary at all, and the effort to understand a given system by analyzing the equations governing its parts or constituents may fail. Then, observe that deterministic and stochastic descriptions are deeply different ontologically, but strong similarities can nevertheless be found between certain behaviours of deterministic chaotic systems and processes governed by stochastic laws. Such similarities are relevant on the practical level, e.g. when modeling complex systems. On the other hand, chaos presents both ontic and epistemic questions[3] which may generate confusion about the real conceptual relevance of chaos. We shall see that chaos allows us to unambiguously introduce probabilistic concepts in a deterministic world. Such a possibility is not merely the consequence of our limited knoweledge of the state of the system of interest. Indeed, in order to account for this limited knowledge, one usually relies on a coarse-grained description, which requires a probabilistic approach, and finds that many important features of the dynamics do not depend on the scale ε of the graining, if it is fine enough. At the same time, many results for the $\varepsilon \to 0$ limit do not apply to the cases with $\varepsilon = 0$. Therefore, the probabilistic description of chaotic systems reveals one more instance of singular limits.

In the following, we would like to clarify some aspects of deterministic chaos which, in our opinion, are often misunderstood, leading to scientifically, as well as philosophically, questionable and confused claims. We begin by considering the relationship between determinism and predictability. Then, we shall consider the role of chaos in the statistical description of complex phenomena, for which statistical mechanics providing an important setting.

[3] We shall see how determinism refers to ontic descriptions, while predictability (and in some sense chaos) has an epistemic nature.

5.1.1 Determinism and Predictability

Among the aspects of mechanicism that have continued to influence scientific thought up to the present day, the impact of Laplace's statement reported in Chap. 3 is remarkable. This statement is a milestone of scientific thought, whose legacy has survived despite the advent of the quantum description of physical phenomena. Unfortunately, it has also often been misunderstood, in its technical and conceptual content cf. (Gleick 2008).

We believe that a fair interpretation of Laplace's "mathematical intelligence" was likely due to his desire to stress the importance of prediction in science, as it appears from a famous anecdote, probably apocryphal but frequently cited. We report here the version given by Cohen and Stewart (1994). After seeing Laplace's masterpiece, *Méchanique Céleste*, Napoleon addressed Laplace saying:

> *[t]hey tell me you have written this large book on the system of the universe, and have never even mentioned its Creator*. Laplace answered: *Sire, I have no need of this assumption*. To that Napoleon replied: *Ah! That is a beautiful assumption, it explains many things*, and Laplace: *This hypothesis, Sire, explains everything, but does not permit the prediction of anything. As a scholar, I must provide you with works permitting predictions.*

The Laplace's ideas (sometimes distorted) in the nineteenth century originated a widely accepted view of the science based on three elements (Kojeve 1990):

(a) *Determinism*: the metaphysical assumption of a deterministic causal structure of nature.[4] In mathematical terms, Laplace assumed that every phenomenon is described by a vector $\mathbf{X}$ (the system state) that evolves according to a deterministic law: i.e. if the state $\mathbf{X}(0)$ at the initial time, $t = 0$, is knonwn, then the state at every later instant $\mathbf{X}(t)$, $t > 0$, is uniquely determined. In modern terms, determinism is assured, in a world governed by Newtonian mechanics, by Cauchy's theorem on the existence and uniqueness of the solution of ordinary differential equations.[5]

(b) *Exact predictability*: the practical possibility of making predictions through mathematical laws. This is a delicate point, because it requires an explicitly computable rule for the evolution of $\mathbf{X}(t)$, once the initial state $\mathbf{X}(0)$ is known with arbitrary accuracy.

(c) *Mechanistic reductionism*: the possibility of explaining (at least in principle) any phenomenon from the motion of its elementary constituents, thought to interact through suitable forces.

---

[4] The idea of causality explicitly enters Laplace's as well as our reasoning. Indeed, the strict notion of "causality" leads to considerable difficulties from epistemological and ontological points of view. This does not concern us. In its evolution, classical mechanics has developed a principle of legal determinism, in which the notions of cause and effect are not explicitly invoked. The idea was anticipated by Kant, who stated that the *geschiet* (the effect) presupposes an antecedent (*worauf*) from which it follows according to a rule. The adjective "causal" is still used in the same sense.

[5] We stress the importance of identifying the state vector $\mathbf{X}$ which fully describes the phenomenon under investigation. For instance, in classical mechanics, $\mathbf{X}$ is given by the positions and velocities of particles. This is a fundamental step which took a long time to be understood. For instance, in Aristotelian physics only the positions were considered.

Together, items (a), (b) and (c) summarise Laplace's view, which can be called mechanistic determinism. The followers of mechanistic determinism are reductionists and expect a scientific theory to describe reality in mathematical terms. Given the equations ruling the temporal evolution of a system, and given its initial conditions, the knowledge of the system at any future time can then be obtained. It is important to stress that Newtonian mechanics, which was founded on such premises, was not restricted to a small class of phenomena; it was believed capable, *in principle*, of yielding predictions in all conceivable phenomena: from the orbit of the moon to the motion of falling apples. This fundamental idea is the very essence of mechanistic (or causal) determinism.

Unfortunately, except for extremely simple cases such as the motion of two gravitationally interacting bodies, or the periodic behaviour of pendulums, the time evolution of a system is typically hard to determine explicitly. However, in principle, the equations of motion can be solved, with more or less complicated mathematical calculations. Indeed, generations of astronomers have computed with incredible patience and perseverance the orbits of planets and asteroids, from the equations of classical mechanics. Their successes were numerous, beginning with the derivation of Kepler's laws from the principles of mechanics and the universal law of gravitation, which was given by Newton himself. After obtaining strong agreement between the theoretical calculations and the observations, this approach was systematically confirmed in astronomy. One of its sensational successes was the discovery of the planet Neptune in the nineteenth century. A series of observations indicated a significant deviation of the motion of Uranus from the positions predicted by Newtonian mechanics. Assuming that this discrepancy was not due to a shortcoming of the Newtonian theory, but the presence of an unknown planet, the laws of dynamics and of gravitation led astronomers to calculate the position of this hypothetical planet which, sure enough, was observed by telescope a short while later.

In less rhetorical terms, the essence of Laplace's famous statement is that the laws of classical physics are perfectly deterministic: if the state of a system at a given time is known exactly, its subsequent evolution is uniquely determined. The calculations that took years in the past, when astronomers could only rely on pen and paper, are today performed very rapidly on computers, which determine with great precision the motion of celestial bodies and artificial satellites. The successful use of modern computers in the exploration of space can be seen as another confirmation of Laplace's idea. At least, indeed, popular writers have taken it to be this way.

In 1867, after 20 years of pen-and-paper work, the French astronomer Delaunay completed the calculation of the position of the moon as a function of time, with an accuracy never reached before. In 1970, Deprit, Henrard and Rom checked that calculation with one of the earliest computer algebra systems. The verification took twenty hours and found only three minor mistakes. It is interesting that computers first allowed people to find Delaunay's minor errors, while today the roles are reversed: the analytical calculation of the great astronomer is used to check the accuracy of the new computer algebra systems[1] (Pavelle et al. 1981).

Despite its undisputed success, however, the mechanistic deterministic approach appears to contradict everyday evidence, where there is no way of predicting many

events. Consider the evolution of the weather, falling leaves, or a stone rolling down a slope. How can we reconcile the fundamental assumptions made by Laplace with the apparent irregularity of most phenomena? The simplest way is to think that irregular phenomena appear so only because they require the solution of a very large number of equations, which may also be very complicated. In such cases, which are actually very frequent, it is not possible to solve the problem by pen and paper, and one may think that a sufficiently powerful computer could provide the answer with the desired accuracy.

It is therefore necessary to distinguish the questions concerning the deterministic nature of the laws of physics from those posed by the possibility of making predictions. This is essential to avoid confusion. For example, unlike the majority of physicists and mathematicians, by *deterministic system* Popper (1992) means a system governed by a deterministic evolution law, whose evolution can be in principle predicted with arbitrary accuracy.

Determinism amounts to the metaphysical doctrine that same events always follow from same antecedents. But, as Maxwell had already pointed out in 1873, it is impossible to confirm this fact, because nobody has ever experienced the same situation twice:

> It is a metaphysical doctrine that from the same antecedents follow the same consequences. No one can gainsay this. But it is not of much use in a world like this, in which the same antecedents never again concur, and nothing ever happens twice... The physical axiom which has a somewhat similar aspect is "that from like antecedents follow like consequences". But here we have passed .... from absolute accuracy to a more or less rough approximation.[6]

In these few lines, Maxwell touches on issues which will be later investigated, and anticipates their solution. The issues are:

1. the impossibility of proving (or refuting) the deterministic character of the laws of nature;
2. the practical impossibility of making long-term predictions for a class of phenomena, referred to here as chaotic, despite their deterministic nature.

About 30 years after Maxwell, Duhem (1991), making a remark on a result obtained by Hadamard, concerning a case of what is currently called deterministic chaos, reached the same conclusion. Very similarly to Maxwell, he noted that mathematical deductions are not useful to physicists if they merely state that a proposition, rigorously true, implies the exact truth of another. To be useful to physicists, the mathematical argument must also prove that the second proposition approximately holds if the first is only approximately verified. More formally, Duhem stressed the importance of the fact that solutions of differential equations enjoy a continuous dependence on initial and boundary data, if they have to be of practical interest e.g. in physics.

---

[6] From the conference *Does the progress of Physical Science tend to give advantage to opinion of Necessity* (or Determinism) over that of the Contigency of Events and the Freedom of the Will?, cf. Campbell and Garnett (1882) Chap. XIV.

After the development of quantum mechanics, many think that discussing the deterministic nature of the laws of physics is too academic an exercise to deserve serious consideration. For instance, in a speech motivated by the heated controversy on chaos and determinism between philosophers and scientists, Kampen (1991) bluntly said that the problem does not exist, as it is possible to show that:

> the ontological determinism à la Laplace can neither be proved nor disproved on the basis of observations.[7]

While we fully agree with this statement, we think that the dichotomy concerning determinism and chaos deserves attention well beyond mere scholarly discussions. For instance, there are uncountably many situations lacking a solid mathematical model, such as those common in biology, in which our question has practical implications.

Popper (1992) was an avowed non-determinist, in the sense that he did not accept what he called "scientific determinism": the doctrine according to which the world can be rationally predicted, to any desired degree of accuracy, if a sufficiently precise description of past events, along with all the laws of nature is available. But it is worth remarking that the Popperian definition of "determinism" is different from the one commonly used in physics, because it includes an arbitrarily precise predictability, not required in physics. Apart from questions of terminology, which can be clarified, Popper has made a very important contribution to the issues related to determinism and predictability, since he has convincingly shown that a possible determinism of the laws of nature would not suffice to produce a forecast from "inside". In other words, assuming that Laplace's infinitely capable Intelligence is part of our world, it should predict itself: but no Intelligence can predict all the results of its own forecasts. Nevertheless, a prediction by an external agent remains possible, requiring that the Laplacian Intelligence be placed outside the world, hence requiring that it does not affect the evolution of the world. The discovery of deterministic chaos gave new impulse to these questions.

We have thus argued that determinism and predictability constitute two quite distinct issues, and the former does not imply the latter.

Roughly speaking, determinism can be traced back to a vision of the nature of causality and can be cast in mathematical terms, by saying that the laws of nature are expressed by ordinary differential equations. It is fair to say that most macroscopic phenomena can be described in this way, as confirmed, for instance, by the impressive successes of astronomy in the past and by technological realisations today. However, as noted by Maxwell, the objectively ontological determinism of the the laws of nature cannot be proven; but one might find it convenient to use deterministic descriptions. Moreover, even at a macroscopic level, many phenomena are chaotic

---

[7] In brief, van Kampen's argument is the following. Suppose the existence of a world A which is not deterministic and consider a second world B obtained from the first using the following deterministic rule: every event in B is the copy of an event occurred one million years earlier in A. Therefore, all the observers in B and their prototypes live the same experiences despite the different natures of the two worlds.

and, in some sense, appear to be "random". The meaning of these terms will be clarified shortly. On the other hand, the microscopic phenomena described by quantum mechanics, fall directly within a probabilistic framework. They appear ontologically and epistemologically non-deterministic.

Concerning predictability, the presence of "chaos" in phenomena governed by deterministic laws and the logical aporia proposed by Popper shows that predictability is far from trivial. Two main issues arise: are deterministic phenomena always predictable? And what does prediction mean?

5.2 An Excursus on Chaos

Ironically, in spite of the success of Newtonian mechanics in the discovery of Neptune, the first clear example of what today we call chaos was found in celestial mechanics, the science of regular and predictable phenomena *par excellence*. This is the case of the long standing three-body problem: the motion of three gravitationally interacting bodies, such as the moon, Earth and sun, which was a nightmare for many great early mathematicians as Newton, Euler and Lagrange. In spite of its deterministic nature, Poincaré (1982) found that the evolution of the three-body system can be chaotic, meaning that small perturbations in the initial state, such as a slight change in the initial position of one of the three objects, may lead to dramatic differences in the later states of the system. As a vivid example of sensitivity to initial conditions, we mention the effect of a very distant single electron on massive bodies (Berry 1978). An electron at the limit of the observable universe (a distance of $O(10^{10})$ light years) will lead in just a few collisions to a complete breakdown of the predictability of systems of billiard balls.

There is a widespread vulgate, see e.g. Gleick (2008), which claims that the line of scientific research opened by Poincaré remained basically neglected until 1963, when meteorologist Lorenz rediscovered deterministic chaos while studying the evolution of a simple model of the atmosphere. Therefore, it is often claimed that the new paradigm of deterministic chaos originated in the sixties. This is not true; mathematicians never forgot Poincaré's legacy, although it was not so well known to physicists, (Aubin and Dalmedico 2002).

Here, we briefly recall the essential characteristics of a deterministic chaotic system[8]:

(i) the evolution is given by a deterministic rule, for example, by a set of differential equations;
(ii) solutions sensitively depend on the initial conditions: i.e. two initially almost identical states $\mathbf{X}(0)$ and $\mathbf{X}'(0)$, characterised by a very small initial displacement $|\mathbf{X}(0) - \mathbf{X}'(0)| = \delta_0$, separate at an exponential rate:

$$|\mathbf{X}(t) - \mathbf{X}'(t)| \sim \delta_0 e^{\lambda t}, \tag{5.1}$$

[8] We consider systems whose phase space is bounded.

where λ is positive and is called the Lyapunov exponent[9];

(iii) the evolution of the state $\mathbf{X}(t)$ is not periodic and appears quite irregular, similar in many respects to that of random systems.

Let us start from point (iii) and its relevance to the issue of reductionism. In the deterministic mechanistic approach, the undeniable irregularity of many natural phenomena is thought to be only "apparent". For instance, it is seen as due to a very large number of causes, which are individually thought to be simple. An example of this interpretation of irregular phenomena, which we might call the philosophy of the "simple elementary brick", is afforded by Landau's theory of the onset of turbulence (Landau 1944). This theory states that the very complicated behaviour of a turbulent fluid arises from the superposition of many periodic oscillations, whose individual behaviour is simple by definition. This influential philosophical point of view was, however, refuted by the discovery made by Lorenz (1963), one of the pioneers of the modern theory of chaos. While investigating a minimal model for the dynamics of the atmosphere, he unequivocally realised that the erratic behaviour, typical of turbulent fluids, is not necessarily due to a large number of variables, since it can be found in quite simple and low dimensional dynamics, as a consequence of deterministic chaos. This led to the important conclusion that the elementary bricks are not always "simple". Within this new vision, Ruelle and Takens (1971) showed some years later that the onset of turbulence was not due to a superposition of simple oscillations.

The sensitive dependence on the initial conditions drastically limits the potential to make predictions: if the initial state is known with a certain uncertainty δ_0, the evolution of the system can be accurately predicted with precision Δ only up to a time that depends on the Lyapunov exponent. This quantity is inherent in the system and does not depend on our ability to determine the initial state; hence, recalling 5.1, the time within which the error on the prediction does not exceed the desired tolerance is given by:

$$T_p \sim \frac{1}{\lambda} \ln \frac{\Delta}{\delta_0}. \tag{5.2}$$

Deterministic systems, which are often fairly good models for macroscopic phenomena, can display a behaviour which is chaotic. Their sensitivity to initial conditions introduces an error in predictions which grows exponentially in time. As the exponent is an intrinsic characteristic of the system, predictions remain meaningful only within a time given by 5.2. It is well evident, therefore, that a deterministic nature does not imply the possibility of an arbitrarily accurate prediction.

Furthermore, this major result holds even for simple low-dimensional systems, which leads to a second major conclusion: the reductionistic idea that complex systems can be analysed as an agglomerate of simple elements is incorrect. In general, complex systems cannot be *reduced* to a sum of simple elementary constituents.

[9] Equation (5.1) holds for infinitesimal distance. Because the phase space is bounded, the distance between the two trajectories cannot grow forever and reaches its maximum in a finite time.

5.3 Chaos and Complexity

We have seen that chaos has major consequences for predictability. However, noting that T_p, Eq. (5.2), could be made arbitrarily large by reducing δ_0, though at great costs, because of the slow divergence of the logarithm, it might seem that the problem is only of practical order and not intrinsic to chaotic evolutions. In other words, the limitations on predictability may appear simply epistemological and not ontological, which would imply that the transformation of a deterministic mechanistic problem into a probabilistic one has to be blamed only on our technical inability to sufficiently reduce the error on the initial conditions.[10] This a point of crucial importance.

We shall give some evidence of the impossibility of circumventing this problem, simply by asserting that a deterministic system is, in principle, predictable, on the grounds that the desired accuracy at any given (finite) time t merely requires sufficiently accurate knowledge of the initial conditions with the necessary (finite) precision. Let us consider a deceptively simple dynamical system:

$$x(t+1) = 2x(t) \ \ mod\ 1. \tag{5.3}$$

This system is chaotic and its Lyapunov exponent is $\lambda = \ln(2)$. This means that a small error in the initial conditions doubles at every step. Suppose that $x(0)$ is a known real number in the interval [0, 1], it can be expressed by an infinite sequence of 0 and 1, because it can be written as

$$x(0) = \frac{a_1}{2} + \frac{a_2}{4} + \cdots + \frac{a_n}{2^n} + \cdots$$

where every a_n takes either the value 0 or the value 1. It is also interesting to note that the above binary notation allows us to determine the time evolution by means of a very simple rule: at every step, move the "binary point" of the binary expansion of $x(0)$ by one position to the right and eliminate the integer part. For example, take

$$x(0) = 0.11001010010110010010100101110\ldots \tag{5.4a}$$

Then

$$x(1) = 0.10010100101100100101001011 10\ldots \tag{5.4b}$$

$$x(2) = 0.00101001011001001010010111 0\ldots \tag{5.4c}$$

$$x(3) = 0.01010010110010010100101110\ldots \tag{5.4d}$$

[10] It is worth stressing how dramatically chaos affects our predictions. Because of the logarithm in 5.2, increasing the predictability time T_p by a factor 5 increases the required precision of the initial conditions by five orders of magnitude, e.g. from metre-order precision to micrometre-order precision. For all relevant phenomena this is and will forever remain impossible to be achieved. This is why our local weather forecast are restricted to 5–7 days predictions (roughly speaking the time given by the Lyapunov exponent) and one cannot hope to greatly improve on that by making more accurate measurements of the initial conditions.

and so on. In terms of the sequence $\{a_1, a_2, \ldots, a_n, \ldots\}$, it becomes quite clear how crucially the temporal evolution depends on the initial condition.

Let us now make a brief digression on the notion of "complexity" of a binary sequence. Generally speaking, different types of sequences are possible, for example consider the following ones:

$$11111111111111\ldots \tag{5.4}$$

$$10101010101010\ldots \tag{5.5}$$

$$00101000110100\ldots \tag{5.6}$$

One would presumably state that sequences (5.4) and (5.5) appear to be "ordered", whereas sequence (5.6) seems "complex". Why should one classify the sequences in this way?

In the case of (5.4) and (5.5) the knowledge of the first n values $a_1, a_2, a_3, \ldots, a_n$ appears to be sufficient to predict the following values $a_{n+1}, a_{n+2}, \ldots$. This is not true for sequence (5.6), which seems to be generated by a stochastic, rather than a deterministic, rule. In this case, one could think that the sequence of 0 and 1 is generated tossing a coin, and writing 1 for heads and 0 for tails. One way to formalise this intuitive concept of "complex" behaviour is to associate it with the lack of a constructive rule; then the cases of (5.4) and (5.5) are not complex because they can be generated by means of very simple rules. On a computer, for instance, (5.4) can be generated through a single statement:

```
WRITE 1 N TIMES
```

and similarly for (5.5):

```
WRITE 10 N/2 TIMES.
```

By contrast, (5.6) seems to require a program of the kind:

```
WRITE 0  WRITE 0  WRITE 1  WRITE 0  WRITE 1  ...
```

A precise mathematical formalisation of the complexity of a sequence has been proposed independently in 1965 by Kolmogorov, Chaitin and Solomonoff (Li and Vitanyi 1992). Given the sequence $a_1, a_2, a_3, \ldots, a_n$, among all possible programs which generate this sequence, one considers that with the smallest number of instructions. Denoting by $K^{(N)}$ the number of these instructions, the algorithmic complexity of the sequence is defined by

$$K = \lim_{N\to\infty} \frac{K^{(N)}}{N}. \tag{5.7}$$

Therefore, if there is a simple rule, which can be expressed by a few instructions, the complexity vanishes. If there is no explicit rule, which is not just the complete list of 0 and 1, the complexity is maximal, that is 1. Intermediate values of K between 0 and 1 correspond to situations with no obvious rules, but such that part of the information necessary to do a given step is contained in the previous steps.

To give an intuitive idea of the concept of complexity, let us consider a situation related to the transmission of messages (Chaitin 1990): A friend on Mars needs the tables of logarithms.[11] It is easy to send him the tables in binary language; this method is safe but would naturally be very expensive. It is cheaper to send the instructions necessary to implement the algorithm which computes logarithms.

However, if the friend is not interested in mathematics, but rather in football or the lottery, and wants to be informed of the results of football matches or the lottery draw, there is no way of compressing the information in terms of an algorithm whose repeated use produces the relevant information for the different events; the only option is the transmission of the entire information.

To sum up: the cost of the transmission of the information contained in the algorithm of logarithms is independent of the number of logarithms one wishes to compute. On the contrary, the cost of the transmission of football or lottery results increases linearly with the number of events. One might think that the difference is that there are precise mathematical rules for logarithms, but not for football matches and lottery drawings, which are then classified as random events.

Let us now analyse the problem of transmission, with accuracy Δ, of a sequence $x(t), 0 < t < T$, generated by the rule (5.3). At first glance, the problem seems similar to sending the tables of logarithms, and we could opt for transmitting $x(0)$ and the rule (5.3), which costs a number of bits independent of T. The friend on Mars would then be left with the task of generating the sequence $x(1), x(2), \ldots, x(T)$. However, we must also choose the number of bits to which $x(0)$ should be specified. From (5.2), the accuracy Δ at time T requires accuracy $\delta_0 \sim 2^{-T}\Delta$ for $x(0)$, hence that the number of bits specifying $x(0)$ grows with T. Again, we have to tackle the problem of the complexity of a sequence of symbols, $\{a_0, a_1, \ldots\}$. The fact is that there are "simple" initial conditions, of the type (5.4) or (5.5), which can be specified by a number of instructions independent of the length of the sequence, but there are complex sequences as well.

The determination of the algorithmic complexity of a sequence is impossible, as implied by Gödel's incompleteness theorem. Notwithstanding this impossibility, a result of Martin-Löf (1966) shows that "almost all" binary sequences, which express the real numbers in [0, 1], are complex. Therefore, the major conclusion is that the details of the time evolution are well hidden in the initial condition and that, in general, is complex.

The immediate and striking consequence of these facts is that determining with arbitrary precision of the initial conditions is hopeless. Hence, long-term predictions with the desired accuracy are impossible in principle, despite seeming quite

[11] In the pre-computer age, numerical computations relied on tabulated numbers for, e.g. logarithmic and trigonometric functions.

reasonable for such a simple time evolution. Insisting nonetheless on following such a path, one faces an infinite regression[12] and ineluctably runs into an impossibility which is not merely practical (Li and Vitanyi 1992).

So far, we have observed that deterministic systems, even with just a few degrees of freedom, may exhibit chaos (that is, a sensitive dependence on initial conditions). This fact strongly impacts on the possibility of making accurate predictions beyond a certain predictability time T_p. In turn, most initial conditions are complex, hence the predictability of chaotic systems is intrinsically limited in time. One may then state that predictable deterministic systems and chaotic unpredictable ones are related by a singular limit. Indeed, as mentioned earlier, the conclusions one may draw for vanishing and for arbitrarily small, but finite errors, are completely different. This singularity highlights the relevance of chaos to reductionism. It also shows that elementary constituents of a given object may indeed have very complex behaviour themselves. Moreover, this singularity clarifies how, in many situations, stochastic macroscopic properties emerge from chaos. We will return to this point at the end of the chapter.

5.4 Chaos and Probability

Because of their irregular behaviour, deterministic chaotic systems share many features with stochastic processes. In particular, the unpredictability of a chaotic system calls for statistical or probabilistic approaches, analogously to the case of stochastic processes. For instance, trying to predict the motion of a fluid particle in a turbulent flow is meaningless, while it is possible and appropriate to predict its statistical features such as its average velocity, kinetic energy, etc. This fact is very interesting from a practical point of view, but it is rather subtle and can lead to confusion. An important characterisation of the dynamics, on a coarse-grained scale, is given the Kolmogorov–Sinai (K–S) entropy, defined as follows.[13]

Let $\mathscr{A} = \{A_1, \ldots, A_N\}$ be a finite partition of the phase space, made up of the N disjoint sets A_i, and consider the sequence of points

$$\{\mathbf{x}(0),\ \mathbf{x}(1),\ \mathbf{x}(2),\ \ldots, \mathbf{x}(n), \ldots\} \tag{5.9}$$

which constitutes the trajectory with initial condition $\mathbf{x}(0)$. This trajectory can be associated with the symbol sequence

$$\{\sigma(\mathbf{x}(0)),\ \sigma(\mathbf{x}(1)),\ \sigma(\mathbf{x}(2)), \ldots, \sigma(\mathbf{x}(n)), \ldots\} = \{i_0, i_1, i_2, \ldots, i_n, \ldots\} \tag{5.10}$$

[12] In philosophical language the classical trilemma of Agrippa: if we are asked to prove how we know something, we can provide a proof or an argument. Nonetheless, a proof of the proof can be then asked and so on, leading to an infinite process which never ends.

[13] For the sake of simplicity, we restrict ourselves to the case of discrete-time dynamical systems, but continuous systems may be treated analogously.

where $i_n \in \{1, 2, \ldots, N\}$ and $\sigma(\mathbf{x}(k)) = i_k$ if $\mathbf{x}(k) \in A_{i_k}$. The coarse-grained properties of chaotic trajectories can be therefore studied through the discrete time sequence (5.10). Let $C_m = (i_1 i_2 \ldots i_m)$ be a "word" of length m and probability $P(C_m)$. The quantity

$$H_m(\mathscr{A}) = -\sum_{C_m} P(C_m) \ln P(C_m) \tag{5.11}$$

is called the block entropy of the m-sequences.[14] In the limit of infinitely long sequences, the asymptotic entropy increment

$$h_S(\mathscr{A}) = \lim_{m\to\infty} (H_{m+1} - H_m)$$

is called the Shannon entropy, and depends on the partition $\mathscr{A}$. Taking the largest value over all possible partitions we obtain the so-called Kolmogorov–Sinai entropy:

$$h_{KS} = \sup_{\mathscr{A}} h_S(\mathscr{A}).$$

A more tractable and intuitive definition of h_{KS} starts from the partition $\mathscr{A}_\varepsilon$ made of a grid of hypercubes of sides of length ε, and takes the following limit:

$$h_{KS} = \lim_{\varepsilon\to 0} h(\mathscr{A}_\varepsilon).$$

Although h_{KS} and K are conceptually very different characterisations of a symbol sequence,[15] their numerical values are simply related:

$$h_{KS} = \lim_{N\to\infty} \frac{\langle K^{(N)} \rangle}{N \ln 2}, \tag{5.12}$$

where $\langle . \rangle$ denotes an average over all sequences of length N. This leads to the following maxim:

Complex = Incompressible = Unpredictable.

[14] Shannon (1948) showed that, once the probabilities $P(C_m)$ are known, the entropy (5.11) measures, under natural conditions, the *surprise* or information carried by $\{C_m\}$.

[15] Consider the following two m-sequences, produced by tossing a fair coin:

01010101010 . . . 010101
01001010110 . . . 101001

One finds that the first sequence is compressible, while the second appears to be stochastic, in spite of the fact that both occur with probability 2^{-m}. This shows that algorithmic complexity, which characterises a single sequence, and information, which amounts to a probabilistic notion, are conceptually different.

which is valid for stochastic processes, e.g. Markov chains, as well. In the final section of this chapter, we will return to this similarity of chaotic deterministic systems and random sequences.

We conclude this section noting that initial conditions play a key role even in deterministic chaotic dynamics, just as they do in the problem of irreversibility. In Chap. 4, we saw that reversible large mechanical systems display an irreversible behaviour, for almost all nonequilibrium initial conditions. Analogously, in deterministic chaotic systems, unpredictable evolutions arise for almost all initial conditions, apparently in conflict with the deterministic nature of the dynamics. Hence, both chaotic systems and systems with many degree of freedoms enjoy a complex nature, requiring probabilistic approaches. Both are characterised by the transition to a complex probabilistic state through a singular limit, which is $\varepsilon \to 0$ for chaotic systems and $N \to \infty$ for systems with many degree of freedom.

5.5 Quarrels on Chaos and Determinism: Chaos and Probability Revisited

The discovery of chaos, in particular the impossibility of making long-term predictions for deterministic systems, has generated a debate about determinism, randomness and, more generally, complexity. The debate has often been heated (Amsterdamski et al. 1990). Here, it suffices to recall some of its most interesting aspects. In his long-lasting diatribe against Prigogine, the father of catastrophe theory, the mathematician Thom, argues in uncompromising terms that being attracted by the charm of randomness is the symptom *par excellence* of an anti-scientific attitude, since it largely proceeds from admiration to confusion. According to Thom, humanists could be forgiven for such an attitude, but not scientists, who should be accustomed to the rigour of scientific rationality. He insists with great determination that randomness is a negative concept, hollow, and devoid of any scientific interest, whereas determinism is an object of fascinating richness (Amsterdamski et al. 1990).

Although not in complete agreement with all positions taken by Thom, we share the concern that chaos could be used as some sort of anti-science passkey. Unlike what some may think, deterministic chaos, and its inability to make predictions, does not provide any proof of the weakness of "classical" or "standard" science, which would have been eventually forced to abandon determinism. Chaos provides no evidence of the inability of official science to deal with the complexity of the real world; nor does it highlight any difficulty that calls for an alternative science.

The impossibility for a deterministic science to make long-term, arbitrarily accurate predictions, is indeed a consequence of deterministic chaos, but by no means does it lead to the impossibility of any form of accurate prediction. In particular, according to Thom, any model of a real phenomenon must be deterministic, in order to teach us something.

Because of chaos, the role of probability in physics takes further meanings. In the nineteenth century, this role was acknowledged by Maxwell and Boltzmann in relation to thermodynamics. To understand the properties of a gas starting from the microscopic details of the dynamics of its molecules is not only hard but is also misleading; only a statistical description, which takes advantage of the huge number of particles involved, and describes the gas in terms of a few macroscopic variables such as pressure, temperature, etc. is appropriate. Indeed, in Chap. 3 we observed that a system containing a very large number of particles is described by particular emerging laws: the so-called statistical laws, which are due to the large number of particles constituting the body,[16] and which cannot in any way be derived from purely mechanical laws. Although the elementary constituents of a system with a large number of degrees of freedom obey the same laws of mechanics as those of a system with a small number of degrees of freedom, the large number implies *qualitatively* different new laws (Landau and Lifshitz 1980).

As a consequence of the large number of particles, the macroscopic level is characterised by a sort of "statistical determinism", as in transport equations: the Navier-Stokes equations for the velocity of a fluid, Fourier's law for the temperature, Fick's law for diffusion are all deterministic, and result from the fact that the statistical analysis is exact with such large numbers of interacting objects. An example of "statistical determinism" in our daily lives is the sweeping of dust with a broom. In order to collect the dust into a corner, one tries to push the dust grains with the broom horsehair. Of course, the single hair cannot push a grain. However, the broom has lots of hair, so sweeping many times eventually achieves the goal.

Regardless of the rivers of ink shed in philosophical controversies, perhaps the greatest technical significance of the discovery of deterministic chaos is that it reveals that the statistical approach is necessary also in systems with few degrees of freedom. A statistical approach is obviously necessary if the number of degrees of freedom is very large, but in the presence of deterministic chaos it is necessary, independent of the number of variables involved.

An example is given by the Lorentz gas, further idealised by the Sinai Billiard, in which a particle moves with constant speed bouncing over fixed circular obstacles (Dorfman 1999), see Fig. 5.1.

Because the dynamics is unstable, the motion looks very similar to Brownian motion. Indeed, with regard to diffusion in the long-time limit, it is possible to prove that the particle in the Sinai billiard enjoys exactly the same statistical properties as a Brownian particle. In the latter case the irregularity of the motion is due to the presence of many fluid molecules randomly colliding with it; the motion in the Sinai billiard has no randomness, but trajectories are chaotic, due to the instability. The overall effect, as far as diffusion is concerned, is the same.

As observed above about the statistical description of both thermodynamic and chaotic systems, the probabilistic approach could be considered as merely a method to cope with our limited ability to accurately control the systems of interest. In statistical

[16] To be rigorous, this is true for particles interacting through a potential, i.e. in all cases of physical interest.

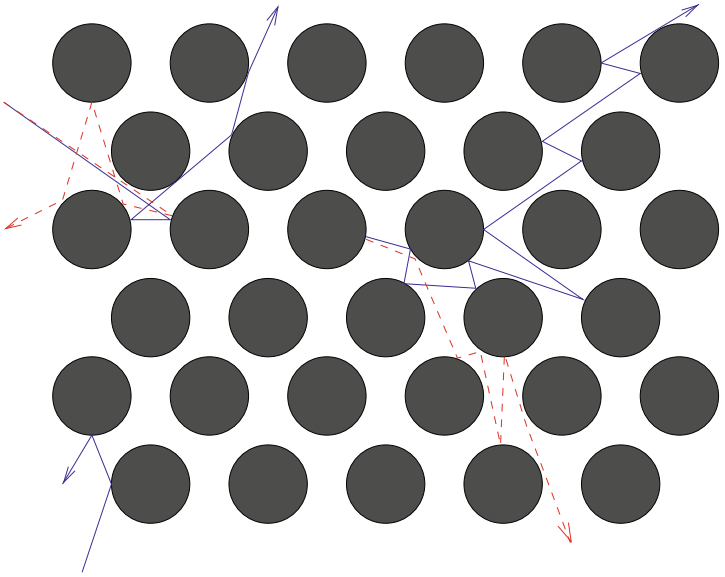

Fig. 5.1 Examples of trajectories of particles bouncing over fixed *circular* obstacles; note the divergence of initially close trajectories

mechanics, the difficulty is due to the large number of degrees of freedom, whereas in chaotic systems it arises from the sensitive dependence on initial conditions. By contrast, quantum mechanics is intrinsically stochastic; the position and momentum of the system cannot be determined with arbitrary precision, because a bound is imposed by Heisenberg's uncertainty principle. Hence, probabilities are unavoidable.

However, in light of our arguments, it seems fair to claim that the vexed question of whether the laws of physics are deterministic or probabilistic has, and will have, no definitive answer. On the sole basis of empirical observations, it does not appear possible to decide between these two contrasting arguments:

(i) Laws governing the universe are inherently random, and the determinism that is believed to be observed is in fact a result of the probabilistic nature implied by the large number of degrees of freedom;
(ii) the fundamental laws are deterministic, and seemingly random phenomena appear so due to deterministic chaos.

Basically these two positions can be viewed as a reformulation of the endless debate on quantum mechanics: thesis (i) expresses the inherent indeterminacy claimed by the Copenhagen school, whereas thesis (ii) illustrates the hidden determinism advocated by Einstein (Pais 2005).

5.6 Concluding Remarks

By way of conclusion, we would like to make a few remarks on the importance of chaos from a conceptual point of view and in the context of present-day research.

The most important findings are the following:

1. Deterministic systems, even with just a few degrees of freedom, may be sensitive to initial conditions, hence unpredictable except in the short term.
2. Chaotic systems are complex. Complexity can be rigorously defined in terms of algorithmic complexity, which is a notion of incompressibility hence of unpredictability. Moreover, *almost all* initial conditions of a generic deterministically chaotic system are complex, hence almost all trajectories are complex.
3. The elementary bricks of complex systems may have far from elementary behaviour, and be complex themselves.
4. A probabilistic description is needed both for chaotic systems and for systems with many degrees of freedom. In both cases, new *statistical* laws emerge from the underlying deterministic framework.
5. If a given phenomenon appears irregular or disordered, it is practically impossible to check whether this is due to chaos, to the presence of many interacting degrees of freedoms, or to some intrinsic randomness.
6. Analogously to the case of the singular limit of statistical mechanics, the singular nature of the chaotic limit allows neither practically nor conceptually the reduction of chaotic macroscopic phenomena to deterministic mechanistic laws. From a philosophical perspective, this is another case of strong emergence.

The discovery of an irregular chaotic behaviour in systems with few degrees of freedom and apparently innocent evolution laws seems to contradict the supporters of the "elementary brick" philosophy. A revealing example of the difficulties faced by this philosophy is given by Lorenz's celebrated model: if we reduce the hydrodynamic equations to elementary, or simple, structures we do not necessarily find simple behaviour, hence we do not necessarily increase our understanding of (for example) turbulence. This teaches us two general lessons, which are of practical importance:

(a) complex (unpredictable) behaviours are not necessarily produced by complicated structures, such as structures made of many components, but are common in simple and low dimensional dynamics;
(b) the methodological approach ["micro-reductionism" in the words of Smith (1998)], which seeks to understand and control dynamics by determining the equations ruling the interactions of its parts, can fail. We may say that: *knowing the Navier-Stokes equation does not solve the problem of understanding turbulence*.

It is a matter of fact that finding solutions, or merely approximate solutions, to the classical "initial value" problem (i.e. to differential equations once the initial state is given) is not a viable approach in many interesting situations characterised by complex (or complicated) behaviour. Even when detailed knowledge of the evolution

laws is given, or presumed to be given, the presence of chaos and/or large numbers of degrees of freedom foils the initial value problem, because an unlimited amount of information on the initial state would be required.

Therefore, rather than considering the properties of specific trajectories originating from given initial states, one is forced to adopt a new strategy based on the statistical information carried by an ensemble of trajectories. This task is usually accomplished with the aid of computers, which have thus played a key role in developing the theory of dynamical systems and chaos. Indeed, the wealth of behaviour of nonlinear systems has been unveiled and systematically and quantitatively characterised thanks only to the fast computations and visualisations made possible by computers. Understanding the practical and conceptual problems posed by chaotic dynamics has led to a shift towards probabilistic or, at times, qualitative approaches, in science.

To better appreciate this recent shift in approach, consider the paradigmatic example of pre-chaos approaches to complex systems, constituted by von Neumann's belief that powerful computers and a clever use of numerical analysis would eventually lead to accurate forecasts, and even to the control, of weather and climate:

> The computer will enable us to divide the atmosphere at any moment into stable regions and unstable regions. Stable regions we can predict. Unstable regions we can control.[17]

The great scientist von Neumann was wrong, but he did not know the phenomenon of deterministic chaos.

Despite the exponentially fast growth of computing power, the forecasting ability of even the largest weather forecasting centres advances rather slowly (Yoden 2007). Modern weather forecasters have two goals: ever more accurate and detailed predictions, and advances in cognition and qualitative understanding. However, even the standard activity is carried out with perspectives different from von Neumann's. The intrinsic limitations on predictability, inherent in the chaotic nature of the atmosphere, require meteorologists to run series of forecasts, known as ensemble forecasts, each member of which starts from a slightly different initial condition, in order to produce data for a probabilistic concept of the forecasts. Is this surrendering before the tasks of prediction and detailed description of weather and climate? In fact, we simply believe that this change of perspective is dictated by the evidence that, in the field of complex systems, one may only investigate problems that are physically well-posed.

Because detailed predictions are impossible in chaotic systems, one wonders whether the study of oversimplified chaotic models of physical phenomena improves our understanding of the behaviour of real-world systems, or is irrelevant to that end. This raises, in turn, the general question of the relationship between scientific theories and the part of the real world they aim to describe, including, in particular, the role played in this relationship by mathematical models and numerical simulations.

Roughly speaking, we can identify two main categories of numerical simulation, although their boundaries are sometimes blurred:

---

[17] Cited in Dyson (2009).

(i) Accurate numerical simulations which approximate the solution of equations representing, or thought to represent, a given phenomenon.
(ii) Numerical implementations of models which, retaining the basic features of a real system, are crude simplifications, or phenomenological caricatures of "realistic" models.

Class (i) includes, for example, standard direct numerical simulations of the Navier-Stokes equations, or the full N-body gravitational problem in celestial mechanics. This computational approach is the most obvious, and reflects the etymological origin of the term "computer": from the Latin computare "to count", "to sum up". The idea underlying this use of computers is that systems can be completely known and reproduced *in silico*, once the equations representing their properties are solved numerically.

Class (ii), instead, presupposes some kind of modelling activity. As an explicit connection between models and reality is not available or, more generally, is not even required, the results of numerical computations only concern the abstract mathematical structures of the model. As such, they can be considered as mere metaphors for of the original phenomenon. Typical examples are: Lorenz's model, which is a caricature of Boussinesq's equation; coupled map lattices that constitute a prototype for spatially extended systems, but are far from representing any of them; the Lotka-Volterra equations, that describe some basic mechanisms of competition between prey and predator species, whose real dynamics is unknown (Cencini et al. 2009).

One should also beware of the possible confusion between ontic and epistemic descriptions, when studying the problems of chaos.

Determinism simply means that: given the same initial state $\mathbf{X}(0)$, one always finds the same evolved state $\mathbf{X}(t)$, at any fixed later time $t > 0$. Therefore, determinism refers exclusively to ontic descriptions, and it does not deal with predictions. This has been clearly stressed by Atmanspacher, in a paper by the rather eloquent title *Determinism is ontic, determinability is epistemic*, (Atmanspacher 2002). This distinction between ontic and epistemic descriptions was obvious to Maxwell; after having noted the metaphysical nature of the problem of determinism in physics, he stated that:

> There are certain classes of phenomena... in which a small error in the data only introduces a small error in the result...There are other classes of phenomena which are more complicated, and in which cases of instability may occur.[18]

On the contrary, Popper (1992) confused determinism and prediction:

> Scientific determinism is the doctrine that the state of any closed physical system at any future instant can be predicted.

In the previous section, we considered arguments, e.g. by van Kampen, which deny that determinism may be decided on the basis of observations. This conclusion is also

[18] From the conference *Does the progress of Physical Science tend to give advantage to opinion of Necessity* (or Determinism) over that of the Contingency of Events and the Freedom of the Will?, see Campbell and Garnett (1882).

reached from detailed analyses of sequences of data produced by the time evolutions of interest. Computing the so-called ε-entropy and the Finite-Size Lyapunov Exponents, at different resolution scales ε, one cannot distinguish potentially underlying deterministic dynamics from stochastic ones. The analysis of temporal series can only be used, at best, to pragmatically classify the stochastic or chaotic character of the observed signal, within certain scales (Cencini et al. 2009).

At first, this could be disturbing: not even the most sophisticated time-series analysis that we could perform reveals the "*true nature*" of the system under investigation, the reason simply being the unavoidable finiteness of the resolution we can achieve. More sophisticated instruments, with the much higher resolution that can be envisaged for the future, will not change this fact, as their resolutions will nevertheless always be finite. On the other hand, one may be satisfied with a non-metaphysical point of view, in which the *true nature* of the object of investigation is not at stake. The advantage is that one may choose whatever model is more appropriate or convenient to describe the phenomenon of interest, especially considering the fact that, in practice, one observes and wishes to account for only a limited set of coarse-grained properties. These properties are typically equivalently obtained from a variety of different underlying dynamics.

Chaotic systems and, more precisely, those which are ergodic, naturally lead to probabilistic descriptions in the presence of deterministic dynamics. In particular, ergodic theory justifies the *frequentist* interpretation of probability, according to which the probability of a given event is defined by its relative frequency. Therefore, assuming ergodicity, it is possible to obtain an empirical notion of probability which is an objective property of the trajectory (von Plato 1994).

There is no universal agreement on this issue; for instance, Popper (2002) believed that probabilistic concepts are extraneous to a deterministic description of the world, while Einstein held the opposite view, as expressed in his letter to Popper:

> I do not believe that you are right in your thesis that it is impossible to derive statistical conclusions from a deterministic theory. Only think of classical statistical mechanics (gas theory, or the theory of Brownian movement).[19]

Naively, one might consider the statistical properties of chaotic systems to be illusory, because they only result from observational limitations. Apparently, such a conclusion is confirmed by the fact that important measures of the dynamical complexity, such as the Lyapunov exponent λ and the Kolmogorov–Sinai entropy h_{KS}, are defined via finite, albeit arbitrarily high, resolutions. For instance, in the computation of λ one considers two trajectories, which are initially very close $|\mathbf{X}'(0) - \mathbf{X}(0)| = \delta_0$ and diverge in time from each other. Similarly, h_{KS} is computed introducing a partition of the phase space, whose elementary cells have a finite size ε. However, in the small-ε limit, the value of h_{KS} asymptotically tends to a value that no longer depends on ε, as happens to λ in the small-δ_0 limit. Therefore, these measures of the chaotic properties of given dynamics can be considered intrinsic properties of the dynamics themselves: they do not depend on our observation ability, provided it is finite, i.e. provided ε and δ_0 do not vanish.

[19] The letter is reprinted in Popper (2002).

According to Primas (2002), measures of stability, such as the Lyapunov exponent, concern ontic descriptions, whereas measures of information content or information loss, such as the Kolmogorov–Sinai entropy, relate to epistemic descriptions. We agree as far as stability is concerned.

Regarding the epistemic character of h_{KS}, we observe that the Shannon entropy of a sequence of data, as well as the Kolmogorov–Sinai entropy, enjoy an epistemic status from a certain viewpoint, but not from another. The epistemic status arises from the fact that information theory deals with transmission and reception of data, which is necessarily finite. On the other hand, h_{KS} is definitely an objective quantity, which does not depend on our observational limitations, as demonstrated by the fact that it can be expressed in terms of Lyapunov exponents.[20] Therefore, the Kolmogorov–Sinai entropy can be considered as a concept which links deterministic and stochastic descriptions.

References

Amsterdamski, S. et al.: La Querelle du déterminisme: Philosophie de la science d'aujourd'hui. Gallimard, Paris (1990)
Atmanspacher, H.: Determinism is ontic, determinability is epistemic. In: Atmanspacher, H., Bishop, R. (eds.) Between Chance and Choice. Imprint Academic, Thorverton (2002)
Aubin, D., Dalmedico, A.D.: Writing the history of dynamical systems and chaos: longue durée and revolution, disciplines and cultures. Hist. Math. **29**(3), 273 (2002)
Bauman, Z.: Liquid Modernity. Polity, Cambridge (2000)
Berry, M.V.: Regular and irregular motion in topics in nonlinear mechanics. In: American Institute of Physics Conference Proceedings, vol. 46, p. 16. AIP (1978)
Campbell, L., Garnett, W.: The Life of James Clerk Maxwell. MacMillan and Co., London (1882)
Cencini, M., Falcioni, M., Olbrich, E., Kantz, H., Vulpiani, A.: Chaos or noise: difficulties of a distinction. Phys. Rev. E **62**(1), 427 (2000)
Cencini, M., Cecconi, F., Vulpiani, A.: Chaos: From Simple Models to Complex Systems. World Scientific, Singapore (2009)
Chaitin, G.J.: Information, Randomness and Incompleteness: Papers on Algorithmic Information Theory. World Scientific, Singapore (1990)
Cohen, J., Stewart, I.: The Collapse of Chaos: Discovering Simplicity in a Complex World. Viking Penguin, New York (1994)
Dorfman, J.R.: An Introduction to Chaos in Nonequilibrium Statistical Mechanics. Cambridge University Press, Cambridge (1999)
Duhem, P.: The Aim and Structure of Physical Theory. Princeton University Press, Princeton (1991)
Dyson, F.: Birds and frogs. Not. AMS **56**(2), 212 (2009)
Gleick, J.: Chaos: Making a New Science. Penguin, Harmondsworth (2008)
Huxley, A.: Brave New World. Buccaneer Books, New York (1932)
van Kampen, N.G.: Determinism and predictability. Synthese **89**, 273 (1991)

[20] The so-called Pesin formula:

$$h_{KS} = \sum_{i:\lambda_i>0} \lambda_i,$$

expresses the Kolmogorov–Sinai entropy as the sum of the positive Lyapunov exponents, (Cencini et al. 2009). The first Lyapunov exponent λ_1 is the λ introduced in (5.1).

Kojeve, A.: L'idée du déterminisme dans la physique classique et dans la physique moderne. Librairie générale française (1990)
Landau, L.D.: On the turbulence. Dokl. Akad. Nauk. SSSR **44**, 339 (1944)
Landau, L.D., Lifshitz, E.M.: Statistical Physics, Part 1. Pergamon, New York (1980)
Li, M., Vitanyi, P.: Inductive reasoning and Kolmogorov complexity. J. Comput. Syst. Sci. **44**(2), 343 (1992)
Lorenz, E.N.: Deterministic nonperiodic flow. J. Atmos. Sci **20**(130), 130 (1963)
Martin-Löf, P.: The definition of random sequences. Inf. Control **9**(6), 602 (1966)
Orwell, G.: Ninety Eighty Four. Secker and Warburg, London (1949)
Pais, A.: Subtle is the Lord: The Science and the Life of Albert Einstein. Oxford University Press, Oxford (2005)
Pavelle, R., Rothstein, M., Fitch, J.: Computer algebra. Sci. Am. **245**, 136 (1981)
Poincaré, H.: Les méthodes nouvelles de la mécanique céleste. Gauthier-Villars, Paris (1982)
Popper, K.R.: The Open Universe: An Argument for Indeterminism. From the Postscript to the Logic of Scientific Discovery. Routledge, London (1992) (Reprint edition)
Popper, K.R.: The Logic of Scientific Discovery. Routledge, London (2002)
Primas, H.: Hidden determinism, probability, and times arrow. In: Bishop, R., Atmanspacher, H. (eds.) Between Chance and Choice, p. 89. Imprint Academic, Exeter (2002)
Ruelle, D., Takens, F.: On the nature of turbulence. Commun. Math. Phys. **20**(3), 167 (1971)
Shannon, C. E.: A note on the concept of entropy. Bell System Tech. J **27**, 379–423 (1948)
Thuillier, P.: La revanche du dieu chaos. Recherche **232**, 542 (1991)
von Plato, J.: Creating Modern Probability. Cambridge University Press, Cambridge (1994)
Yoden, S.: Atmospheric predictability. J. Met. Soc. Jap. **85**, 77 (2007)

Chapter 6
Quantum Mechanics, Its Classical Limit and Its Relation to Chemistry

> *Physics is mathematical, not because we know so much about the physical world, but because we know so little: it is only its mathematical properties that we can discover. For the rest our knowledge is negative.*
>
> B. Russell

In 1929 P.A.M. Dirac, one of the fathers of quantum mechanics, wrote the following celebrated sentence, which is often quoted in discussions of the reduction of chemistry to physics:

> The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved.

It is true that it is possible to find the explicit solution of the Schrödinger equation for the hydrogen atom and therefore to explain with high accuracy the phenomena observed in experimental spectroscopy. On the other hand, nobody is able to repeat this exact mathematical analysis on systems with many electrons. Nevertheless, with the help of computers and suitable numerical methods, it seems that there are no particular limitations to determining approximate solutions even for heavy atoms[1] and, in general, complex materials.

In 1998 the Nobel prize for chemistry was awarded to the physicists W. Kohn and J. Pople, for their contributions to a rather powerful mathematical technique, for many-body systems, known as density functional theory, which allows us to treat quantum systems made of a large number of particles. Many theoretical physicists interpreted this success as a confirmation of the correctness of the point of view of Dirac, a triumph of the imperialist process in which a science (chemistry in this case) that borders physics is brought to order via the use of the first principles of quantum mechanics.

[1] Basically the solution of the Schrödinger equation for the hydrogen atom is the starting point of approximations for heavier atoms, e.g. using the Hartree-Fock or Thomas-Fermi methods.

S. Chibbaro et al., *Reductionism, Emergence and Levels of Reality*,
DOI: 10.1007/978-3-319-06361-4_6,

In spite of the opinion of the great English scientist and of the members of the Nobel prize committee, which legitimately awarded to physicists the prize for chemistry, the view that chemistry is nothing but applied quantum mechanics is rather questionable. We will see that the use of physics for chemical phenomena is not a merely deductive affair which, starting from quantum mechanics and using approximations introduced just for "practical reasons", predicts specific quantitative behaviour. On the contrary, the construction of modern chemistry has been developed largely independently of quantum mechanics Scerri (2008).

Before entering a specific discussion on the topics of the reduction of chemistry to physics (namely to quantum mechanics) we address another issue which is important for chemistry, besides being interesting of its own merits: the (non-trivial) relation between quantum mechanics and classical mechanics, cf. the book by Bokulich (2008) and the insightful review by Berry (2010, 2013).

6.1 Classical Versus Quantum Mechanics

Classical mechanics and quantum mechanics are two superbly successful theories that describe, in their respective ranges of applicability, real world phenomena. Textbooks often suggest that classical mechanics, after all, is nothing but the $\hbar \to 0$ limit of quantum mechanics. This would mean that quantum mechanics is more fundamental than classical mechanics, in the sense that that the latter theory is just an approximation of the "correct" quantum mechanical description. We would have then produced a paradigmatic example of reductionism, as defined in the texbooks of philosophy of science.

These kinds of claims can be appropriate from a pedagogical point of view, by sparing young students from too brutal a shock at the outset of their studies of modern physics. On the other hand, any reflection that is not too superficial would convince us that the scenario cannot be so simple. It suffices to remind oneself that classical mechanics is a deterministic theory, while quantum mechanics has an intrinsically indeterminate character. Moreover, chaotic behaviour is very common in classical systems, while the time evolution is typically periodic, or quasiperiodic, in quantum mechanics.

It is well known that Einstein, as well as Schrödinger, vehemently disagreed with the so-called Copenhagen school. Even disregarding the arguments of the founding fathers, we find it remarkable that the very supporters of the "orthodox intepretation" of Copenhagen have quite different views on some basic aspects of quantum mechanics.

Let us start with the opinions of N. Bohr who can be considered a sort of patriarch of quantum mechanics. He repeatedly emphasised that quantum mechanics requires a classical framework, in order to be formulated:

> According to the view of the author, it would be a misconception to believe that the difficulties of the atomic theory may be evaded by eventually replacing the concepts of classical physics by new conceptual forms .... No more is it likely that the fundamental concepts of the

> classical theories will ever become superfluous for the description of physical experience. (Bohr 1929)

This point of view is expressed very clearly by Heisenberg (1963):

> Newtonian mechanics is a kind of a priori for quantum theory. It is a priori in the sense that it is the language which enables us to say what we observe.[2]

and repeated (sometimes with a certain embarrassment) by influential texbooks, such as that of Landau and Lifshitz (1978):

> It is in principle impossible ... to formulate the basic concepts of quantum mechanics without using classical mechanics.
>
> Thus quantum mechanics occupies a very unusual place among physical theories: it contains classical mechanics as a limiting case, yet at the same time it requires this limiting case for its own formulation.

The above quotations show that the relation between quantum mechanics and classical mechanics is not so innocent: it is too naive just to consider quantum mechanics as the fundamental theory and classical mechanics as its $\hbar \to 0$ limit.

Even the oft-quoted correspondence principle does not help. Despite many statements in the physics literature, this principle was not intended as a requirement for classical mechanics to be recovered from quantum mechanics, in the limit of large quantum numbers, or other kinds of limits:

> This Correspondence Principle must be regarded purely as a law of quantum theory, which can in no way diminish the contrast between the postulates and electrodynamic theory. Bohr (1924)

Far from being a mere scholar curiosity, we deem it interesting to cite the non-converging opinions of Heisenberg and Dirac:

> We no longer say "Newtonian mechanics is false and must be replaced by quantum mechanics which is correct." Instead we adopt the formula "Classical mechanics is a self-contained scientific theory. It is a strictly 'correct' description of nature wherever its concepts can be applied." (Heisenberg 1948)
>
> My own opinion is that we ought to search for a way of making fundamental changes not only in our present Quantum Mechanics, but actually in Classical Mechanics as well. Since Classical Mechanics and Quantum Mechanics are closely connected, I believe we may still learn from a further study of Classical Mechanics. In this point of view I differ from some theoretical physicists, in particular Bohr and Pauli. Dirac (1951)

In the nice book of Bokulich (2008), readers interested in the history of physics may find a detailed discussion on the epistemological views of three fathers of quantum mechanics: W. Heisenberg, P.A.M. Dirac and N. Bohr.

Heisenberg's philosophical opinion can be classified as a "strong form theoretical pluralism".[3] The most central aspect of Heisenberg's philosophy is the concept of

[2] One can say that for Heisenberg, in the context of quantum mechanics, the classical mechanics is a Kantian "a priori".

[3] Quite similar to the "metaphysical nomological pluralism" of N. Cartwright: different domains of nature are ruled by different systems of laws.

"closed theory" i.e. of a system of axioms, laws and so on describing a certain set of phenomena in a correct self-consistent fashion. Heisenberg (1948) identifies four closed theories in physics: classical mechanics, electromagnetism (including optics and special relativity), statistical mechanics (including thermodynamics) and quantum mechanics:

> for each of these realms there is a precisely formulated system of concepts and axioms, whose propositions are strictly valid within the particular realm of experience they describe.

Roughly speaking, Heisenberg's views amount to a picture of science which is not unified, but consists of non-overlapping fields, each with its own methods and range of applicability.

On the contrary, Dirac (1962) considers quantum mechanics and classical mechanics as open theories, not immune from future revision:

> I rather got to the idea that everything in nature was only approximate, and that one had to be satisfied with approximations, and that science would develop through getting continually more and more accurate approximations, but would never attain complete exactness.

Surely Heinsenberg and Dirac, although from different standpoints, cannot be considered reductionists, with no particular interest in the unity of science.

6.1.1 Is Classical Mechanics Nothing but a Limit of Quantum Mechanics?

A claim commonly found in textbooks is that quantum mechanics reduces to classical mechanics, in the $\hbar \to 0$ limit or, more precisely, in the $\hbar/A \to 0$ limit, where A is the action of the system. As an example, one may recall Planck's formula for the energy of blackbody radiation at frequency ν:

$$U(\nu) = \frac{8\pi V}{c^3} \frac{h\nu^3}{e^{h\nu/kT} - 1} \tag{6.1}$$

where V, c, k, and T are, respectively, the volume of the cavity, the speed of light, Boltzmann's constant, and the temperature. In the $h\nu \ll kT$ limit, one easily obtains

$$U(\nu) = \frac{8\pi V}{c^3} kT\nu^2, \tag{6.2}$$

the classical Rayleigh-Jeans formula which does not contain h. The above example, although correct, does not imply that the $h \to 0$ limit recovers the classical regime for all physical systems.

Another connection often invoked between classical and quantum mechanics is Ehrenfest's theorem, which states that the mean values of coordinates and momenta of a quantum system evolve according to classical mechanics, under certain conditions.

Let us briefly discuss this point, considering a particle of mass m moving in one dimension, under the action of a potential $V(x)$. In the Heisenberg picture one has the following equations for the position and momentum operators:

$$\frac{d\hat{q}}{dt} = \frac{\hat{p}}{m}, \frac{d\hat{p}}{dt} = F(\hat{q}) \tag{6.3}$$

where $F(x) = -dV(x)/dx$. Taking the average of the previous equation, one gets

$$\frac{d\langle\hat{q}\rangle}{dt} = \frac{\langle\hat{p}\rangle}{m}, \frac{d\langle\hat{p}\rangle}{dt} = \langle F(\hat{q})\rangle. \tag{6.4}$$

So we have the "correct" Ehrenfest theorem: *given a quantum mechanical wave function the evolution equation of the correspondent classical system is obtained replacing the position, the momentum and force acting on the particle by the expected values of these quantities*.

One wonders, however, whether $\langle F(\hat{q})\rangle$ is correctly approximated by $F(\langle\hat{q}\rangle)$, in the classical limit, which is the case if the Wigner function[4] is sharply peaked around $(\langle x\rangle, \langle p\rangle)$. Therefore, one may start in a (semi)classical situation, i.e. with a Wigner function initially concentrated on $(x(0) = \langle x\rangle, p(0) = \langle p\rangle)$, so that

$$\langle F(\hat{q})\rangle \simeq F(\langle\hat{q}\rangle) \tag{6.5}$$

for a certain amount of time, during which classical mechanics is substantially correct.

At larger times, say after a certain t_C, the quantum nature of the system will prevail, and one would like to know the dependence of t_C on $\hbar$ and on the parameters of the classical Hamiltonian. This can be obtained considering an initial state whose Wigner function is smooth and localised on scales much larger than $\hbar$. Denote by Δp_0 and Δq_0, respectively, the initial widths of the momentum and position distributions (e.g. a Gaussian packet), and let $\Delta q_0 \Delta p_0 = A_0 \gg \hbar$. Initially, the quantum effects are small, hence the evolution of the averages is ruled by classical mechanics.

If the classical system is chaotic, the initially smooth packet will be exponentially stretched in the directions with positive Lyapunov exponents. Since the phase space volume must be preserved (Liouville theorem) the Wigner function will develop

[4] Just for notation simplicity we write the Wigner function $W(q, p, t)$ for one-dimensional systems:

$$W(q, p, t) = \frac{1}{h}\int dq' \psi^*(q + q', t)\psi(q - q', t)e^{ipq'/\hbar}$$

where $\psi(q, t)$ is the wave function which obeys to the Schrödinger equation. By integrating $W(q, p, t)$ over either q or p, one obtains the probability density for the other variable. Note that $W(q, p, t)$ is not positive definite, sometimes it is called a "quasi-probability" distribution, however the expectation value of an operator $G(\hat{q}, \hat{p})$ can be written as

$$\langle\hat{G}\rangle = \int dq\, dp G(q, p)W(q, p, t).$$

dendritic structures at small scales, which decrease exponentially in time as $\varepsilon_0 e^{-\lambda t}$ where λ is the Lyapunov exponent and ε_0 is $O(\Delta p_0)$ (or $O(\Delta q_0)$). When the smallest scale reaches order $O(\hbar^{1/2})$, i.e. $\Delta q_0 \Delta p_0 e^{-2\lambda t} \sim \hbar$, the quantum effects become relevant. This leads to the following estimate for the crossover time t_C:

$$t_C \sim \frac{1}{\lambda} \ln \frac{A_0}{\hbar}. \tag{6.6}$$

Differently, non-chaotic systems lead to the power law dependence

$$t_C \sim \tau \left(\frac{A_0}{\hbar} \right)^{\alpha}, \tag{6.7}$$

where τ is the characteristic time of the classical system and α depends on the details of the dynamics (Berry 2001).

As $A_0/\hbar \to \infty$, the crossover time t_C diverges in both Eqs. (6.6 and 6.7), and one may be tempted to conclude that the classical behaviour is then a clear consequence of quantum mechanics. On the other hand, in the case of chaotic classical systems, the divergence is so slow that t_C remains definitely small, in practice.

6.1.2 *Quantum Mechanics, Classical Chaos and Planetary Dynamics*

The severe limit of the prediction Eq. (6.6) is well illustrated by the chaotic tumbling of a small satellite of Saturn, Hyperion. This minor celestial body has an irregular, potato-like shape which extends for an order $O(10^2)$ km. The irregular (chaotic) motion of Hyperion, due to the interaction with Saturn and the big moon Titan, was quite clearly observed during the Voyager 2 mission.

The classical instability time (the inverse of the Lyapunov exponent) is $\lambda^{-1} \sim 10^2$ days, while the order of magnitude of the classical action is $A \sim 10^{58}\hbar$, a quantity that sounds enormous. However if we use the above numerical values for A and λ the estimate Eq. (6.6) gives

$$t_C \simeq 37 \text{ years} \tag{6.8}$$

a time interval which is really very small. Nobody seriously thinks that in a few or many decades, astronomers will see quantum effects take over in Hyperion, and its classical chaotic behaviour turn in some quantum quasiperiodic motion Berry (2001). What is the solution of this paradox?

The answer lies in the so-called decoherence effect: Hyperion does not interact with Saturn and Titan only, but also with a myriad of other objects, such as the other moons of Saturn, cosmic dust, photons from the Sun and so on. Therefore, the (naively) expected quantum suppression of classical chaos, which leads to Eq. (6.8),

is contrasted by other quantum effects, making the classical behaviour persist over time scales exceedingly larger than the value estimated by Eq. (6.8).

The proper way to take these effects into account would be to enlarge the system of interest to include the Sun's radiation field and all other relevant interactions as well. Then, formally, one would have to eliminate the degrees of freedom which do not pertain to Hyperion, by tracing over them.

In order to have an estimate of the effects of decoherence, it is enough to treat the environment as a random disturbance. Noting that the quantum suppression of classical chaos involves the interference of waves and their phases it is necessary to understand how the external environment changes the phase of the wave function. Simple considerations show that after a time $t_D \ll t_C$ one has decoherence of the phase.

The argument, due to Zurek (1991, 2003), is the following. The interactions of the system with the environment change the evolution equation for the Wigner function; this contribution can be fairly reflected by a term of the form $D\partial^2 W/\partial p^2$, where the diffusion coefficient D is given by the Einstein formula $D = 2m\gamma k_B T$, $\gamma^{-1} = \tau_R$ being the typical relaxation time of the classical system.

Since the interference term is modulated by functions of the form $\cos(p\Delta x/\hbar)$ where Δx is the typical length of the initial wave function, the decoherence time t_D is estimated as the time at which, starting with a Wigner function of the form $\cos(p\Delta x/\hbar)$, the term $D\partial^2 W/\partial p^2$ is relevant, therefore

$$\frac{1}{t_D} \sim D\frac{\Delta x^2}{\hbar^2} = \frac{2mkT}{\tau_R \hbar^2}(\Delta x)^2 .$$

For a system with $\Delta x = 1$ cm, $m = 1$ g and $T = 300$ k, one has $t_D/\tau_R \sim 10^{-40}$. Therefore even if τ_R is very large, say the age of the universe ($\sim 10^{17}$ s), the quantum coherence would be destroyed in only $\sim 10^{-23}$ s.

The above amounts to the notion that including the environment, in quantum dynamics, produces a smoothing effect that leads to classical behaviour. Indeed, Hyperion behaves as a classical object over time intervals which are well beyond 37 years, not as a mere consequence of the fact that $A_0/\hbar \sim 10^{58} \gg 1$, but because of the emergent semiclassical phenomenon, in which the interactions with the external environment play a fundamental role.

We conclude this brief discussion with a quotation from Dirac (1958), who could not be knowledgeable about chaos but had precise ideas on quantum and classical mechanics as open interesting theories:

> We should ... expect to find that important concepts in classical mechanics correspond to important concepts in quantum mechanics, and, from an understanding of the general nature of the analogy between classical and quantum mechanics, we may hope to get laws and theories in quantum mechanics appearing as simple generalisations of well-known results in classical mechanics.

Dirac had been prophetic: the example of Hyperion and the role of decoherence show how results from classical mechanics can be used to guide further development

on the range of validity of quantum dynamics. In addition we can mention the cross-fertilisation of methods and ideas between mathematics, chemistry and physics, e.g. the perfect parallel between the equations governing hydrogen ionisation in crossed electric and magnetic field and celestial mechanics (Porter and Cvitanovic 2005), and the semiclassical computations to extract information in classical phase space structure (Ezra 1998).

6.1.3 An Interlude: Discrete Versus Continuous Descriptions and the Semiclassical Limit

We conclude this section with a brief discussion of the coarse-grained description of systems with a finite number of states: such an issue is relevant to certain aspects of the semiclassical limit and decoherence (Mantica 2000; Falcioni et al. 2003).

The first observation is that chaos, in the sense of positive Lyapunov exponents or positive Kolmogorov-Sinai (KS) entropy,[5] h_{KS}, can be realised only in systems with bounded and continuous state space. Indeed, the trajectories of a deterministic system with a number M of states are asymptotically periodic, and characterised by vanishing h_{KS}. The reason is that it takes M steps at most to return to a state that has been previously visited. At that point, the determinism of the evolution rule implies that the evolution repeats itself periodically. The period T cannot be larger than M, but in typical applications of physical interest one has $T \sim \sqrt{M}$ (Coste and Hénon 1986).

If the system at hand has a continuous state space, i.e. it has uncountably many states, one could try to approximate its dynamics by means of systems with a large but finite number M of states, and recover the original dynamics in the $M \to \infty$ limit. This issue is closely related to the semiclassical limit, as quantum dynamics is necessarily periodic (or quasiperiodic) in a finite state space, but it is believed to yield classical mechanics, which has a continuous state space, in the $\hbar \to 0$ limit. We have already argued that such a limit is not as trivial as it may seem, because it is, in fact, a singular limit.

Let us now consider a deterministic map

$$x(t+1) = f(x(t)), \tag{6.9}$$

with continuous state space $[0, 1]$, and with discrete time t. Introduce a spatially discrete approximation of the map, defining:

$$m(t+1) = F_M(m(t)), \tag{6.10}$$

[5] These are the exponential rates at which nearby trajctories separate, and at which information on the initial state is lost (which are strictly related concepts).

where $F_M(m(t)) = [Mf(m(t)/M)]$, and $[\cdot]$ indicates the integer part of a real number. This dynamics now concerns the integer states $\{0, 1, ..., M-1\}$, which represent the original ones with accuracy $\eta = 1/M$. Whether the dynamics of f are chaotic or not, those of F_M are periodic, apparently at variance with the naive idea that the $M \to \infty$ limit, i.e. the $\eta \to 0$ limit recovers the continuous dynamics. This fact is similar to the classical limit of quantum mechanics: because quantum mechanics is governed by a linear evolution law, the Schrödinger equation, chaos cannot exist in quantum mechanics, while it is common in (non-linear) classical mechanics. This, however, should be recovered from quantum mechanics in the $\hbar \to 0$ limit.

It is important to note that although classical mechanics is typically described by nonlinear equations, it is formally analogous to quantum mechanics in many respects. Indeed, the Liouville equation of classical mechanics affords a linear theory for the evolution of probabilities, at the cost of switching from a finite dimensional phase space to an infinitely dimensional function space, analogously to the description based on the Schödinger equation. We stress that even systems with linear evolution laws, such as the Schödinger equation, can exhibit nontrivial behaviours and complex features.[6]

This apparent contradiction has led to many different conclusions, see e.g. Ford et al. (1991), but can be explained in terms of the accuracy with which both continuous and discrete dynamics are observed—the *coarse-graining* level—and thanks to the introduction of randomness in the quantum mechanical description.

If the accuracy ε is much larger than the lattice spacing of the discrete dynamics ($\varepsilon \gg \eta$), the discrete and the continuous dynamics are practically indistinguishable for a time that grows logarithmically with the period T. Analogously, the semiclassical limit applies within a time that grows as $\hbar$ decreases: the time that a wave-packet takes to spread over a large distance. Therefore, the quantity $\eta = 1/M$ in the $M \to \infty$ limit, the limit of infinitely many states, plays the same role played by $\hbar$ in the $\hbar \to 0$ limit.

Introduce now some randomess in the spatially discrete approximation of the spatially continuous dynamical system. Consider, for instance, the probabilistic automaton:

$$m(t+1) = F_M(m(t)) + \sigma(t)\,, \tag{6.11}$$

with independent random variables $\sigma(t)$. If the randomness is insufficient, the dynamics of Eq. (6.11) does not accurately reproduce the complexity of the continuous dynamics. But the chaotic behaviour may be restored, and one obtains features

[6] For instance, if we consider the passive advection of a scalar field $\theta(\mathbf{x}, t)$ in a given velocity field $\mathbf{u}(\mathbf{x}, t)$

$$\partial_t \theta + \mathbf{u} \cdot \nabla\theta = D\Delta\theta,$$

it is easy to show that an initial uncertainty $\delta\theta(\mathbf{x}, 0)$ cannot grow *forever* in time. On the other hand, the solution of this equation can be nontrivial (e.g. the spatial gradient of $\theta(\mathbf{x}, t)$ can grow quickly) if the equation

$$\frac{d\mathbf{x}}{dt} = \mathbf{u}(\mathbf{x}, t)$$

is chaotic (Crisanti et al. 1994; Berry 1992).

practically indistinguishable from those of the original chaotic dynamical system, if $\sigma(t)$ is sufficiently random (Falcioni et al. 2003).

This result shows that the decoherence required for quantum mechanics to lead to the classical macroscopic behaviour has a simple origin, which is completely independent of the interpretations of quantum mechanics: classical mechanics can be seen as a sort of emergent property of quantum mechanics where the interactions with the environment are modelled using some random variable.

6.2 Chemistry Is Not Just Applied Quantum Mechanics

If Dirac's claim (or dream?) is correct, chemistry does not exist as an autonomous science; it is just a branch of applied quantum mechanics. The only difficulty in recognising this should be of a technical nature and related either to the difficulties encountered in solving the Schrödinger equation, or in finding sufficiently realistic approximations:

> It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation (Dirac 1929).

The simple-minded reductionist finds no reason to dispute Dirac's dream to suppress the pluralism of chemical theories and to derive (at least in principle) all chemical phenomena from the unifying principles of quantum mechanics. To uphold this point of view, the work of L. Pauling is often referred to, because Pauling contributed seminal works to modern chemistry, through an ingenious use of quantum mechanics. However a close analysis of his papers and books shows that this conclusion is quite superficial. As a matter of fact, the views of this great American scientist, on the relation between quantum mechanics and chemistry, were not very solid and changed over time. In the late 1920s, his opinion was the following:

> With the development of the quantum mechanics it has become evident that the factors mainly responsible for chemical valence are the Pauli exclusion principle and the Heisenbeg-Dirac resonance phenomenon. (Pauling 1928)

In spite of the above strong claim, Pauling produced a clever mix of chemical intuition and quantum mechanics, hardly reducible to mere quantum mechanical calculations. For instance, in his celebrated series of papers in the 1930s on the quantum mechanical theory of chemical bonds in polyatomic molecules, Pauling gave a list of rules, some of which were based on quantum mechanics and others suggested only by phenomenological arguments, based on his prior knowledge of chemistry (Bouguerra 2002). Some years later, he eventually came to emphasise the role of chemistry as an autonomous science:

> The theory of resonance is a part of the chemical structure theory which has an essential empirical (inductive) basis; it is not just a branch of quantum mechanics. (Pauling 1956)

This opininion is shared by other pioneers of quantum chemistry, like Coulson (1960), who stated that:

> It is not unfair to say that ... in practically the whole of theoretical chemistry, the form in which mathematics is cast is suggested, almost inevitably, by experimental results. This is not surprising when we recognise how impossible is any exact solution of the wave equation for a molecule. Our approximations to an exact solution ought to reflect the ideas, intuitions and conclusions of experimental chemistry.

There is no unanimous consensus on this delicate topic. For instance, Slater disagrees with Pauling and Coulson, and writes:

> I felt here we had a fundamental theory and a challenge to be able to explain everything we see around us in terms of that fundamental theory. So I have already discarded anything that didn't fit into that.[7]

We will see that the connection between chemistry and quantum mechanics is not so unquestionable, as is sometimes claimed, and that the difficulties in reducing one to the other, far from being merely technical, involve serious conceptual issues. To understand that, one should avoid confusion between the usefulness of computational quantum mechanics and a complete theory, from first principles. Numerical computations based on the Schrödinger equation are definitely important, but cannot replace chemical insight.

6.2.1 Quantum Mechanics Does Not Explain Chemistry

The shape of molecules holds a central role in chemistry. The shape of a diatomic molecule is rather simple, and determined by the distance between the two atoms. For a triatomic molecule, one has a triangle, while in general the molecular shape is determined by the length and orientation of the chemical bonds. These concepts have a distinct classical nature and, as admitted even by Heisenberg, they are inconsistent with the fundamental principles of quantum mechanics (Primas 1981; Scerri 2008).

Let us briefly sketch the typical procedure used to study molecules in the framework of quantum mechanics. In a first principle approach to the chemistry of a given molecule, one must consider the Schrödinger equation of the N_1 nuclei and N_2 electrons interacting via the Coulomb potential. This makes the computation of the wavefunction of a typical molecule prohibitive. For instance the benzene molecule consists of 12 nuclei and 42 electrons, so the Schrödinger equation is a partial differential equation in 162 variables (the spatial coordinates of the electrons and nuclei), which is clearly beyond the capability of any existing computer.

The assumption which makes the calculation possible is known as the Born-Oppenheimer approximation of the wave-functions, which was proposed in the early days of quantum mechanics and is still indispensable in quantum chemistry Thijssen (2007).

Denote by $\mathbf{r}$ and $\mathbf{R}$ the electronic coordinates and the nuclear coordinates respectively, the global wave function $\psi_T(\mathbf{r}, \mathbf{R})$ is factorised as the product of the electronic and nuclear wave-functions

[7] Cited by Schweber (1990).

$$\psi_T(\mathbf{r}, \mathbf{R}) = \psi_E(\mathbf{r}, \mathbf{R}) \times \psi_N(\mathbf{R}) \tag{6.12}$$

In the first step the nuclei are fixed in a certain configuration $\mathbf{R}$ and the Schrödinger equation for the electrons is solved, yielding the wavefunction ψ_E.

The electronic features (e.g. energy levels) thus computed become functions of the coordinates of the nuclei. Varying these positions and solving the electronic Schrödinger equation for each $\mathbf{R}$, ψ_E may be used to construct an effective potential for the interactions of the nuclei $V_{\text{eff}}(\mathbf{R})$. Once $V_{\text{eff}}(\mathbf{R})$ is known, the second step requires the solution of the Schrödinger equation for the nuclear motion.

Usually, for practical purposes, it suffices to study the nuclear motion in the harmonic limit:

$$V_{\text{eff}}(\mathbf{R}) \simeq V_{\text{eff}}(\mathbf{R}^*) + \frac{1}{2} \sum_{i,j} A_{i,j}(R_i - R_i^*)(R_j - R_y^*) \tag{6.13}$$

where $\mathbf{R}^*$ is determined by the minimum of $V_{\text{eff}}(\mathbf{R})$ and $A_{i,j} = \partial^2 V_{\text{eff}} / \partial R_i^* \partial R_j^*$. The result is the total energy of the molecule, including contributions from electrons, nuclear vibrations, and rotations as well as translations of the molecule. The justification of the Born-Oppenheimer approximation lies in the large ratio of nuclear to electronic masses, which is of order $O(10^3)$.

In the framework of the Born-Oppenheimer approximation, the shape of the molecules is determined by the minimum of the effective potential $V_{\text{eff}}(\mathbf{R})$. Since at room temperature the first excited levels, both rotational and vibrational, are basically empty, one can consider the nuclear structure as a rigid system; so it is enough to study the electronic properties for $\mathbf{R} = \mathbf{R}^*$.

In spite of its evident practical merits, the Born-Oppenheimer approximation presents various criticalities, from a conceptual point of view. Indeed, the principles of physics, and of quantum mechanics in particular, imply that:

(a) nuclei cannot have fixed positions; and
(b) nuclei and electrons interact.

While point (a) may be neglected in first-order approximations, as is commonly done in the study of high temperature systems, point (b) is much harder to dispense with.

6.2.2 Why Does the Born-Oppenheimer Approximation Fail to Predict the Shape of Molecules?

The fact is that the Born-Oppenheimer approximation decouples the nuclear motions form the electronic ones, but this is conceptually inconsistent with the holistic character of quantum mechanics, which is well illustrated be the experimentally confirmed EPR paradox (Bell 2004).

Such a paradox refers to an apparently paradoxical property of quantum mechanics. It was introduced as a thought experiment, to reveal the contradictions of the Copenhagen interpretation of quantum mechanics.

Consider two systems, A and B (e.g. two free particles), whose wave-functions are known. If A and B interact for a short period of time, one can determine the wave-function resulting from this interaction, by solving the Schrödinger equation. Assume that A and B are subsequently moved far apart, so that they do not interact anymore.

Einstein, Podolsky and Rosen posed the following question: what happens if one makes a measurement of system A? What happens if one measures its momentum? From the conservation of momentum and our knowledge of the system before the interaction, one can then infer the momentum of system B. Thus, by making a momentum measurement of A, one can also determine the momentum of B. Recall now that A and B are spacelike separated, and cannot communicate in any way. This separation means that B must have had the inferred value of momentum not only after, but before the measurement as well. If, on the other hand, the measurement at A had somehow caused B to enter into a particular momentum state, then A ought to have acted upon B, so as to immediately inform B that a measurement had taken place, in spite of the distance which prevents any instantaneous communication between the two systems.

If one examines the wave-function at the moment just before the measurement at A is made, one finds that there is no certainty as to the momentum of B because the combined system is in a superposition of multiple momentum eigenstates of A and B. So, even though system B must be in a definite state before the measurement at A takes place, the wave-function description of this system cannot determine its momentum. Therefore, since system B has a definite momentum which quantum mechanics cannot predict, it seems that quantum mechanics is inconsistent or, more precisely, it is not a "complete theory".

Today, after the performance of accurate experiments, such as those of A. Aspect, we may conclude that the EPR "paradox" shows how deeply quantum mechanics violates classical intuition.

At odds with Einstein's intentions, the EPR paradox reveals a characteristic quantum mechanical phenomenon, known as entanglement: the fact that a quantum system made of various parts is, in reality, a whole which cannot be split into its elementary components. At times, this is expressed by saying that measurements performed on spatially separated parts of a quantum system instantaneously affect one another, suggesting that quantum mechanics violates the prescriptions of relativity. This is incorrect; one should rather say that measurements reveal the pre-existing entangled nature of a quantum system, giving rise to a kind of nonlocal effect that, however, cannot be used for any superluminal transmission of information.

The equation obtained from the Born-Oppenheimer approach is not a mere approximation of quantum mechanics, which is accurate whenever m_n/m_e is sufficiently large, where m_n and m_e are the masses of the nuclei and of the electrons, respectively. On the contrary, it has a mixed status, being partially quantum and partially classical, or semiclassical, in nature. In particular, the Born Oppenheimer approach

cannot be exactly valid in any limit; treating the nuclear and electronic motions as independent of each other breaks the quantum entanglement of nuclei and electrons. Such a decoupling can never be valid, not even in the $m_n/m_e \to \infty$ limit.

The fact is that the Born-Oppenheimer approximation describes a quantum mechanical system: the electrons anchored on a classical rigid structure, the frame of the nuclei. This method produces highly accurate solutions; however some aspects of it do not belong to quantum mechanics. Therefore, one may only say that molecular structures are justified, but not explained, in terms of an "approximate solution" of the Scrödinger equation, via the Born-Oppenheimer approximation.

An example of the difficulty in obtaining chemical structures directly from quantum mechanics is afforded by the case of the C_3H_4 molecule, which takes different shapes, called allene, cyclopropene and methylacetylene. The three molecules above have the same number of electrons and nuclei, so they are described by the same Hamiltonian. However the structures of the three isomers are very different and characterised by different classical configurations (the positions of the nuclei) that are missing in the quantum description.

From a naive point of view one could think that the different isomeric structures are nothing but different solutions of the Born-Oppenheimer approximation, i.e. different minima of the effective potential. But such a scenario meets rather serious problems: in the case that $V_{\text{eff}}(\mathbf{R})$ has multiple minima, quantum mechanics implies delocalised structures. In other words the quantum mechanics of an isolated molecule does not predict the existence of allene, cyclopropene and methylacetylene but a unique C_3H_4 molecule obtained as a superposition of the three isomers. Therefore, the concept of molecular structure seems to be in strong disagreement with the basis of quantum mechanics.

6.2.3 A Look at a Specific Problem: The Pyramidal Molecules

From experimental evidence, we know that there exist pyramidal molecules of the type XH_3 where X can be, for instance, N (nitrogen), P (phosphorus) or As (arsenic), in which cases we have ammonia (NH_3), phosphine (PN_3) or arsine (AsH_3). A scketch of this kind of molecules is shown in Fig. 6.1: three hydrogen atoms lie in a plane (say $x = 0$) and the fourth atom is on the top of the pyramid.

Because of the symmetries of the problem, the fourth atom can either sit over the plane $x = 0$ (as in Fig. 6.1) or in the corresponding symmetric configuration, below that plane. Since the two configurations have the same energy, quantum mechanics implies that the pyramidal molecules appear delocalised, as in Fig. 6.2.

This is plainly incompatible with chemistry, which finds arsine to be localised, and, in any event, definitely different to the delocalised structure.

How can we explain this apparent inconsistency between quantum mechanics and chemistry? A nice paper by Claverie and Jona-Lasinio (1986) shows, once more, that the environment plays a fundamental role in the emergence of such classical behaviour as the localisation of ammonia, phosphine and arsine. The essence of the

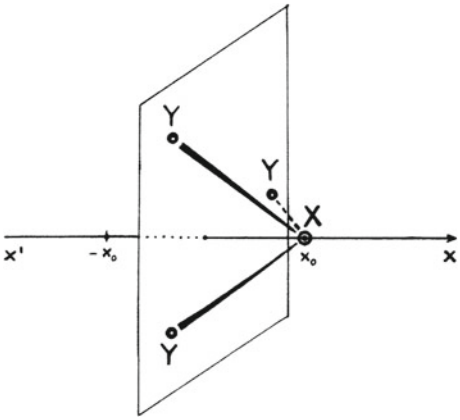

Fig. 6.1 A pyramidal molecule XY_3 with its inversion axis x' and the two nuclear equilibrium configurations $-x_0$ and x_0

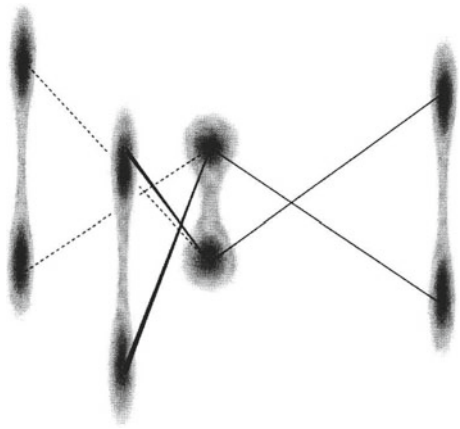

Fig. 6.2 Ground state of a pyramidal molecule (according to quantum mechanics description of an isolated molecule)

argument is the following. Molecules of a given chemical element are not isolated; they interact with the rest of the universe and, in the first place, with each other. If one takes a single isolated molecule XH_3, in the semiclassical limit, the position of its fourth atom is described by the double-well potential, shown in Fig. 6.3, whose two minima separately correspond to the position of the fourth atom in Fig. 6.1. If the molecule is not isolated, the potential resulting from the interaction of the fourth atom with the other three hydrogen atoms and with the rest of the universe is not perfectly symmetric.

This is one example of the symmetry breaking phenomenon, which is common in physics. For instance, below a certain temperature, the magnetisation of certain objects does not vanish, hence its magnetic moments have a preferential orientation, despite the symmetry of the Hamiltonian, even in the absence of an external

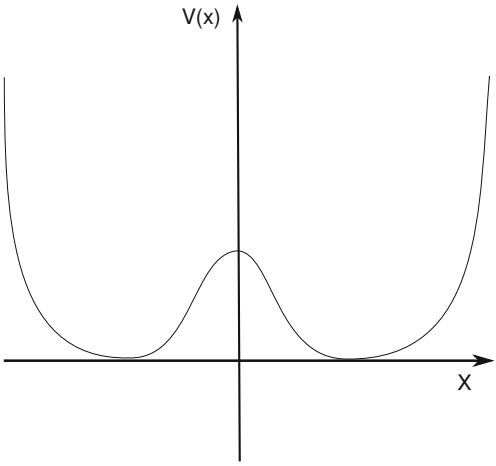

Fig. 6.3 The one-dimensional double-well potential

magnetic field.[8] It is important to stress that the phenomenon of sharp symmetry-breaking, deduced from the mathematical arguments, can only happen in systems with an infinite number of components; but macroscopic objects are made of so many microscopic elements, that they closely approach the mathematical prediction.

As a consequence, any small perturbation "selects" one of the two configurations localising the fourth atom, due to the external environment. Of course, the strength of this symmetry-breaking perturbation depends on the details of the systems. The localisation observed of arsine molecules is thus a cooperative effect, resulting from the interactions with the environment.

Let us emphasise that some aspects of the pyramidal molecule are not peculiar to the specific cases discussed above, but are typical of large objects whose spectrum is almost degenerate near the ground state (e.g. molecules of the order of 100 Daltons). Then, because of the very small difference in the energy levels, one cannot neglect the perturbations due to the environment, even if they are small, and the state realised in nature does not correspond to an eigenstate of the isolated system: its symmetry is broken.

In our opinion the result by Claverie and Jona-Lasinio, completed by Jona-Lasinio (2010) in terms of a fully quantum mechanical treatment, supports the point of view of Woolley (1978, 1986), who considers the interaction of individual molecules with the environment as instrumental in the generation of chemical structures.

[8] In a semi-serious fashion, we may claim that the symmetry breaking is somehow a solution of Buridan's ass paradox. In a (large) magnetic system, the spins "choose" a preferred orientation, and, in an analogous way, the ass chooses one of the two haystacks, if they are large enough (formally infinite).

6.2.4 Beyond the Born-Oppenheimer Method

Perhaps some readers have the impression that the Born-Oppenheimer approximation is an old-fashioned subject of marginal relevance to modern research and numerical computations of large quantum systems. There have indeed been important advances in the last decades, especially those related to density functional theory and to the (Car and Parrinello 1985) method, which have proven highly successful in the so-called *ab initio* calculations of the structure of matter. However, leaving aside the technical details, we observe that these methods belong to the same class of the Born-Oppenheimer approach, from a conceptual standpoint.

In density functional theory (DFT) (Jones and Gunnarsson 1989), the nuclei of the system of interest, e.g. molecules or whole solids, are seen as fixed in positions $\mathbf{R}$, and generate a static potential, $V(\mathbf{R})$ that determines the dynamics of the electrons. DFT does not consider the complete wave-function $\psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$ of the N electrons, which is a function in $\boldsymbol{R}^{3N}$; rather, it develops a description of the one-body wave-function in $\boldsymbol{R}^3$, which depends in a self consistent way on the electronic density $n(\mathbf{r})$ and confers a statistical character to the calculations.[9]

Let $\psi_k(\mathbf{r})$ be the wave function of the k-th electron; it depends on the electron density $n(\mathbf{r}) = \sum_j |\psi_j(\mathbf{r})|^2$, determined by the state of the N electrons. Then, one has to solve the one-body Schödinger equation,

$$\hat{H}(\{\mathbf{R}\}, n(\mathbf{r}))\psi_k = e_k\psi_k \tag{6.14}$$

where the Hamiltonian operator $\hat{H}(\{\mathbf{R}\}, n(\mathbf{r}))$ depends on the positions of the nuclei $\mathbf{R}$ and on the electron density $n(\mathbf{r})$.

Equation (6.14) is a set of nonlinear equations whose solution is not easily determined. Although there are some rigorous results,[10] approximations are required in order to overcome the mathematical difficulties.

The conclusion is that *ab initio* calculations, i.e. from first principles, can be performed only in a few cases and under controlled approximations. On the other hand, as explicitly noted by Pople (1999), *the principal limitation of DFT models is that there is no clear route for convergence of methods to the correct answer* and a considerable amount of empirical parametrisations, deduced from experimental data and physical intuition, must be used.

The situation is similar with the powerful method introduced by Car and Parrinello, which combines classical molecular dynamics, for the nuclei, and DFT for the electrons. This approach has its physical justification in the time scale separation between nuclei and electrons, and it is, at the conceptual level, similar to the Born-Oppenheimer approximation.

---

[9] For some aspects, DFT is similar to the Boltzmann equation, which describes the one-particle distribution.

[10] For instance, the Hohenberg-Kohn theorem states that the properties of the electronic ground state depend only on the density $n(\mathbf{r})$ Jones and Gunnarsson (1989). But the determination of $n(\mathbf{r})$ remains exceedingly difficult.

6.3 Summary and Conclusions

Let us conclude with a brief summary of the main aspects of the classical limit of quantum mechanics and the connection between chemistry and quantum mechanics.

There exists a rather deep and non-hierarchical relation between classical and quantum mechanics:

(a) quantum mechanics cannot be formulated without classical mechanics: in Heisenberg's words *"Newtonian mechanics is a kind of a priori for quantum theory"*;
(b) classical mechanics is not just the $\hbar \to 0$ limiting case of quantum mechanics, and the actual consequences of the Ehrenfest correspondence principle are not obvious;
(c) for a chaotic classical system, the proper classical limit can be obtained from quantum mechanics only taking into account the influence of the outer environment;
(d) an analysis of systems with discrete states suggests that, once a coarse-graining procedure and randomness are introduced, classical mechanics can be seen as a sort of emergent property of quantum mechanics.

Despite Dirac's claim and the availability of the advancements made possible by powerful modern supercomputers and the introduction of powerful numerical methods (DFT and Car-Parrinello), quantum chemistry has been largely created by chemists rather than physicists; it is basically a subdiscipline of chemistry and not merely applied quantum mechanics. Chemistry is mainly the result of intuition and imagination, and not just empirical facts and quantum mechanical computations.

Molecular structures seem to be justified as approximate solutions of the Schrödinger equation, via the Born-Oppenheimer approximation, but this is a vicious circle: the results obtained from the Born-Oppenheimer approximation are accurate precisely because one introduces a number of subtle assumptions that cannot be obtained as limiting cases from quantum mechanics:

(a) classical variables for the positions of the nuclei;
(b) disentangling of electronic and nuclear motion; and
(c) symmetry breaking if $V_{\text{eff}}(\mathbf{R})$ has more than one minimum.

Therefore molecular structures are not fully determined by quantum mechanics: it is more correct to say that quantum mechanics + (semi)classical physics is able to describe the observable features of molecules. We stress that the word "describe" is much more appropriate than "explain".

In our opinion the relation between quantum mechanics and chemistry is much weaker than the link between Newtonian mechanics and astronomy. For instance, in the latter case the theory was able to predict the existence of a previously unknown planet (Neptune). On the contrary, as far as we know, there is nothing similar in chemistry regarding the prediction of a new element solely on the basis of quantum mechanics.

Such a limitation of the predictive power of quantum mechanics for chemistry is quite evident for the structural isomers, i.e. molucules with the same atoms but

different molecular structures. We can mention dimethyl ether and ethyl alcohol, which both have the same compositional formula C_2H_6O, and are both described by the same Hamiltonian; but they are represented by two distinct structural formulas: $(CH_3)_2O$ and C_2H_5OH. The two isomers differ radically in physical properties such as melting points and boiling points as well as patterns of chemical reactivity. Ethanol is extremely soluble in water, whereas dimethyl ether is only partially soluble in water. Ethanol boils at 78.4° C, while dimethyl ether boils at 34.6° C. Drinking ethanol leads to intoxication, while drinking dimethyl ether has no such effect.

On the other hand, as is clear in the case of pyramidal molecules, the (semi)classical limit of the molecules does not follow in a straightforward way from quantum mechanics but is a consequence of the interaction of single molecules with an external environment consisting of a very large number of components.

It is remarkable that the emergence of molecular structures and the semiclassical behaviour in the presence of chaos have a common origin: in both cases, the basic ingredients are the presence of large objects and their interaction with an environment containing a large number of microscopic constituents.

References

Bell, J.S.: Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy. Cambridge University Press, Cambridge (2004)

Berry, M.V.: Chaos and the Semiclassical Limit of Quantum Mechanics is the Moon There When Somebody Looks? In: Russell, R.J., Clayton, P., Wegter-McNelly, K., Polkinghorne, J. (eds.) Quantum Mechanics: Scientific Perspectives on Divine Action, pp. 41. Vatican Observatory CTNS Publications, (2001)

Berry, M.V.: Classical limits. In: Gross, D., Henneaux, M., Sevrin, A. (eds.) The Theory of the Quantum World Proceedings of the 25th Solvay Conference on Physics, pp. 52 (2013)

Berry, M.V.: True Quantum Chaos? an Instructive Example. In: Abe, Y., Horiuchi, H., Matsuyanagi, K. (eds), New Trends in Nuclear Collective Dynamics, pp. 183. Springer (1992)

Berry, M.V.: Alisa Bokulich, Reexamining the quantum-classical relation: beyond reductionism and pluralism. Br. J. Philos. Sci. **61**, 889 (2010)

Bohr, N.: On the Application of the Quantum Theory to Atomic Structure. Cambridge University Press, Cambridge (1924)

Bohr, N.: Atomic Theory and the Description of Nature. Cambridge University Press, Cambridge (1929)

Bokulich, A.: Reexamining the Quantum-Classical Relation. Cambridge University Press, Cambridge (2008)

Bouguerra, M.L.: Linus Pauling: Génie Rebelle et Humaniste. Belin, Paris (2002)

Car, R., Parrinello, M.: Unified approach for molecular dynamics and density-functional theory. Phys. Rev. Lett. **55**(22), 2471 (1985)

Claverie, P., Jona-Lasinio, G.: Instability of tunneling and the concept of molecular structure in quantum mechanics. Phy. Rev. A **33**(4), 2245 (1986)

Coste, J., Hénon, M.: Invariant cycles in the random mapping of *n* integers onto themselves. Comparison with Kauffman binary network. In Bienestock, et al., (eds.), Disordered Systems and Biological Organization, pp. 361. Springer, Berlin, (1986)

Coulson, C.A.: Valence. Oxford University Press, Oxford (1960)

Crisanti, A., Falcioni, M., Mantica, G., Vulpiani, A.: Applying algorithmic complexity to define chaos in the motion of complex systems. Phys. Rev. E **50**(3), 1959 (1994)
Dirac, P.A.M.: Oral interviews with P.A.M. Dirac by Thomas S. Kuhn. In Archive for the History of Quantum Physics. Deposit at Harvard University, Cambridge (1962)
Dirac P.A.M.: The relation of classical to quantum mechanics. In: Proceedings of the Second Canadian mathematical Congress, Vancouver, pp. 10. University of Toronto (1951)
Dirac, P.A.M.: Quantum mechanics of many-electron systems. Proc. Roy. Soc. London. Ser. A **123**(792), 714 (1929)
Dirac, P.A.M.: The Principles of Quantum Mechanics. Clarendon Press Oxford, Oxford (1958)
Ezra, G.S.: Classical-quantum correspondence and the analysis of highly excited states: Periodic orbits, rational tori, and beyond. In Hase, W.L. (ed.) Advances in Classical Trajectory Methods, vol. 3, pp. 35. JAI Press (1998)
Falcioni, M., Vulpiani, A., Mantica, G., Pigolotti, S.: Coarse-grained probabilistic automata mimicking chaotic systems. Phys. Rev. Lett. **91**(4), 44101 (2003)
Ford, J., Mantica, G., Ristow, G.H.: The Arnol'd cat: failure of the correspondence principle. Phys. D: Nonlinear Phenomena **50**(3), 493 (1991)
Heisenberg W.: Oral history interview of W. Heisenberg by Thomas Kuhn. In Archive for the History of Quantum Physics. Deposit at Harvard University, Cambridge, (1963)
Heisenberg W.: The notion of a Closed Theory in Modern Science. In Heisenberg, W. (ed.) Across the Frontiers, pp. 39, Harper & Row, New York (1948)
Jona-Lasinio, G.: Spontaneous symmetry breaking—variations on a theme. Prog. Theretical Phy. **124**(5), 731 (2010)
Jones, R.O., Gunnarsson, O.: The density functional formalism, its applications and prospects. Rev. Mod. Phys. **61**(3), 689 (1989)
Landau, L.D., Lifshitz, E.M.: Quantum Mechanics Non-Relativistic Theory. Pergamon Press Oxford, Oxford (1978)
Mantica, G.: Quantum algorithmic integrability: The metaphor of classical polygonal billiards. Phy. Rev. E **61**(6), 6434 (2000)
Pauling, L.: Nature Chemical Bond. Cornell University Press, New York (1928)
Pauling, L.: The nature of the theory of resonance. In: Todd, A. (ed.) Perspectives in Organic Chemistry, p. 1. Interscience, London (1956)
Pople, J.A.: Nobel lecture: Quantum chemical models. Rev. Mod. Phys. **71**(5), 1267 (1999)
Porter, M.A., Cvitanovic, P.: Ground control to Niels Bohr: Exploring outer space with atomic physics. Not. AMS **52**(9), 1020 (2005)
Primas, H.: Chemistry Quantum Mechanics and Reductionism Perspectives in Theoretical Chemistry. Springer, Berlin (1981)
Scerri, E.R.: Collected Papers on Philosophy of Chemistry. World Scientific, Singapore (2008)
Schweber, S.S.: The young John Clarke Slater and the development of quantum chemistry. Hist. Stud. Phy. Biol. Sci. **20**(2), 339 (1990)
Thijssen, J.M.: Computational Physics. Cambridge Univ Press, Cambridge (2007)
Woolley, R.G.: Must a molecule have a shape? J. Am. Chem. Soc. **100**(4), 1073 (1978)
Woolley, R.G.: Molecular shapes and molecular structures. Chem. phy. lett. **125**(2), 200 (1986)
Zurek, W.H.: Decoherence and the transition from quantum to classical. Phys. Today, **44**(10), (1991)
Zurek, W.H.: Decoherence, einselection, and the quantum origins of the classical. Rev. Mod. Phy. **75**(3), 715 (2003)

Chapter 7
Some Conclusions and Random Thoughts

Cause when the goin' gets tough... The tough get goin'!

J.B. Blutarsky 1978

In the previous four chapters we have considered certain issues of theory reduction in the specific subjects of statistical mechanics, macroscopic phenomena, chaos, quantum mechanics and chemistry, which are united by the presence of singular limits and emergent properties. Nevertheless, these subjects appear to be almost completely independent of each other, and of interest to scholars only as separate fields.

Some readers may perhaps be disturbed by this fragmentation of science, which can be seen as an unavoidable consequence of our non-reductionist point of view.[1]

The main purpose of this chapter is to emphasise that even a non-reductionist approach may be consistent with a unitarianview of science, albeit non-hierarchical, not irrational and nor holistic.

7.1 Unity of Science Beyond Reductionism

Let us analyse some issues revealing often intricate overlapping of different sciences and techniques, which we consider to be the real signature of the unity of science.

7.1.1 Common Practice in Statistical Mechanics

Statistical mechanics is intimately interwoven with quantum mechanics and it often makes heavy use of experimental data. For instance the microscopic dynamics of a

[1] We characterise our approach as non-reductionist rather than anti-reductionist.

S. Chibbaro et al., *Reductionism, Emergence and Levels of Reality*,
DOI: 10.1007/978-3-319-06361-4_7,

liquid at room temperature is substantially classical, and its thermodynamic properties can be obtained within the framework of classical statistical mechanics. Then, under the assumption of thermodynamic equilibrium, standard textbooks describe the common practice of statistical mechanicists as follows: given a system made up of N particles with known interaction potential $U(|\mathbf{q}_i - \mathbf{q}_j|)$, contained within a vessel of volume V, one computes specific heat, free energy etc. Indeed, there exist rather powerful methods for accurately estimating, or reasonably approximating, the values of thermodynamic observables from knowledge of V, N and $U(r)$.

For example, dilute gases are described by the following equation of state:

$$\frac{p}{k_B T} = \rho + b_2(T)\rho^2 + b_3(T)\rho^3 + \dots \tag{7.1}$$

where $\rho = N/V$ and the virial coefficients $b_2, b_3, \dots$ are determined by $U(r)$.

On the other hand, despite the fact that the statistical mechanical problem is classical, the potential $U(r)$ has a genuine quantum mechanical origin and, in a first-principles approach, it should be determined by the solution of the Schrödinger equation, using for instance the methods illustrated in Chap. 6 (the Born-Oppenheimer approximation or the Car-Parrinello method). This is a rather elaborate and difficult task and only in a few special cases, e.g. in the presence of spherical molecules, can it be performed to a satisfactory degree of accuracy (Barker and Henderson 1976).

In reality, however, the common analytical and numerical approaches to simple liquids—e.g. liquids made of rare-gas atoms—starts from the assumption that the interaction potential is of Lennard-Jones type:

$$U(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right], \tag{7.2}$$

where the two parameters ε and σ respectively denote the characteristic energy and length scales. In this potential, the r^{-6} term, which represents the attractive force dominating the long-range interaction, is indeed derived from quantum mechanical calculations, accounting for dipole-dipole interactions. But the r^{-12} contribution, which represents a repulsive force that diverges as $r \to 0$, has been introduced on purely phenomenological grounds. Moreover, only a very limited class of systems allows the constants ε and σ to be determined from the Schrödinger equation, under the Born-Oppenheimer approximation or analogous techniques. Therefore, it is common practice to adopt the form of Eq. (7.2) for $U(r)$, and to fit it to experimental data, by matching the values of ε and σ. In brief, this approach consists of the following steps for simple liquids:

- the Lennard-Jones potential Eq. (7.2) is assumed to hold,
- thermodynamic properties (e.g. the virial coefficients $b_2, b_3, \dots$) are computed from this potential,
- ε and σ are determined to fit the experimental data, via Eq. (7.1).

In more complex systems, the Lennard-Jones potential is often replaced by more elaborate effective potentials that are typically suggested by physical intuition and

contain a certain number of free parameters, to be determined by data fitting. As a well known example, consider the Tosi-Fumi potential, for molten salts:

$$U(r) = Ae^{-r/\lambda} - \frac{C}{r^6} - \frac{D}{r^8},$$

where A, C, D and λ are adjusted to fit the experimental data (Hansen and McDonald 1986).

In the vulgate of naive reductionism, one starts from the most fundamental theory (quantum mechanics, in our case) and derives the most phenomenological one (thermodynamics): the solution of the Schrödinger equation produces the effective potential $U(r)$ and then statistical mechanics computes the thermodynamics properties of interest. In practice, one follows instead the opposite route: starting with the experimental data, which may be expressed e.g. by Eq. (7.1), the classical potential is determined.

At first glance, one could think that this practical protocol of statistical mechanics essentially prevents its falsification. This conclusion, however, is incorrect: statistical mechanics predicts non-trivial features of the physics of matter, which can be experimentally tested and which, remarkably, sometimes do not depend on the details of the microscopic interactions. For instance, one may recall:

- the Maxwell-Boltzmann distribution of the velocity of molecules at a given temperature; recall that such a distribution holds independently of the shape of $U(r)$; and
- the existence of phase transitions, and the scaling features of critical phenomena, in which only a few quantitative features of the potential are relevant (the universality classes).

7.1.2 *From Boltzmann's Disputes with Zermelo to Models Built from Data*

In Chap. 3 we illustrated Boltzmann's reply to Zermelo's objection based on Poincaré recurrences. Boltzmann *merely* observed that the typical recurrence time T_R for a macroscopic object is exceedingly large: $T_R \sim \tau_0 C^N$, where τ_0 is a characteristic time of the system made up of N particles and $C > 1$. Consequently, no recurrence will ever be observed in any macroscopic system. Such a conclusion, reached by Boltzmann thanks to his remarkable physical intuition, anticipated Kac's lemma, now a well known result of Ergodic theory.

For the sake of simplicity, let us consider a discrete time dynamical system, formally defined by:

$$\mathbf{x}_1 = \mathscr{S}\mathbf{x}_0, \mathbf{x}_2 = \mathscr{S}\mathbf{x}_1 = \mathscr{S}^2\mathbf{x}_0, \ldots, \mathbf{x}_k = \mathscr{S}^k\mathbf{x}_0,$$

and let A be a measurable set in the phase space. The recurrence time $\tau_A(\mathbf{x})$ is defined by:

$$\tau_A(\mathbf{x}) = \inf\{k \geq 1 : \mathbf{x} \in A, \quad \text{and} \quad \mathscr{S}^k\mathbf{x} \in A\}.$$

If μ denotes the invariant measure, the average recurrence time is expressed by:

$$\langle\tau_A\rangle = \frac{1}{\mu(A)} \int_A d\mu(\mathbf{x})\tau_A(\mathbf{x}).$$

For an ergodic system, the following classical result due to Kac (1957) holds:

$$\langle\tau_A\rangle = \frac{1}{\mu(A)}. \tag{7.3}$$

Taking A in the form of a hypercube of linear size ε, in a phase-space-volume preserving system with N degrees of freedom, whose typical range of variation is of order $O(L)$ for each component of $\mathbf{x}$, one has

$$\langle\tau_A\rangle \sim \left(\frac{L}{\varepsilon}\right)^N, \tag{7.4}$$

an exponentially long (in N) return time, as Boltzmann substantially understood.

Poincaré's theorem holds for a generic ergodic system, even a dissipative one. In such a case, the invariant measure is singular, and typically has a fractal nature whose main feature is its fractal dimension D.[2] Writing the measure of a set of linear size ε, e.g. a small ball of radius ε around a point $\mathbf{y}$ belonging to the attractor, $S_\mathbf{y}(\varepsilon) = \{\mathbf{x} : |\mathbf{y} - \mathbf{x}| < \varepsilon\}$, we have:

$$\mu(S_\mathbf{y}(\varepsilon)) = \int_{S_\mathbf{y}(\varepsilon)} d\mu(\mathbf{x}) \sim \left(\frac{\varepsilon}{L}\right)^D.$$

Therefore, in dissipative chaotic systems Eq. (7.4) is still valid, as long as N is replaced by D.

Equation (7.4) has important consequences for the predictability of physical phenomena, and for the possibility of building models on the basis of experimental data. Briefly, to build a predictive model from the available data, one is supposed to search for a past state similar to the present state of a given pheonomenon of interest. Then, looking at the sequence of events that followed the past state, one may infer by analogy the evolution that will follow the present state. In other words, given a known sequence of "analogues", i.e. of past states $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_M$, which resemble each other closely in pairs, so that $|\mathbf{x}_i - \mathbf{x}_j| \leq \varepsilon$, with ε reasonably small, one makes

[2] In general, a single dimension is not enough to fully characterise such measures, which typically have a *multifractal* nature; infinitely many dimensions (the Renyi dimensions) are necessary for that. This technical issue is not important for our purpose.

the approximate prediction:

$$\mathbf{x}_{M+1} = \mathbf{x}_{k+1} \ ,$$

if $\mathbf{x}_k$ is an analogue of $\mathbf{x}_M$. One may then proceed to build a model of the phenomenon, i.e. to determine a function $\mathbf{f}$ such that the sequences of states is well approximated by the dynamical system $\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k)$. The application of this method requires knowledge of at least one analogue. It is not difficult to realise that such knowledge requires sufficiently long sequences, at least of duration of order $O\left([L/\varepsilon]^D\right)$. Therefore, one may take into consideration only cases with rather small D, say less than 4 or 5 (Kantz and Schreiber 1997; Cecconi et al. 2012).

This shows that, while the exponential growth of T_R as a function of D (or N) plays a positive role in the explanation of foundational issues of statistical mechanics, such as irreversibility, it impacts rather negatively on our ability to make predictions, solely relying on previously acquired data. One can say that D larger than 6 renders the approach described here useless, because it makes it practically impossible to observe the "same" state twice, i.e. within an acceptable accuracy ε. Indeed, Maxwell had already understood this fact, noting that:

> It is a metaphysical doctrine that from the same antecedents follow the same consequents. [...] But it is not of much use in a world like this, in which the same antecedents never again concur, and nothing ever happens twice...
>
> The physical axiom which has a somewhat similar aspect is "That from like antecedents follow like consequents."

It is puzzling how long the time-series-analysis community took to understand the practical relevance of Kac's lemma and its consequences. The experts of nonlinear time series reached the same conclusions that Maxwell had reached more than a century earlier!

Let us stress an important point, before developing our argument. Our analysis is intimately related to the existence of different levels of reality and thus to observations. In this sense, we are mainly concerned about epistemic or procedural reductionism, since only that has a precise scientific meaning. One might object that more demanding forms of reductionism only apply to ideal theories, and in that respect Boltzmann's argument against Zermelo would not be definitive. We believe that this standpoint is weak. First, formalised ultimate theories do not exist in real world. Second, theories without an observation level are physically irrelevant. Science is intrinsically linked to observations. In this sense, Boltzmann's argument, which turns exact in the limit of an infinite number of molecules, has to be considered much more than a merely practical consideration. In conclusion, idealistic objections do not refer to natural sciences (maybe they apply to mathematics).

7.1.3 Law Without Law?

Wheeler (1983) introduced the maxim *law without law*, to express the possibility of building part of physics without recourse to detailed treatments in terms of fundamental theories, simply following *regularity principles* implied by large numbers and by few *relevant aspects*. In Chap. 3, we discussed two of the examples considered by Wheeler, namely the Maxwell-Boltzmann distribution for the velocities of the particles of systems in equilibrium at given temperature, and the universality of the critical exponents.

Another interesting example of *law without law* illustrated by Wheeler is the statistics of heavy nuclei. A heavy nucleus contains several hundred interacting nucleons (protons and neutrons), whose interaction potential is only vaguely known. Therefore, a quantum mechanical treatment of such nuclei is practically impossible. Nevertheless, according to Landau and Smorodinsky's seminal ideas, and Wigner's intuition, the statistical properties of the corresponding energy spectra, which are properties of highest physical interest, can be reproduced by large Hermitian matrices $\mathbf{H}$ with random elements, mimicking the Hamiltonian operators (Brody et al. 1981).

The fundamental ingredient of this conjecture is clear: provided one is not interested in the tiny details of the phenomenon, it is reasonable to assume that the large number of degrees of freedom makes the overall behaviour depend only on a few properties of the system. This approach has a genuine probabilistic character and gives good results if the statistical properties of the elements of the random matrices are based on proper general principles of quantum many-body theory. Basically, the following must hold:

- the ensemble must depend only on a few assumptions, including the hermiticity (or unitarity) of the matrices;
- the statistical properties of an individual (large) matrix should be close to the ensemble average (which is an ergodic property).

Under the above assumptions, it is possible to derive the celebrated semicircular and circular theorems for the distribution of the eigenvalues, as well as the statistical features of the level spacing between two nearest-neighbors random matrices $\mathbf{H}$ in the limit $N \gg 1$.

Remarkably, the experimental results on heavy nuclei are in good agreement with the predictions of the random matrices approach (Brody et al. 1981).

Note that both in critical phenomena and in the spectral properties of heavy nuclei, the details of the Hamiltonian do not play a major role, and "first principles" (i.e. the partitition function provided by quantum many body theory) are not used. The reason is that one focuses on:

- few qualitative properties, i.e. on the universality class for critical phenomena, and the hermitian (or unitary) character of the random matrices;
- the presence of a large number of degrees of freedom.

At the end of his paper, Wheeler writes:

Oh, how beautiful and simple it all is! How could we ever have missed it so long!

These results are certainly very beautiful, but we do not at all think that they are simple. Only *a posteriori* could they appear so. Indeed, they follow from results of probability theory that are far from trivial.[3] In practice, Wheeler describes as *regulatory principles* certain limit theorems of probability theory:

- The Boltzmann-Gibbs law follows from the microcanonical distribution, which is based on the ergodic hypothesis.[4]
- The universality of the statistical features of heavy nuclei is a sort of ergodicity: the single (large) nucleus is well described by its average properties.
- The universal properties of critical phenomena can be seen as a generalisation of the central limit theorems.[5]

7.2 It From Bit?

The success of Wheeler's approach is intriguing, and may suggest the possibility that physical results might be obtained merely from information theory and statistical inference, without any physical insight. Indeed, Wheeler (1992) himself formulated the aphorism *"it from bit"*, speculating on the possibility that information is the ultimate building block of our understanding of the world. We believe that this is an interesting but easily misinterpreted point of view.

7.2.1 Statistical Mechanics as Statistical Inference?

According to a radically anti-dynamical point of view, statistical mechanics is but a form of statistical inference rather than a theory of objective physical reality. In this

[3] There are cases in which emergent features are almost trivial consequences of smoothing mechanisms. For instance the bouncing on the asphalt of a small bead of, say, half a centimetre radius, is quite irregular, because the roughness of the asphalt is of a size comparable to that of the bead. By contrast, a tennis ball bounces regularly, as if the asphalt were smooth, because its diameter is significantly larger than the scale of the asphalt roughness.

[4] The Boltzmann-Gibbs law:

$$\rho(\mathbf{X}) = Const.\, e^{-\beta H(\mathbf{X})}$$

describes the statistical features of a system which exchanges energy with an external environment at given temperature $1/\beta$, and can be obtained from the microcanonical distribution in the large-N limit.

[5] Critical phenomena can be treated by means of the powerful renormalisation group method, which has played an important role in the physics of the second half of the twentieth century, both as a conceptual and a computational tool. Remarkably, the renormalisation group has a probabilistic interpretation as a generalisation of the central limit theorem to the case of non independent variables. This fact has been understood in the mid-seventies by Bleher and Sinai, and has been widely investigated by Jona-Lasinio (2001).

approach, probabilities measure the degree of truth of a logical proposition, rather than describing the state of a system.

In this context, Jaynes proposed the maximum entropy principle (MEP) as a general rule for finding the probability of a given event when only partial information is available. If the mean values of m independent functions $f_i(\mathbf{x})$ are given:

$$c_i = \langle f_i \rangle = \int f_i(\mathbf{x})\rho(\mathbf{x})d\mathbf{x} \quad i = 1, \ldots, m,$$

the MEP rule determines the probability density ρ of the events compatible with these mean values, by maximising the "entropy"

$$H = -\int \rho(\mathbf{x}) \ln \rho(\mathbf{x})d\mathbf{x},$$

under the constraints $c_i = \langle f_i \rangle$. Using the Lagrange multipliers one easily obtains

$$\rho(\mathbf{x}) = \frac{1}{Z} \exp \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x})$$

where $\lambda_1, \lambda_2 \ldots \lambda_m$ depend on $c_1, c_2, \ldots, c_m$. Therefore, when applied to the statistical mechanics of systems with a fixed number of particles and the unique constraint of the energy mean value, the MEP leads to the usual canonical distribution in a very simple fashion.

As a technical but rather important detail, we note that this holds only if $\mathbf{x}$ is the vector of the canonical coordinates (i.e. positions and momenta of the particles). Analogously, for systems of varying numbers of particles, the grand canonical distribution is obtained by additionally constraining the mean number of particles. Many find in these facts an unquestionable proof of the validity of the MEP, but we will see that it is just a matter of fortunate coincidence, related to the choice of canonical coordinates.

The most frequent objection to the MEP is summarised by the maxim *Ex nihilo nihil*, or citing the title of one of Peres's paper, "Unperformed experiments have no results" Peres (1978), i.e. it is not possible that we infer something about a real phenomenon, thanks to our ignorance (Ma 1985). Apart from this very general observation, the weakest technical aspect of the MEP approach is the dependence of the results on the choice of the variables.

For simplicity's sake, consider a scalar random variable X, ranging over a continuum, whose probability distribution function is p_X. It is easy to realise that the "entropy" $H_X = -\int p_X(x) \ln p_X(x)\, dx$ is not an intrinsic quantity of the phenomena concerning X. With a different parametrisation, i.e. using the coordinates $y = f(x)$ with an invertible function f, rather than x, the entropy of the same phenomenon would now be given by

$$H_Y = -\int p_Y(y) \ln p_Y(y)\, dy$$

with $p_Y(y) = p_X(f^{-1}(y))/|f'(x = f^{-1}(y)|$. Therefore, one has

$$H_Y = H_X + \int p_X(x) \ln |f'(x)|\, dx$$

The MEP gives different solutions if different variables are adopted to describe the very same phenomenon.

In order to avoid this dependence on the choice of variables, Jaynes later proposed a more sophisticated version of the MEP, in terms of the relative entropy:

$$\tilde{H} = -\int \rho(\mathbf{x}) \ln\left[\frac{\rho(\mathbf{x})}{q(\mathbf{x})}\right] d\mathbf{x},$$

where q is a known probability density. Of course, $\tilde{H}$ depends on q; but, at variance with the entropy, it does not depend on the chosen variables. On the other hand, one must decide how to select q, and this issue is equivalent to the problem of choosing the "proper variables". Therefore, even this more elaborate method is non-predictive, and we see no reason to pursue the MEP approach further in the field of statistical mechanics. In any event, the interested reader is referred to the existing extensive literature, e.g. (Jaynes 1967; Uffink 1995), for further details.

As a general remark, let us stress that science concerns nature, it is not an ideal construction. Consequently, some form of empiricism cannot be avoided in the development of correct models and theories.[6]

7.2.2 Algorithmic Complexity: A Key to Understanding Nature?

Why is the physical world comprehensible? It is self-evident that the external world is not just a jungle of irregular facts: on the contrary, many regularities are quite clear. In brief, we can say that there exist laws that capture at least certain aspects of reality, and nature in particular.[7] We do not enter into the debated and open problem of whether the laws of nature are merely human constructs, or have an autonomous status; in other words, whether the laws of nature are discovered or invented by us.

Typically, scientists engaged with fundamental research, e.g. in the study of elementary particles and cosmology, are inclined to believe that the laws of nature exist

[6] Logical empiricists went certainly too far in this direction, but they appreciated this fundamental point. Some contemporary scientists and philosophers (notably *postmodernists*) seem very far from getting the message.

[7] Understanding the existence of the laws of nature was a great achievement brought about by the scientific thought. The path towards such an important step has been long and tortuous, often affected by theological views, such as in Leibniz's philosophy.

independently of scientific enquiry, while the comunity of applied sciences usually preferes the other point of view.

McAllister (2003), for instance, states that the aim of science is to organise in the most economical fashion the data collected from experiments.[8] A scientific law, once established, allows us to predict the outcomes of all related experiments. For instance, rather than keeping in mind all individual instances of refracted light, we can simply summarise them all with Snell's law. This way of considering the laws of nature was quite popular at the end of the nineteenth century. For instance, Duhem (1991) basically shared Mach's view and neither Mach nor Duhem can be considered reductionists: they both opposed the atomistic hypothesis.

Recently, their point of view has been reconsidered in the framework of algorithmic complexity (Li and Vitanyi 1992). For instance, Solomonoff, one of the fathers of the theory, considers (without any reference to Mach) a scientific law, and more generally a theory, as an algorithm for compressing the results of experiments, providing a mathematical formalisation of the idea of science as an "economy of thought" (Solomonoff 1964):

> *The laws of science that have been discovered can be viewed as summaries of large amounts of empirical data about the universe. In the present context, each such law can be transformed into a method of compactly coding the empirical data that gave rise to the law.*

In brief, the laws of nature are compression algorithms, and the theory of everything can be seen as the ultimate algorithm for the whole universe. It is rather ironic that the approach to science as economy of thought,developed by two anti-reductionists and anti-atomists such as Mach and Duhem, should have been reformulated in modern terms and used today to support extreme reductionism.

On the other hand there is a subtle(and often ignored) aspect which should be considered: the role of initial conditions, which are usually independent of the laws of nature. Such an important point had been already realised by Newton (1957)[9] who noted that all the planets move in the same direction on concentric orbits, while the comets move in eccentric orbits, concluding that such a property of the solar system cannot be a mere coincidence, but a choice: the initial condition.

Interestingly, sometimes the initial conditions are not just a frozen accident, but can be explained. For instance, the fact that all the planets move in the same direction as well as their distances from the sun (Bode's law) can be explained within the theory of the formation of the solar system from a spinning cloud of gas. In an analogous way, large scale structures of fluids in planetary atmospheres, such as the well known red spot of Jupiter (a giant persistent storm discovered by Galileo with

---

[8] Such a view of the science is shared by many scientists in the positivism or neopositivism currents. As an interesting exception, we may recall Born (1948), who ironically noted: *if we want to economise thinking, the best way would be to stop thinking at all*, and then *the expression "economy of thinking" may have an appeal to engineers or others interested in practical applications, but hardly to those who enjoy thinking for no other purpose than clarify a problem.*

[9] Wigner (1963) considers the understanding of the distinction between laws and initial conditions as the most important contribution that Newton made to science. He considers this even more important than the laws of gravitation and dynamics.

his telescope) are self-generating from generic initial states by the bi-dimensional Navier-Stokes equations, properly modified to take rotation into account (Miller et al. 1992). Another example is the focussing collisions in which neutrons with high kinetic energy but random orientation are converted into lower velocity neutrons with a preferred direction (Silsbee 1957).

In Chap. 5, we discussed the deep connection between chaos and algorithmic complexity. We stress now that the presence of chaos does not imply the impossibility of understanding the evolution law of the system, and therefore to "compress" the phenomenon. In our opinion, there is a certain confusion on this point. For instance, Davies (1990) writes

> *there is a wide class of physical systems, the so-called chaotic ones, which are not algorithmically compressible.*

What is not compressible is the time sequence generated by chaotic systems, and this is due to the non-compressibility of a generic initial condition, see Chap. 5. However, this simply implies the practical impossibility of reconstructing the evolution law merely from experimental data, as discussed in Sect 7.1. But this is just a practical limitation, and perhaps a concern for the naive inductivist approach. In fact, the evolution laws of chaotic systems are typically obtained in non-inductive ways, for instance within a well-established and formalised theory, and are typically compressible. Turbulent flows, a paradigm case of complex systems, are governed by the Navier-Stokes equations which can be written in two lines.

We stress again that disregarding the distinction between initial conditions and laws of nature can lead to great confusion. For example, McAllister (2003) observes that *empirical data sets are algorithmically incompressible*, concluding that the task of scientific laws and theories does not consist in compressing empirical data. The remark on the incompressibility of generic empirical data is surely correct: it is nothing but Martin-Löf's result on random sequences, which we mentioned in Chap. 5, and is not surprising at all. On the contrary the conclusion that scientific laws and theories do not constitute algorithmic compressions of empirical data sets, collected from observations and measurements, is questionable.

Regarding the opinion that scientific laws constitute a compression of empirical data, McAllister claims that no scientist has ever made such a statement. Apart from the historical aspects, we want to discuss the following example: consider a series of light-refraction experiments, in which $\{\alpha_1, \ldots, \alpha_N\}$ are the angles of the incident rays, and $\{\beta_1, \ldots, \beta_N\}$ the angles the refracted rays. The sequences $\{\alpha_1, \ldots, \alpha_N\}$ and $\{\beta_1, \ldots, \beta_N\}$ may or may not be compressed. This is a *frozen accident* which depends on the protocol followed by the scientist while preparing the experiment.[10] However, once the values $\{\alpha_1, \ldots, \alpha_N\}$ are known, the sequence $\{\beta_1, \ldots, \beta_N\}$ is simply determined by the Snell's law: $sin\alpha/sin\beta = n$, and this is a genuine form of compression.

[10] For instance, in the case of the protocol $\alpha(n) = \alpha(n-1) + \delta_\alpha$, the sequences can be compressed. For intrisically random processes, they cannot.

A less trivial instance concerns the Navier-Stokes equation for fluids: even in chaotic systems, whose time sequences are not compressible, the laws have a predictive power, in the sense that they generate the results of observations and experiments.

The claim that *the world is comprehensible because it is algorithmically compressible* is, in our opinion, a truism which is equivalent to saying that the laws of nature exist. The reason why our minds successfully represent many aspects of nature remains an enigma. However, some understanding arises from the fact that, fortunately:

- Physical laws obey spatially and temporally local rules, i.e. a given phenomenon is not affected too much by events which are distant in time and/or in space. Practically all the equations of physics, like the laws of Maxwell, Schrödinger, Newton etc., obey the locality assumption.
- Despite the enormous complexity and the intricate interconnections of different phenomena, often there is a *scale separation* which allows separate treatments and descriptions of the different levels on which reality may be considered. For instance, fluids, which are comprised of atoms obeying quantum mechanics, are well described by the Navier-Stokes equations originally derived assuming the hypothesis of the continuity of matter.

It is important to note that the main effort in building a theory concerns the formulation of the proper questions, and the identification of the appropriate variables.[11] Unfortunately, we have no definitive method for selecting the right ideas or the proper variables.

7.3 Concluding Remarks

Most likely, apart perhaps from a few high energy physicists and cosmologists, nobody really believes that the ultimate mission of science is to pass to deeper and deeper levels of reality, in search of the smallest constituents of matter, strings or other even more exotic objects, while the rest— chemistry, macroscopic phenomena and so on—should be considered nothing more than a shadow of that deepest level. On the other hand, although the reductionist goal cannot be achieved in practice, in principle it could realise the unity of science, because that would produce a unique body of knowledge, developed from the most fundamental discipline (string theory or more exotic descriptions). Science would not split into many disconnected branches and, conceptually at least, would have a unique hierarchical structure founded on the most fundamental level of description.

[11] Such a difficulty is well known in statistical physics; it has been stressed e.g. by Onsager and Machlup (1953) in their seminal work on fluctuations and irreversible processes, with the caveat: *how do you know you have taken enough variables, for it to be Markovian?* In a similar way, Ma (1985) notes that: *the hidden worry of thermodynamics is: we do not know how many coordinates or forces are necessary to completely specify an equilibrium state*.

Could we pursue the unity of science in a non-reductionist context? Perhaps this question is not totally relevant. Nevertheless, we briefly argue that the possible unity of science does not need to be based on a supposed hierarchical structure of nature. The unity of science simply has historical, conceptual and technical origins.

In the previous section, specific examples showed that:

- In real research activity, one faces intricate overlapping of different scientific fields.
- Certain techniques play important roles in different sciences.
- The slogan *law without law* may be interpreted in terms of probability theory.

This reveals, for instance, the cross-fertilisation mechanism which links physics and mathematics. Among the many, we may pick some other relevant examples:

- The ergodic hypothesis, introduced by Boltzmann as a technical tool of statistical mechanics, has developed into a new important chapter of the (mathematical) measure theory.
- The modern theory of stochastic processes sprang out of Einstein, Smoluchowski and Langevin's investigation of Brownian motion.
- Pre-existing mathematical objects, such as Hilbert spaces and linear operators, played a major role in the formulation of quantum mechanics.
- Random matrices, initially introduced in nuclear physics as technical tools, proved to be deeply connected with number theory, which is one of the most abstract of mathematical fields.

Although mathematics and physics acquired their distinctive flavours during the nineteenth century, the separation was far from absolute (Dorato 2010; Jona-Lasinio 2005). Some eminent scientists explicitly claim that there is no real distinction between physics and mathematics. The famous mathematician V. Arnold believed that: *Mathematics is a part of physics. Physics is an experimental science, a part of natural science. Mathematics is the part of physics where experiments are cheap.* Perhaps, Arnold only meant to provoke his audience, but his view contains a lot of truth, as confirmed by the experience of many great mathematicians. For instance, Riemann, who was deeply interested in physics, made important contributions to the theory of wave propagation in compressible gases. Hilbert studied general relativity in the same period as Einstein, and the mathematical aspects of the Boltzmann equation. Interestingly, his approach to the kinetic theory of gases was rather physical and became the basis for the Chapman-Enskog method. Kolmogorov formulated his theory of fully developed turbulence after an accurate study of experimental data, following Richardson's very physical intuition. To stress this difficulty in separating mathematics and physics, we can quote Atkins (2003): *Determining where mathematics ends and science begins, is as difficult and as pointless, as mapping the edge of the morning mist*.

References

Atkins, P.: Galileo's Finger: the Ten Great Ideas of Science. Oxford University Press, Oxford (2003)

Barker, J.A., Henderson, D.: What is "liquid" ? understanding the states of matter. Rev. Mod. Phys. **48**(4), 587 (1976)

Born, M.: Natural Philosophy of Cause and Chance. Read Books, Vancouver (1948)

Brody, T.A., Flores, J., French, J.B., Mello, P.A., Pandey, A., Wong, S.S.M.: Random-matrix physics: spectrum and strength fluctuations. Rev. Mod. Phys. **53**(3), 385 (1981)

Cecconi, F., Cencini, M., Falcioni, M., Vulpiani, A.: Predicting the future from the past: an old problem from a modern perspective. Am. J. Phys. **80**, 1001 (2012)

Davies, P.: Why is the physical world so comprehensible? In: Zurek, W.H. (ed.) Complexity, Entropy and the Physics of Information, pp. 61. Santa Fe Institute (1990)

Dorato, M.: Why is the language of nature mathematical? In: Antonini, G., Altamore, A. (eds) Galileo and the Renaissance Scientific Discourse, pp. 65. Nuova Cultura (2010)

Duhem, P.: The Aim and Structure of Physical Theory. Newblock Princeton University Press, Princeton (1991)

Hansen, J.P., McDonald, I.R.: Theory of Simple Liquids. Academic Press Inc., London (1986)

Jaynes, E.T.: Foundations of probability theory and statistical mechanics. In: Bunge, M. (ed.) Delaware Seminar in the Foundations of Physics, vol. 1, p. 76. Springer-Verlag, Berlin (1967)

Jona-Lasinio. G.: La matematica come linguaggio delle scienze della natura. In Matematica cultura e societá, pp. 93. Quaderni della Scuola Normale Superiore di Pisa (2005)

Jona-Lasinio, G.: Renormalization group and probability theory. Phys. Rep. **352**(4), 439 (2001)

Kac, M.: Probability and Related Topics in Physical Sciences. American Mathematical Society (1957)

Kantz, H., Schreiber, T.: Nonlinear Time Series Analysis. Cambridge University Press, Cambridge (1997)

Li, M., Vitanyi, P.: Inductive reasoning and Kolmogorov complexity. J. Comput. Syst. Sci. **44**(2), 343 (1992)

Ma, S.K.: Statistical Mechanics. World Scientific, Singapore (1985)

McAllister, J.W.: Algorithmic randomness in empirical data. Stud. Hist. Philos. Sci. Part A **34**(3), 633 (2003)

Miller, J., Weichman, P.B., Cross, M.C.: Statistical mechanics, Euler's equation, and Jupiter's red spot. Phys. Rev. A **45**(4), 2328 (1992)

Newton. I: Optics. Dover Publications, New York (1957)

Onsager, L., Machlup, S.: Fluctuations and irreversible processes. Phys. Rev. **91**(6), 1505 (1953)

Peres, A.: Unperformed experiments have no results. Am. J. Phys **46**(7), 745–747 (1978)

Silsbee, R.H.: Focusing in collision problems in solids. J. Appl. Phys. **28**(11), 1246 (1957)

Solomonoff, R.J.: A formal theory of inductive inference. part i and ii. Inf. Control **7**(1), 1 (1964)

Uffink, J.: Can the maximum entropy principle be explained as a consistency requirement? Stud. Hist. Philos. Mod. Phys. **26B**(3), 223 (1995)

Wheeler, J.A.: On recognizing a law without law, Öersted medal response at the joint APS-AAPT meeting, Am. J. Phys. **51**, 398, 25 January 1983, New York, (1983)

Wheeler, J.A.: Recent thinking about the nature of the physical world: it from bit. Ann. N. Y. Acad. Sci. **655**(1), 349 (1992)

Wigner, E.P.: Events, laws of nature, and invariance principles. Nobel Lectures in Physics, (1963)