Part II of *Complex Methodological Individualismm* with F. Di Iorio

The Micro-Meso-Macro Dialectic in the Natural and Social Sciences

Jean Petitot

2026

Contents

1 Epistemological remarks **1**

1.1 Complex Methodological Individualism (CMI) . . . 1
1.2 Complex systems . . . 5
1.3 Computational syntheses and conceptual analyses . . . 6
1.4 An example of transdisciplinary universality: Pareto's laws . . . 8
1.5 Reminders about some debates concerning MI . . . 9
1.6 Towards modeling . . . 11
1.6.1 A classic example that is off-topic but very instructive . . . 11
1.6.2 The Coleman's boat . . . 15

2 CMI in Physics **17**

2.1 Macroscopic thermodynamics (MH) . . . 17
2.1.1 The relevance of returning to physics . . . 17
2.1.2 Description of the "holistic" thermodynamical macro level (MH) . . . 18
2.1.3 Ideal gas equation of state . . . 23
2.1.4 Phase transitions . . . 23
2.1.5 The van der Waals equation (1873) (MH with a touch of meso-CMI) . . . 27
2.2 Microstatistical thermodynamics (CMI) . . . 29
2.2.1 The revolution of statistical physics. Gas kinetics . . . 29
2.2.2 Statistical computations . . . 30
2.2.3 Thermodynamic irreversibility VS mechanical reversibility . . . 33
2.2.4 The "micro-canonical" case (purely CMI) . . . 33
2.2.5 The "canonical" case and Gibbs' formula (combination of CMI and MH) . . . 35
2.2.6 The perfect gas . . . 37
2.2.7 The van der Waals equation . . . 38
2.2.8 The Boltzmann equation . . . 40
2.2.9 Hilbert's Sixth Problem. Part I: Meso → Macro . . . 46

2.2.9.1 Mesoscale → macroscale: the hydrodynamic limit. . . . 46
2.2.10 Hilbert's Sixth Problem. Part II: Micro →Meso 49
2.2.10.1 Lanford's Theorem 50
2.2.10.2 Clarifications. 55
2.2.10.3 The Deng, Hani, Ma theorem 61
2.2.10.4 A scientific epic. 62
2.3 Magnetic media 62
2.3.1 Phase transitions and mean fields (CMI and MH) 62
2.3.2 Onsager's computations and scaling laws 64
2.3.3 The renormalization group (purely CMI) 67

3 Natural self-organized complex systems 71
3.1 Spin glasses 71
3.2 Neural networks and CMI in neuropsychology 72
3.3 Sandpiles, sanddunes and self-organized criticality 76
3.4 Towards a social physics of collective behaviors 77
3.5 Operativity and efficiency of CMI 81

4 From ethology to Game Theory 83
4.1 Hayek's Sensory order 83
4.2 Ethology: flocks, herds, schools, swarms 84
4.3 Collective intelligence 87
4.3.1 Social insects 87
4.3.2 Stigmergy 87
4.3.3 Eusociality 89
4.3.4 Cognitive VS Eusocial complexification 89
4.3.5 Mandeville's Fable of the Bees 90
4.3.6 Eusociality in humans 92
4.3.7 Ober's analysis of ancient greek poleis (cities-states) . . . 93
4.4 Evolutionary game theory 95

5 Conclusion of Part II 101

Bibliography 105

Chapter 1

Epistemological remarks

1.1 Complex Methodological Individualism (CMI)

The purpose of this second part is to contextualize the question of *methodological individualism* (MI) in social sciences into a broader and more general scientific context:

(i) first, by reformulating it within the framework of complex systems and thus treating it as a *complex methodological individualism* (CMI); and

(ii) in including physical, biological and ethological natural sciences and not just social sciences.

In addition to our first part, we refer to numerous articles in the *Handbook of Methodological Individualism* (Bulle-Di Iorio (eds) 2023) for a discussion of the specifically social, political, economic, cultural, and institutional aspects of MI, as well as the difficult epistemological problems they raise in the humanities. In this second part, we will instead focus on the *intrinsically* transdisciplinary nature of the CMI issue, which we have already briefly addressed in (Petitot 2016 and Petitot 2023). CMI concerns the emergence of global collective properties (structures, organizations, processes) at the macroscopic level in populations composed at the microscopic level of a very large number of elementary individuals interacting with each other.

As Jon Elster emphasizes in "What's the Alternative" (Elster 2023), IM is simply a "methodological injunction" prescribing the search for micro foundations of macro states. It makes no "ontological" assumptions, either micro or macro. For example, in the psycho-social human sciences, it does not make any

> "substantive psychological assumptions about human behavior" (p.118)

Elster discusses the macro-level application of purely micro-level concepts, a kind of hypostasis that risks being unintelligible if it is not very finely elaborated. Obviously, subjects are immersed in systems of norms, values, and institutionalized rules that are given to them by their education, history, and

culture. But these rules function as external macro "fields" that are internalized and shared by individual subjects possessing social "dispositions." They cannot be thought of in terms of "ontological realism." We will return to this point below.

In CMI, it is therefore assumed that the system under consideration has at least *two levels of reality* (micro/macro) and that these two levels are *autonomous* and are therefore described by different empirical and theoretical vocabularies, which may even be *conceptually incommensurable*. It is assumed that their variables and state functions, dynamics, rules, laws, and rationalities are specific and radically heterogeneous.

The mechanical vocabulary of the motion of molecules in a gas has nothing to do semantically with the thermal vocabulary of thermodynamics; the granular vocabulary of grains rolling over each other has nothing to do semantically with the geometric and morphological vocabulary of sand dunes; the vocabulary of action potentials (spikes) propagating along neural axons through mechanisms of opening/closing sodium or potassium ion channels has nothing to do semantically with the vocabulary of the perceptual and cognitive performances of complex neural networks (whether biological or artificial); the vocabulary of elementary local actions of social insects such as bees, ants, or termites, or also fish or birds, has nothing to do semantically with the global geometric and architectural vocabulary of the combs of beehives, anthills, termite mounds, schools, or flocks; the vocabulary of individual interactions and transactions that induce incredible complex connectivity in internet networks has nothing to do, semantically speaking, with the vocabulary used to describe scientific communities, macroeconomic organizations, or political ideologies.

In all these cases, the transition from micro to macro is highly problematic and even amounts to a *conceptual antinomy*, the "mereological" antinomy of whole/part relationships in wholes composed of a very large number of components. This antinomy is very old and dates back to medieval debates on realism and nominalism. In modern times, it is omnipresent in Leibniz, who oscillated throughout his life between various versions:

(i) *Nominalist thesis*. Compound wholes are composed of simple substances (singular and individual "metaphysical atoms") but are not substances as such. They are merely aggregates lacking a specific principle of individuation, assemblages cut out by language and perception in reality. Their unity is only nominal.

(ii) *Realist antithesis*. There are composite substances that are substantially more than aggregates of nominally composite simple substances, which are individuated by a "vinculum substantiale" and possess an "ontological" reality.[1]

For Kant, this is the second antinomy of transcendental dialectics in the *Critique of Pure Reason* (Kant 1781). It contrasts the thesis that all substantial wholes are composed of simple parts with the antithesis that composite substances are not made up of simple parts.

[1] In (Petitot 1999) we commented on André Robinet's (1986) penetrating analysis of Leibniz's theses.

There are a considerable number of modern epistemological studies on these questions of mereology and emergence. They are revived each time a new scientific success of CMI emerges. We will return to them several times in our examples. In the physical and biological sciences, they are omnipresent. In the social sciences, they are much rarer. This is unfortunate because MI cannot do without them. As Nathalie Bulle explains in "Methodological Individualism as Holism of Parts" (Bulle 2023), the characteristic of the social sciences is that micro-agents have subjective interiority and "give meaning" to their actions. The classical naturalistic (physicalist) epistemology of different levels of reality (where entities of level L_n are composed of entities of level L_{n-1} whose properties are independent of the emerging level L_n) must therefore be considerably deepened.

Regardless of these difficult and profound epistemological debates, micro → macro emergence phenomena are *empirically* observed on a massive scale in many different disciplines and must therefore be taken into account. This is the challenge that CMI aims to tackle. It has succeeded in many cases, and whenever it has been able to access the generating mechanisms, it has found that emerging macro structures exhibit a certain *universality* (multi-realizability) in relation to their micro implementation.

In 1981, Jean-Pierre Dupuy and Paul Dumouchel organized a symposium at the International Cultural Center of Cerisy-la-Salle in France, which became a memorable reference. It was entitled *Self-organization: from physics to politics* (Dupuy-Dumouchel 1981). We could paraphrase it and title this second part *CMI: from physics to politics*. Indeed, in agreement with (Dupuy 2004), we believe that CMI is a vector for "the unity of the social sciences."

We will therefore use the terms CMI and methodological holism (MH) in all situations where collective macro properties emerge from interactions between micro constituents, whatever they may be, including atoms, molecules, or grains of sand. "Individuals" do not necessarily have interiority and intentionality. It is sufficient that there are rules governing their behavior, and even physical laws are such rules. A particle moving in a field can be considered an "agent" acting according to rules, even if that agent does not intentionally control its actions. CMI is always agent-based, an agent being a unit acting according to rules and "motivations" described by the rules, which can be extremely varied.[2] In all cases, it is a scientific challenge to overcome the mereological antinomy by performing a *computational synthesis* of macro properties from micro properties.

We can therefore immediately see what CMI is not.

(i) CMI is not a reductionist and eliminativist "atomism" since it focuses on the macro consequences of micro behaviors in order to understand how the emerging level of reality becomes *autonomous*.[3] It is a "nonreductive individualism" in the sense of (Sawyer 2003).

(ii) By definition, CMI is also not a semantic reductionism, since its pre-

[2] On the close links between MI and agent-based models, see (Manzo 2023).

[3] In (Petitot 2010) we give an example of how global *geometric* properties of visual percepts can emerge from neural interactions and, as *geometric* properties, cannot be interpreted in a reductionist way.

cise function is to explain the transition between conceptually incommensurable levels. In a study with Shu-Heng Chen, "Agent-Based Simulation as an Implementation of Methodological Individualism" (Di Iorio-Chen 2019), one of us[4] explained in particular the 1955 debate between Maurice Mandelbaum and John Watkins. According to the former, MI was a semantic reductionism, whereas such a reduction seems impossible. But the argument was biased, since MI has never been a semantic reductionism. Inspired by statistical thermodynamics, its challenge has always been, from the outset, to find the tools to derive macro properties *computationally* from micro properties using tools from probability theory and complexity theories.

(iii) CMI is also not an "ontological individualism" insofar as the difference between micro and macro levels is always *relative*. In sociology, micro states may be the mental states of subjects. But for cognitive neuroscience, these mental states are macro states emerging from micro neural dynamics.

(iv) In the social sciences, CMI is also not a psychological reductionism, which would oppose psychology to sociology. The intentional states of subjects can incorporate any social constraints one wishes. In economics and game theory, CMI has no reason to be preferentially associated with rational choice theory concerning agents whose rationality consists in optimizing their utility. What matters for MI is that individual behaviors are governed by rules, whatever they may be. Altruism and cooperation can obviously be part of this.

(v) However, in the social sciences, since micro units are, let us repeat, agents or subjects endowed with interiority and behavioral "dispositions," CMI obviously requires appropriate assumptions about their internal structure if we want to understand and explain causally the overall collective properties that may emerge from their interactions. Enzo di Nuoscio in "Methodological Individualism, Scientific Explanation, and Hermeneutics" (Di Nuoscio 2023) and Gerald Bronner in "Why Does Comprehensive Sociology Need to Be Enlarged?" (Bronner 2023) have clearly explained this point by referring to the logic of actions developed by Raymond Boudon in his comprehensive sociology and his conception of praxeological rationality based on the situational and contextual analysis of the "good reasons" of subjects (see Boudon 1986). Di Nuoscio unifies von Mises' more rationalist point of view and Boudon's more empiricist one in what he calls the "Mises-Boudon model." As for Bronner, he explains that subjective assumptions about the internal structure of subjects (with their hermeneutic and interpretative dimensions) can be clarified by contemporary cognitive neuroscience without any neurophysical reductionism that would eliminate the question of meaning in the mind.

(vi) It goes without saying that CMI therefore has nothing to do with what is known as "individualism" in the sense of a political, cultural, or ethical choice of behavior, as a maxim for action or a way of being. It is only a methodological choice for scientific explanations of collective phenomena.

To conclude this first section, let us return to the *intrinsic interdisciplinarity* of CMI. It is essential from an epistemological and methodological point of

[4] F. Di Iorio.

view. Indeed, while using concepts with strong physicalist connotations such as "micro," "atomism," "macro," "emergence," "interaction," "causality," etc., most of the criticisms levelled against CMI in the humanities authoritatively assert theses that are in no way validated by the physical, chemical, or biological versions of CMI. It is therefore essential to refer back to the sources and fundamental examples of the spectacular success of CMI in physics. These have taught us that it is *complexity* that lies at the heart of the challenges of CMI.

1.2 Complex systems

CMI is by definition "systemic" and only makes sense beyond a certain level of internal complexity. Talking about CMI means immersing MI issues in social, economic, and political sciences in the much broader context of complex systems sciences. Economic transactions, cognitive processes, individual intentions, actions, agent interactions, beliefs, axiologies, norms, institutions, and even semiotic and hermeneutic dimensions are inserted into a broader paradigm ranging from the natural sciences to NBIC (nanotechnology, biotechnology, information technology, cognitive science) and the social, economic, and political sciences.

The establishment of this global interdisciplinary field represents one of the major scientific advances of the second half of the 20th century. A new world emerged from it. It goes back to 1940-1950 with John von Neumann, Norbert Wiener, Warren McCulloch, Walter Pitts, and the famous Macy Conferences (1942-1953) at the origin of cybernetics, cognitive sciences, and information sciences (see Jean-Pierre Dupuy's book (Dupuy 2000) *The Mechanization of Mind. On the Origins of Cognitive Science*).

The increasingly in-depth and formalized understanding of self-organization and critical phenomena then led many scientists with dual training in statistical physics and social sciences to relaunch such a research program (known as "second cybernetics"). Many research institutes were then developed. In the 1980s, one of the best known was the *Santa Fe Institute for Complex Systems*, founded in 1984, which brought together such important figures as Per Bak (self-organized critical states), Jim Crutchfield (nonlinear dynamics and chaos, critical phenomena, statistical inference, and machine learning), Murray Gell-Mann (elementary particle physics, 1969 Nobel Prize for quark theory), John Holland (genetic algorithms), Stuart Kauffman (Boolean neural networks, out-of-equilibrium self-organization), Brian Goodwin (structuralist behavioral biology), or Christopher Langton (artificial life). Many other institutes of this type were subsequently created, such as the *MIT Media Lab* (with Alex Pentland) and, in Paris, the *Institute for Complex Systems* founded and initially directed by Paul Bourgine and directed now by David Chavalarias, who, in his article "From inert matter to the global society" (Chavalarias 2020), explains that the study of complex systems constitutes

> "a paradigm where it is possible to go up the physical levels of organization of matter, from physics to biology and society, without

resorting to strong reductionism."

1.3 Computational syntheses and conceptual analyses

It is very important to emphasize that CMI is based on *computational syntheses* and not conceptual analyses. To quote the old Kantian adage: "philosophy is knowledge through concepts and science is knowledge through the construction of concepts," (where "construction" means mathematization), we could say that MI concerns scientific knowledge and not philosophical knowledge. This is undoubtedly the reason for the many ambiguities in philosophical debates surrounding MI: they are argued without taking formalisms into account, which makes them aporetic.

Conceptual analysis starts from empirical data gathering and presents them in a rational way. Through successive conceptual abstractions and "subsumptions," it arrives at Categories (in the philosophical sense of the term: causality, etc.) and Principles (conservation laws, etc.). But modeling in formalized sciences actually solves the *inverse problem* of conceptual analysis. It is a computational synthesis that translates concepts into algorithms (for instance, differential equations as in Newton's Mechanics) in order to perform a *virtual reconstruction* of empirical data. Then algorithmic analogies can appear where there are no conceptual analogies.

The intrinsic interdisciplinarity of CMI is therefore not a matter of vague semantic metaphors between heterogeneous domains of reality. It concerns *computational* equivalences between the formal reconstructions of different empirical phenomena.

Such formal intertheoretical isomorphisms are fundamental to the natural sciences. The most striking classic example remains Newton's equivalence, for phenomena of gravitational attraction, between the "supralunar" world of planetary motions and the "sublunar" world of ballistic trajectories, worlds that at the time had been completely separate ontologically for centuries and corresponded to unrelated technologies and professions: on the one hand, the sophisticated instruments of Tycho Brahe and the astronomical calculations of Kepler (in his *Tabulæ Rudolphinæ*) for planetary motions, and on the other, the military engineering of artillerymen for ballistic trajectories. It was the isomorphism of formalisms that made it possible to imagine universal gravitation and to understand that planets "fall" toward the sun just as apples fall to the ground.

It is essential to bear in mind that the computational reconstruction of empirical reality is not at all identical to conceptualization and instead solves the inverse problem of conceptual analysis. It is primarily in relation to the assumptions underlying their computational models that the most formally powerful, experimentally confirmed, and operational theories can pose, and generally do pose, extremely delicate conceptual problems. To continue with Newton's ex-

ample, his theory of gravitation, which is admirable and still allows us to calculate planetary and satellite trajectories today, is nevertheless based on the hypothesis of forces acting instantaneously at a distance, which is conceptually unacceptable. Hence the fierce debates of the time and, in response, the famous "hypotheses non fingo" of the *General Scholium* of the second edition of the *Principia*[5] in 1713, "hypothesis" meaning conceptual speculations not justified by phenomena. It was not until the advent of general relativity that the situation was resolved.

In this second part, CMI is therefore treated as a computational synthesis program, which is precisely why it is "methodological." What matters is that it must be operational. From this perspective, questions of substantialist ontology are of little relevance. For example, we will discuss below the formal isomorphism between magnetic spin glasses and neural networks. It is hardly relevant to point out that neurons are not atoms with spins. What matters is that the transfer of spin glass computations to neural networks has enabled the explosion of artificial intelligence technologies that are revolutionizing all human practices. The epistemology of CMI is more a techno-scientific epistemology of formalized models than a conceptual epistemology of a philosophical nature.

To this we must add that, in the very specific case of the social sciences, all these classic epistemological problems (very old, studied for centuries, and very difficult) present an additional crucial difficulty. For, as Gustav Ramström emphasized in "The Analytical Micro-Macro Relationship in Social Science and its Implications for Individualism-Holism Debate" (Ramström 2023), there is no macro *phenomenality* that is proper, *autonomous*, and manifest as such in the social sphere. There are, of course, social phenomena, but

> "Social phenomena do not have macro-phenomenal expressions" (p. 488).

A property P of a macro entity E can refer to E as a whole (e.g., "a diamond is hard") or to E as a composite entity ("a diamond is a crystalline network of carbon atoms").[6] However, in the social sciences, social properties are, at the conceptual level, *analytically* "micro-based".[7] They refer analytically (in the conceptual sense) to underlying micro multiplicities and compositions. Ramström emphasizes this point: whereas in the natural sciences there is generally a phenomenal micro-macro dualism, it is rather the "hypothesis" of a micro level that is at issue[8] and it is the discovery of a micro phenomenality that constitutes a scientific revolution[9], in the human sciences macro phenomenality is given from the outset as that of the collective properties of a micro

[5] *Philosophiæ naturalis principia mathematica.*

[6] Below, we will give several examples of this type: the thermodynamic properties of gases, the geometry of dunes, the shapes of flocks of birds, etc.

[7] For a discussion of the relationship between analytical sociology and MI, see (Bulle-Phan 2017).

[8] For example, the atomist hypothesis from the Greeks and Romans Democritus, Epicurus, Lucretius to Gassendi, Boyle, Newton, Lavoisier, Dalton, Mendeleev, then to Boltzmann, who was so attacked that he committed suicide in 1906.

[9] For example, the empirical evidence for the "reality" of the atomist hypothesis obtained by Jean Perrin (Nobel Prize 1926) through the analysis of cathode rays, Röntgen X-rays, and

phenomenality. This allowed Jon Elster to open his article (Elster 2023) with the statement

> "I believe that methodological individualism (MI) is trivially true." (p. 117)

1.4 An example of transdisciplinary universality: Pareto's laws

There are well-known precursors to this type of transdisciplinarity. A particularly famous and ubiquitous example in the social sciences is Pareto's laws. We will refer to it several times.

Vilfredo Pareto (1848-1923), Léon Walras' successor in Lausanne, wrote his engineering thesis in Torino on the equilibrium of solid bodies. He provides a remarkable example of the transfer of concepts from a physical domain to a socio-economic domain (see Pareto 1885). His theory of equilibrium is certainly not a good general theory of economics, but it shows that a fragment of a theory of idealized and simplified economic situations can recover the same formalisms as phenomena in other domains such as physics.

As a function of the quantities of goods X_i utilty $U(X_i)$ closely parallels the internal energy of a thermodynamic system, and the partial derivatives $\partial U/\partial X_i$, which are the marginal utilities of the marginalist school of Léon Walras, Stanley Gevons, and Carl Menger, closely parallel thermodynamic coefficients such as specific heat.

Introduced by Pareto in 1895 in relation to the distribution of incomes, Pareto's law states that the proportion $F(x)$ of individuals receiving an income $\geq x$ is $(x/x_0)^{-\alpha}$, where x_0 is the minimum income and the exponent $\alpha > 1$ is the Pareto index. Since $x/x_0 \geq 1$, $F(x)$ is a decreasing function of α. The average income m is $m = (\alpha/(\alpha-1))x_0$. From these formulas and the differential relation $dm = (\alpha/(\alpha-1))dx_0 - x_0/(\alpha-1)^2 d\alpha$, Pareto deduced a whole set of statements about the concentration of wealth and inequality in society. For example, since the distribution of wealth is unequal, in order to enrich the poors, the overall wealth of society must grow faster than the population.

The validity of these theories has been hotly debated and their model has been gradually modified to the present day. The work of Max Lorenz and Corrado Gini, for example, has led to a deeper understanding of inequality models (see e.g. Barbut 1984). The probability $P(X < x)$ is equal to $1 - (x/x_0)^{-\alpha}$. An interesting function is the Lorenz function $L(p)$, which expresses the cumulative percentage of total income held by a cumulative proportion p of the population (individuals being ranked by increasing income). For Pareto's

---

Brownian motion, as well as the calculation of Avogadro's number N_A. Equal to 6.022×10^{23}, Avogadro's number is the fundamental converter from micro to macro in physics and chemistry. It is a universal physical constant that gives the number of particles (ions, atoms, or molecules) in one mole of matter, i.e., approximately (neglecting electrons) the number of nucleons (protons and neutrons) in $1g$ of matter, the mass of the nucleon being $1.67.10^{-24}g$.

law, it is given by the formula $L(p) = 1 - (1 - p)^{(\alpha-1)/\alpha}$. Economist Bernard Zimmer pointed out that the Lorenz function is *formally* the same as that of a thermodynamic model where competition is equivalent to collisions between particles, income to altitude, the Pareto index to temperature, and cumulative percentages to pressure.

What interests us here is that Pareto's laws are typically transdisciplinary. They concern random variables X that can take real values $x \geq x_m > 0$ and are of the form $P(X > x) = (x/x_0)^{-\alpha}$ where P is the probability and α is the Pareto index. We encounter it elsewhere, along with other similar power laws such as Zipf's law (introduced by George Zipf in 1949) or Gibrat's law (introduced by Robert Gibrat in 1931), when analyzing the distribution of city sizes, the sizes of beach cobbles, or the frequency of lexical items in a linguistic corpus. In each case, the micro-mechanisms that generate these emerging formations and hierarchical structural evolutions are specific, but the fact that the distribution follows a Pareto or Zipf law reveals a certain universality of *formal* structures. Once again, these mechanisms cut across regional ontologies and range "from the physical to the political."

1.5 Reminders about some debates concerning MI

Let us repeat that in the social sciences[10], much of the debate concerning MI stems from the fact that they are dealing with units that are *subjects* and therefore individual agents endowed with interiority, intentionality, maxims of action, and freedom.[11] Since Carl Menger (1840-1921) and the marginalist revolution, Georg Simmel (1858-1918) and his theory of the forms of socialization of individual interactions, and above all Max Weber (1864-1920), the great representatives of MI such as Ludwig von Mises (1881-1973, subjective conception of value), Friedrich von Hayek (1899-1992, we will come back to this), Karl Popper (1902-1994, critique of historicism as a finalist and prophetic conception of history), and Raymond Boudon (understanding of individual "reasons" and "composition effects" emerging from the aggregation of individual actions with *unintended* consequences often opposed to the intentions of the agents), MI in the social sciences is *interpretative*, even "hermeneutic," insofar as it is based on motivations for intentional actions that have unintended consequences. The transition from micro to macro is precisely a matter of these unintended consequences which, as Hayek often insisted, are inaccessible to the rationality of subjects and transcend their cognitive abilities.

This micro/macro incommensurability does not raise philosophical problems for social insects, for example. It is easily accepted that bees do not have the capacity to understand the hexagonal geometry of their honeycombs and that

[10] See the first part of the book, chapter 1.

[11] An excellent summary of some of the key moments in MI and CMI (Menger, Mises, Hayek, Weber, Popper, Morin, Dupuy, Boudon) can be found in Alain Laurent's book (Laurent 1994). Another good introduction to the "background, history and meaning" of MI is (Udehn 2001).

this geometry is "unintentional." But it does pose a problem for the social sciences because it brings into play different conceptions of subjectivity.

There are several types of recurring criticisms of MI. A good example is provided by Geoff Hodgson's excellent analysis (Hodgson 1986). Some argue that MI does not take into account the top-down effects of macro structures on micro behaviors. These criticisms are not valid, since all MI models, from thermodynamics to Hayekian conceptions, take them into account extensively. The real question is not that. Everyone agrees that there is autonomy between levels of reality and that collective concepts such as state, market, class, etc. are operational. The question is in fact whether it is justified to *hypostatize* these collective macro concepts into substances causally active as *substances* on micro agents by controlling them. The MI criticizes this substantialist realism, which submits the autonomy of agents to the heteronomy of a "subjectless" determinism.

Other holistic criticisms of MI consist, in order to explain collective actions, of transferring concepts from the micro level to the macro level. An attempt is made to overcome the semantic incommensurability of the levels by collectivizing micro concepts. This is the case with a concept such as "collective intentionality" in John Searle (see Searle 1995) and other authors who, each in their own way, seek to "collectivize" cognitive and affective mental states, whereas intentionality seems at first glance to be a concept that *analytically* refers to a subject, whether psychological or phenomenological (in the Husserlian sense); see for example (Greve 2023). It is clear that belonging to groups, communities, institutions, and other social actors, as well as mimicry, consensus, conventions, institutionalized norms, altruistic moral imperatives, etc., can be distributed among individual intentionalities and thus induce macro-intentionalities. But this does not mean that such macro-intentionalities can be that of a kind of macro-subjectivity.

It can be shown that in language, the *spontaneous* attribution of *intentionality* is a common phenomenon that has been well known and analyzed for a long time, for example in the visual perception of scenes. Perceptual inferences lead to spontaneously ascribing to moving inanimate objects the semantic roles of animate and intentional subjects. The first careful experiments were conducted by Fritz Heider and Marianne Simmel (Heider-Simmel 1944). They studied how the motion (including speed, acceleration, deceleration, change of direction, etc.) of simple completely inanimate shapes (e.g., two triangles, one circle and one bigger rectangle) were spontaneously interpreted as intentional actions verbalized by verbal lexemes such as 'enter', 'exit', 'hide', 'escape', 'run away', 'chase', 'attack', 'give', 'let go', 'force inside', etc. A purely kinematical motion can thus be described by an elaborate agent-based verbal proposition.

We can therefore consider that collective intentionality is simply an attribution of intentionality to collective entities, an attribution that allows us to describe certain social phenomena simply in terms of common sense and "folk sociology" without this in any way implying that these collective "intentionalities" possess true intentionality.

This is why MI criticizes MH in return, insisting that intentionality is not

limited to being an operational concept of common sense, but is a deep psychological process that exists only under certain conditions and refers to experiences filled with existential meaning and to internal subjective states that are the cause of behavior.

It should be noted, however, that MI could be taken to a much more radical level by criticizing psychological intentionality itself as a concept belonging to common sense and "folk psychology" and having no real scientific content despite its great practical usefulness. This would bring us back to the materialistic naturalism of Dan Dennett who, with his concept of the "intentional stance," explains that any intentional state is only apparently intentional and must be reduced to purely functional mental states implemented neurophysiologically. Dennett's argument against the notion of a consciousness that unifies the various highly modularized mental acts of a subject in representations is very similar to that of MI against MH.

1.6 Towards modeling

1.6.1 A classic example that is off-topic but very instructive

After these general epistemological and methodological remarks, and before addressing technical examples of CMI in various natural sciences, let us say a few words about modeling. We said earlier that CMI is not primarily a problem of "subsumption" of empirical diversity under conceptual unity (whether empirical or theoretical) – a conceptual subsumption to be analyzed philosophically – , but first and foremost a problem of "computational synthesis" generating operational models of observed empirical reality. It is therefore inseparable from centuries-old (even millennia-old) reflections on the mathematization of natural phenomena.

The literature on this subject is immense, from the astronomical models of the Greeks (Eudoxus, Hipparchus, Ptolemy) to Copernicus, Tycho Brahe, Kepler, Newton, Lagrange, and up to the present day, a century after Poincaré introduced qualitative dynamical methods for studying the gravitational N-body problem ($N > 2$, $N = 2$ is the Keplerian problem completely solved by Newton).

Let us consider this example for a moment, even though it is apparently outside the scope of our discussion, since classical planetary systems are single-level and involve only a few interacting bodies. However, it already contains many ingredients of model epistemology and has the advantage of being universally known. In a study (Petitot 2023b) based in part on historical works such as Pierre Duhem's famous book (Duhem 1908) *Sôzein ta phainomena. Essai sur la notion de théorie physique de Platon à Galilée*[12], and, on the other hand, Kant's

[12] *Essay on the notion of physical theory from Plato to Galileo.* "Sôzein ta phainomena" was what Greek astronomers called "saving the phenomena," that is, correctly mathematizing, by using basic mathematical and physical principles and hypotheses, the empirically observed "wandering" motions of the planets.

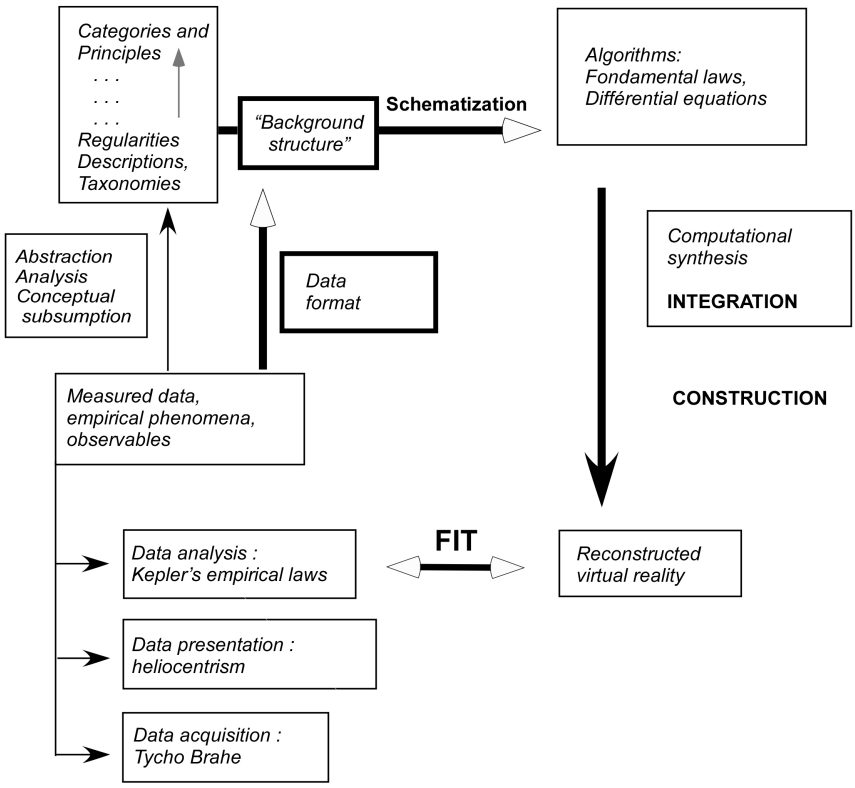


Figure 1.1: The dialectic of conceptual analysis and computational synthesis in the mathematical modeling of empirical phenomena.

Metaphysical Foundations of Natural Science (Kant 1786), we have summarized the main moments of the Keplerian-Newtonian mechanics in the following diagram (figure 1.1).

(1) The left column is that of conceptual "subsumption." It "goes up" from observable empirical diversity to conceptual unity. It begins first with data processing: (a) precise acquisition methods involving high-tech observation instruments (such as those constructed by Tycho Brahe), (b) pre-processing computations that reveal patterns often not apparent in the raw data (such as Kepler's magnificent recalculation of Tycho Brahe's planetary position tables in a Copernican heliocentric reference frame), (c) the discovery of empirical laws governing the regularities revealed by pre-processing (it is from the regularities in the "wandering" planetary trajectories recalculated in an heliocentric frame that Kepler became able to derive his three famous laws).

(2) This yields observed, measured, processed, and standardized data that can, at each level, be subsumed under abstract concepts that can be analyzed conceptually.

(3) Conceptual abstraction also has several levels, ranging from empirical descriptions, classifications, and the conceptualization of regularities to universal categories and principles of scientific explanation such as causality.

(4) In the case of the physical sciences, there is an additional element that is essential for modeling: phenomena (the "wandering" trajectories) are given and observed in a particular "framework," a format, namely *space-time*, which is called a "background structure." This structure can be *mathematized* and has a very rich structure: Euclidean geometry, symmetry groups and relativity groups, differential structure.

(5) Points (1)-(3) in the left-hand column of *conceptual analysis* of empirical diversity are common to *all* sciences.

(6) What is specific to modeled sciences is the *right-hand* column of *computational synthesis*. It involves the algorithmic *reconstruction* of a *virtual* diversity that can be *compared* (fitted) to the *given empirical* diversity.

As we have already mentioned above, this computational synthesis solves the inverse problem of that of conceptual analysis. It "goes down" from fundamental equations to a reconstructed phenomenal diversity. To generate diversity (which is the inverse of "subsumption"), it must use the *generativity* of mathematics, i.e., appropriate algorithms. For the period considered in our example, this involves integral-differential calculus, which allows to *integrate* the basic differential equations expressing physical laws (in particular Newton's universal law $f = m\gamma$ linking forces f and accelerations γ through masses m).

Differential equations are the *infinitesimal generators* of motions, and the reconstruction of trajectories amounts to *iterating* these generators. Today, this can be done on a computer, resulting in "numerical" integration. In particularly simple cases, *explicit formulas* can be obtained for the solutions, and the system can then be explicitly integrated.

It cannot be overemphasized that it is *computational generativity* (synthetic and not analytical in the conceptual sense) that radically distinguishes formalized sciences from simply descriptive and conceptual sciences. Without it, it is impossible to construct a modeled virtual reality that can be compared to a given empirical reality.

(7) A crucial problem is then to link the two columns together. Our diagram does this at the top and bottom. At the bottom, the real empirical data (on the left) are compared with the virtual reality generated by the models. This is a problem of *fit*. In our example, it is remarkable and even often considered "miraculous." To be convinced of this, one need only consider (e.g.) that the trajectory of the ESA (European Space Agency) Rosetta Spacecraft launched in 2004 towards the comet 67P/Churyumov-Gerasimenko was computed with Newton's equations. It implied "gravitational slingshot" maneuvers and flybys of planets (three of Earth and one of Mars) to reach a highly elongated orbit whose apogee is close to the orbit of the giant gas planet Jupiter, about five hundred million miles from the Sun.

(8) But the most profound and interesting relationship between the two columns is the one at the top between categories and principles on the conceptual analysis side and algorithms and differential equations on the computational synthesis side. It essentially uses the possibility of *schematizing* (in a sense inspired by Kant's transcendental schematism) categories and transforming them into principles by means of symmetry groups (i.e., relativity groups) of the background structure that formats phenomena. These are technical issues that it would not be appropriate to discuss here, but interested readers may consult, for example, our 1992 study "Actuality of Transcendental Aesthetics for Modern Physics" (Petitot 1992).

(9) The example of planetary models is very instructive for what can be the possible complexity of a computational synthesis. In the case of a Keplerian gravitational system with two bodies (a large sun and a small planet, the sun being considered fixed), we know how to integrate the system explicitly, and it is easy to recover Kepler's laws. But as soon as we consider a three-body system (a sun (S), a planet (P) and a satellite (L)), the system is generally no longer integrable and, even if we approximate the sun–planet system (S, P) by a Keplerian system, the satellite L can have a very complicated trajectory, such as (i) an initial motion orbiting around S but in an eccentric and irregular manner due to the significant attraction of P, (ii) a sudden shift causing L to orbit around P with irregular epicycles, then (iii) a return to orbiting around S again, etc. An N-body gravitational system is generally so complex that it is non-integrable and can only be integrated numerically. As for its structure, it can only be studied *qualitatively* using the qualitative dynamical tools introduced by Poincaré (and already mentioned above).

(10) The example of planetary models thus gives a good idea of the problems raised by computational syntheses. It seems obviously still very far from the problems we will encounter in the following sections, since it does not seem to involve micro-macro dialectics. Note, however, that such dialectics may still exist in complex gravitational systems such as the rings of Saturn. The micro level consists of a myriad of small pieces of ice (and dust) in gravitational interaction, and the macro level consists of the geometric properties of the rings. These rings are located in an extraordinarily thin disk (less than ten meters thick!) measuring $360,000\ km$. They were observed in detail by NASA's Cassini probe, launched in 1997. It is amazing to see how they divide the disk into several (7) main rings separated by empty bands and how they exhibit "gaps" caused by resonances with Saturn's various satellites. Their dynamics are fascinatingly sophisticated and exhibit all the characteristics of complexity that we will encounter later in our other examples. The computational synthesis of their gravitational interactions represents an extraordinary deepening of that of planetary and satellite motions (with the added effects of electromagnetic forces due to Saturn's structure).

1.6.2 The Coleman's boat

Insofar as the social sciences are rarely mathematized, such reflections are rare, but we would nevertheless like to mention a common reference in the literature dealing with MI in these fields, that of "Coleman's boat," which schematizes the structure of MI models according to James Coleman (Colemen 1987). In "Methodological Individualism and Micro-Macro Modeling", Karl-Dieter Opp (Opp 2023) has discussed the epistemology of the models in question. Reference can also be made to Raub- and van de Rijt (2023).

In a CMI model, we have, let us repeat, (at least) two micro/macro phenomenal levels, each of which has, on the one hand, its own phenomenality, its own tools of observation and description, as well as its empirical regularities and, on the other hand, its theoretical explanations and more or less detailed formalizations. The basic hypothesis is that, by means of "bridge assumptions"[13], macro *descriptions* can be *deduced* from micro *models* through explicit transformations (e.g., probability theory and statistical averages). The central problem is that, even if in the social sciences macro entities are analytically micro-based, the *conceptualities* – the theoretical vocabularies, variables, processes, and laws – used to describe the properties of the two levels remain *incommensurable*, whereas we want, on the contrary, to link mathematically and algorithmically their two formalizations.

The paradox is so obvious that it raises a real challenge.

What we explained above about the "inversion" between conceptual analysis and computational synthesis becomes unavoidable here. Indeed, since computational syntheses are based on long-range mathematical *generativity*, which can be highly *trans*-theoretical, they can *mathematically* link *conceptually* very heterogeneous theories. In other words, it is rational to seek to "factorize" macro computational synthesis through micro computations using bridge assumptions, even if there is no corresponding factorization of conceptual analyses.

Here we must make a philosophical remark about mathematics that we consider decisive for the epistemology of MI. Since the development of logical positivism, many philosophers have adhered to the thesis of the "analyticity" of mathematics, despite the opinion of most creative mathematicians who have always insisted on the contrary on its *synthetic* character. "Syntheticity" means here that there are many formalized bridges connecting very different mathematical theories, for example, algebraic number theory and geometry or topology and algebra. For readers familiar with mathematical "category theory", let us say that there are many "functors" between different categories.[14] The standard example is that introduced by Poincaré, the homotopy group $\pi(M)$ of a topological space M, an algebraic structure that counts the number of 'holes' in M. It was the origin of "algebraic topology."

Thanks to these bridges, functors, topos, etc. the notion of deduction has

[13] Once again, let us repeat that the founding example is in thermodynamics, where a bridge assumption interprets the macro variable of gas temperature as the statistical average of micro molecular vibrations, a hypothesis verified every day in every home by microwave ovens.

[14] More recently, another type of bridge has been developed in what is known as *topos theory*.

been considerably enriched since, in addition to *intra*-theoretical deductions remaining within the axiomatic system of a given theory, it is possible to develop *inter*- or *trans*-theoretical deductions passing from one proof path in a theory T to another theory T' to make another piece of proof path in T' and then return to T by another bridge. By iterating such local pieces of proof and using bridges, proofs can *navigate* in an increasingly complex and *global* way within the *unity* of the mathematical world, rendering the thesis of analyticity completely obsolete. Interested readers can consult our compilations of two exemplary mathematical circumnavigations, that of Andrew Wiles' proof of Fermat's theorem (Petitot 1993) and that of Alain Connes' partial results in the proof of the Riemann hypothesis (Petitot 2022).

Now, although invalidated by the trans-theoretical circumnavigations of major proofs, the analyticity of mathematics defended by logicism has often been interpreted by philosophers as *conceptual* analyticity, leading to a profound misinterpretation of bridge assumptions. As a result, the possibility of demonstrating the existence of microfoundations for macro properties and behaviors tends to be interpreted as semantic reductionism, which is not the case at all.

In short, in our view, an "analytic" interpretation of computational syntheses, when these are precisely the opposite of conceptual analyses, makes MI philosophically unintelligible.

Chapter 2

CMI in Physics

2.1 Macroscopic thermodynamics (MH)

2.1.1 The relevance of returning to physics

As we have said several times, the first, and most celebrated, example of CMI is provided by the statistical interpretation of *thermodynamics*. To present it pedagogically, we will revisit the basics found in classical treatises[1], university textbooks, and Wikipedia-type articles. It initially concerned the purely macro level (MH) and then developed its micro interpretation with statistical physics (CMI).

Thermodynamics was invented by (Nicolas Léonard) Sadi Carnot[2] (1796-1832) whose 1824 treatise focused on the theory of heat acting as a "driving force" (i.e., as energy) in steam engines. It was then developed by great physicists such as Emile Clapeyron (1799-1864, with the key publication Clapeyron 1834), William Thomson (Lord Kelvin, 1824-1907), Rudolf Clausius (1822-1888), and then, as statistical physics, by James Clerk Maxwell (1831-1879, from 1859), Ludwig Boltzmann (1844-1906, with the key publication Boltzmann 1872), and Josiah Willard Gibbs (1839-1903) who was mainly interested in chemistry and published his great treatise *Elementary principles in statistical mechanics* (Gibbs 1902) in 1902.

Thermodynamics made it possible to exemplify and test a vast set of methodological, theoretical, modeling, and even ontological hypotheses. It is also a remarkable example of the difficulties and controversies encountered in the development of a new theory. Sadi Carnot used Lavoisier's caloric theory, which treated heat as a fluid substance capable of acting causally. It was a decisive step to replace it with the universal concept of energy, with heat then becoming interpretable as a "thermal mode" of energy, just as kinetic energy is a "mechanical mode". The different modes are able to transform into one another

[1] Among these, Poincaré's *Leçons* (Poincaré 1908) are particularly interesting.

[2] The son of scientist and statesman Lazare Carnot.

under the constraint of the principle of conservation of total energy.[3]

We will see that statistical thermodynamics provided the first explicit equivalences between macro and micro descriptions, and that it is obviously paradoxical that CMI, justified by one of the greatest successes in physics, a progress whose applications are innumerable and universally verifiable, can nevertheless be considered by some philosophers as philosophically inconsistent.

We will first limit ourselves to the simplest and most phenomenologically obvious cases, those of simple substances such as water that can exist in qualitative states, known as "phases," such as solid (ice), liquid, and gas (vapor). These phases are phenomenological evidences, and even though atomistic hypotheses had existed since antiquity, they were inapplicable to them at the beginning of the 20th century and therefore purely speculative. Before Jean Perrin's work, atomism was only a "hypothesis."

In other words, as we have seen, while in the humanities it is the data on individuals that is obvious and the theorization of their collective "holistic" properties that needs to be constructed, in thermodynamics the opposite was initially true: collective properties were clearly evident, and it was their correlation with the "individual" properties of supposedly unobservable micro-constituents that had to be constructed from scratch. It is therefore ironic for history of science to note that where MI was not phenomenologically self-evident, it was spectacularly successful, while where it is phenomenologically self-evident, it is strongly contested at the very level of its principles.

2.1.2 Description of the "holistic" thermodynamical macro level (MH)

We will now give an overview of basic formal entities. They are elementary and only involve school concepts:

- derivable functions of several variables of the type $f(x, y, z)$;
- partial derivatives of the type $\partial_x f(x, y, z)$ symbolizing (if it exists) the partial derivative of the function f with respect to the variable x.[4] If the context imposes the constraint $z =$ constant, we write it as $\partial_x f(x, y, z)_z$;
- the exponential function $exp(x)$, which is the exponential e^x with base the Euler constant $e\tilde{}2.71828$ and which has the remarkable property of being equal to its derivative;
- the natural logarithm function $log(y)$ for $y > 0$, which is the inverse function of the exponential, i.e. $log(y) = x$ is equivalent to $y = exp(x)$;
- the factorial function $n! = 1 \times 2 \times ... \times n$;

[3] It is as if energy were to the physical world what money is to the economic world.

[4] $\partial_x f$ means $\frac{\partial f}{\partial x}$.

- the symbol $\sum_c$ which denotes the sum of terms indexed by an integer c varying within a certain interval;
- the gradient $grad(f)$ of a function of several variables, which is the vector whose components are the partial derivatives of f;
- differentials of type $df(x,y) = \partial_x f(x,y)dx + \partial_y f(x,y)dx$;
- Leibniz's rule $d(fg) = fdg + gdf$.

To describe macroscopically the state of a substance like water, assumed to be at thermodynamic equilibrium, one uses descriptive concepts and measurable quantities such as absolute temperature T (in $^\circ$ Kelvin[5]), pression P, volume V, and entropy S. Entropy here is a measure of the *homogeneity* of the system: it is at its maximum when the temperature is the same everywhere, when the concentrations of the various components of a mixture are the same everywhere, etc.

T and P are *intensive* state variables (i.e. independent of the size of the system), while V and S are *extensive* state variables (i.e. additive for subsystems). Each intensive variable is "conjugate" to an extensive one, T with S and P with V. These four variables are not independent, and T, P, and V are generally chosen because they are easy to manipulate experimentally.

A major discovery, which required countless observations, remarkable technological innovations in measuring devices, and new methods of organizing and presenting measurement results, was that, when the system is at equilibrium, these three parameters are related by a *state equation* $\Phi(T,P,V) = 0$, defining a state surface Σ_3 in the 3-dimensional (T,P,V)-space. If we take into account the relationships with S, Σ_3 can be embedded as a surface Σ_4 in the 4-dimensional (T,P,V,S)-space.

The concept of a state *function* is subtle and not at all self-evident. In fact, what we measure experimentally are *variations* δM in certain magnitudes that depend on state variables. To say that M is a state function is to say that an infinitesimal variation δM is an *exact* differential dM, in other words that $dM = AdT + BdS + CdP + DdV$, where the coefficients are the respective partial derivatives $A = \partial_T M$, $B = \partial_S M$, $C = \partial_P M$, $D = \partial_V M$. Since the second derivatives are *symmetric*[6], we necessarily have $\partial^2_{TS} M = \partial^2_{ST} M$ and therefore $\partial_T B = \partial_S A$, etc. These strong constraints imply that, if a transformation along a temporal path $\gamma(t) = (T(t), S(t), P(t), V(t))$ leads from an initial state $M_0 = M(t_0)$ at t_0 to a final state $M_1 = M(t_1)$ at t_1, then the variation ΔM, which is the integral of dM along the path γ, is *independent* of the specificity of γ between M_0 and M_1 and simply equal to the difference $M_1 - M_0$

What's more, deep *analogies* were gradually discovered between the state surfaces of different bodies, which showed that there must be underlying *universal* processes. This was a problem of *multirealization* of thermodynamic

[5] 0° Celsius = 273,15° Kelvin.

[6] This is a fundamental property of differentiation. It may no longer necessarily hold true if specific constraints are imposed on the coordinates.

properties by different material substances.[7] It was therefore the differential *geometry* of the state surface Σ that contained the essential information. In particular, it includes the position and orientation of the tangent planes in the surrounding space as well as the apparent contours of Σ on the coordinate planes, the apparent contour being the projection onto the coordinate plane of the "critical locus" of Σ where the projection directions are tangent to Σ.

One therefore had variables describing the systems and constraints expressed by the equations of state. To measure these constraints and experimentally construct the geometry of Σ, physicists then conducted extensive studies of systems at thermodynamic equilibrium and of their variations under reversible infinitesimal transformations. Experimentally, certain variables were fixed and the way in which the others changed was measured.

A deep idea, omnipresent in mechanics since Newton, Euler, Lagrange, and Hamilton, was to base the thermodynamic description on a *variational principle* as the principle of energy minimization. Variational principles dominate physical theories and allow the solutions of certain differential equations (such as Newtonian mechanics equations) to be computed in a very operational way using the "variational calculus." They are "finalist" rather than "causal." In an admirable text published in 1744 (and edited by Constantin Carathéodory), Leonhard Euler dazzlingly expounded and exemplified his new "Method for finding curved lines satisfying maximum and minimum properties, or solutions to isoperimetric problems in the broadest sense of the term" (Euler 1744).[8]

The state surface Σ then becomes the surface of the extrema of a thermodynamic "potential" Π, which is its *generating function*. Thermodynamic potentials were introduced in the 1870s by Willard Gibbs and Pierre Duhem. Gibbs developed the mathematics behind them in his major treatise of 1902 (Gibbs 1902).

In-depth studies showed that in the (T, S, P, V)-space the surface Σ_4 of the equilibrium states can be locally expressed, depending on the experimental situations considered, by means of different generating functions of two basic variables (defined up to an additive constant which can be dropped):

- the *internal energy* $U(S, V)$,
- the *Helmholtz free energy* $F(T, V)$,
- the *enthalpy* $H(S, P)$,
- the *Gibbs free enthalpy* $G(T, P)$.

[7] So, the multi-realizability of macro properties (in this case, differential properties of macro state surfaces) is, from the outset, a property of CMI in physics. It cannot therefore be used as an argument against it, contrary to what, as we saw in the first part and at the beginning of this second part, has been claimed by many philosophers who think of macro emergence in reductionist terms.

[8] It is in this text that Euler's equations appear (which would later become the Euler-Lagrange equations), the method later known as Lagrange multipliers, and the principle of least action as a general and universal principle valid far beyond Fermat's principle in optics.

These functions are in fact functions of the 4 parameters T, V, P, S where the two non mentioned variables can be reconstructed from the other two as partial derivatives. They are the standard thermodynamic potentials.

The basic thermodynamic potential is the internal energy $U(S, V)$. When it was introduced, it was *not known* how to compute it, due to the lack of explicit microstate theories. It was not until the hypotheses of statistical physics were developed that an explicit expression became available. However, even without such an expression, it is possible to compute the *variations* of U during transformations.

This point is very interesting from an epistemological perspective. MH-level entities (potentials) are introduced that cannot be explicitly computed but whose variations are nevertheless observable and calculable and whose partial derivatives yield the *state variables*.

To work with the state function U, we must use the two *principles* of thermodynamics.

(i) The first principle states that, during any transformation of a *closed* system, the variation of its energy is equal to the amount of energy exchanged with the external environment through thermal transfer (heat) and mechanical transfer (work).

(ii) The second principle concerns the irreversibility of thermodynamical processes. Successively deepened by Carnot in 1824, Clapeyron in 1834, Clausius in 1850, Lord Kelvin and then Boltzmann in 1872, it states that the total global entropy {system + external environment} can only *increase* during a transformation of the system: $\Delta S_{glob} = \Delta S_{syst} + \Delta S_{ext} \geq 0$, and that equality only applies to *reversible* transformations.

The entropy variation ΔS_{syst} of the system is the sum of an exchange term and an entropy creation term. We therefore have $\Delta S_{syst} = \Delta S_{exchange} + \Delta S_{creation}$ where entropy creation is > 0 during an irreversible transformation and zero during an infinitely slow reversible "adiabatic" transformation. In the latter case $\Delta S_{syst} = \Delta S_{exchange}$. One can then demonstrate the fundamental result that, if the system is *closed*, $\Delta S_{exchange}$ is the *integral* of $\delta Q/T$ along the transformation. The heat Q exchanged during a transformation *is not* a state function and depends on the *path* taken by the transformation (depending on whether it is irreversible or reversible) between its initial state and its final state. In other words, δQ *is not* an exact differential that would be the differential of a state function. However, $\delta Q/T$ *is* an exact differential $dS_{exchange}$. If the system is closed and the transformation is reversible $dS_{syst} = dS_{exchange} = \delta Q/T$. Mathematically, one says that $1/T$ is an *integrating factor* for δQ because it transforms δQ into an *exact* differential dS.

Let us return to the internal energy state function $U(S, V)$ of the system. Its introduction is fundamental to thermodynamics. *Without* knowing the fine internal structure of matter and the processes taking place within it, one nevertheless introduces, as a *methodological* principle, a global, collective, macroscopic state function, a kind of "holistic" descriptor that cannot be computed directly

but whose variations can nevertheless be measured.[9]. Classical thermodynamics is therefore based from the outset on an MH principle.

Let's clarify things a little. The differential dU during a reversible infinitesimal transformation at equilibrium is $dU = \delta Q + \delta W$ by the first principle of thermodynamics, where δQ is the heat variation and δW the work furnished by the system. As mentioned above, Q and W *are not* state functions. Therefore,neither δQ nor δW are exact differentials.

Let's also assume that the system is in *mechanical* equilibrium (no change in total kinetic energy, no change in external potential energy) and that only the external pressure force P_{ext} works. By the second principle of thermodynamics, as the transformation is reversible, Q has an integrating factor, $\delta Q = TdS$. Moreover, $\delta W = -P_{ext}dV$.The system being at equilibrium, the internal pressure P equilibrates the external pressure.

(1) For the internal energy U we get therefore

$$dU = TdS - PdV \qquad ((1))$$

and $T = \partial_S U(S,V)$, $P = -\partial_V U(S,V)$. The partial derivatives with respect to S and V thus give the state variables T and P. The other partial derivatives give other thermodynamic quantities. For example $c_V = \partial_T U(S,V)_V$ is the isochoric molar specific heat (i.e., at constant V).

(2) For the Helmholtz free energy $F = U - TS$, we get

$$dF = TdS - PdV - TdS - SdT = -SdT - PdV \qquad ((2))$$

and hence $S = -\partial_T F(T,V)$, $P = -\partial V F(T,V)$. The partial derivatives with respect to T and V thus give back the state variables S and P.

(3) For the enthalpy $H = U + PV$ we get

$$dH = TdS - PdV + PdV + VdP = TdS + VdP \qquad ((3))$$

and hence $T = \partial_S H(S,P)$, $V = \partial_P H(S,P)$. The partial derivatives with respect to S and P thus give the state variables T and V. Enthalpy is very useful because in certain changes of state, the parameters U, P, and V vary greatly, and the product PV is a better variable. One will therefore study in particular the variations of H at constant P.

(4) For the Gibbs free enthalpy $G = H - TS = U + PV - TS = F + PV$ we get

$$dG = TdS + VdP - TdS - SdT = -SdT + VdP \qquad ((4))$$

and hence $S = -\partial_T G(T,P)$, $V = \partial_P G(T,\ P)$. The partial derivatives with respect to T and P thus give the state variables S and V.

There are numerous relationships between all these quantities that can be easily calculated using the rules of differential calculus and, depending on the problems being addressed and their experimental setups, allow the most relevant variables and potentials to be selected.

[9] A bit like how, in sociology, polls can be used to measure changes in opinion even if one does not know the psychological reasons behind individuals' choices, or how, in a market economy, prices guide trade.

2.1.3 Ideal gas equation of state

The above descriptions and formulas involve potentials and equations of state between state variables, but none of these are explicit. It took a long time and a lot of deep experimental innovations to find the equations correlating the thermodynamical variables. The first equation was that of (Robert) Boyle and (Émile) Mariotte (in 1662 and 1676, respectively) for *ideal gases*, which are a universal idealization that correctly approximates diluted gases (oxygen, nitrogen, etc.) at low pressure (conditions far from their liquefaction or solidification). Once the crucial concept of absolute temperature had been established, it was expressed as $PV = nRT$, where n is the number of moles (one mole is equal to N_A units, N_A being Avogadro's number as defined above) and R is a universal constant.[10] It summarizes three laws: Boyle–Mariotte's law PV = constant at constant T, Charles's law V/T = constant at constant P (1787), and Gay-Lussac's law P/T = constant at constant V (1802). In the (P, V)-plane, the isotherms are therefore arcs of equilateral hyperbolas that are asymptotic to the axes.

2.1.4 Phase transitions

One of the most remarkable empirical phenomena that classical thermodynamics has succeeded in describing and explaining is that a body can exist in several phases (solid, liquid, vapor) depending on the conditions of pressure, volume, and temperature, and that, for special values –called "critical" values – of these conditions, the system can change phase (boiling, liquefaction, melting, solidification, sublimation). For phase transitions, the volume V is in general considered to be the internal variable and (T, P) the external ones. The singular points at which phase transitions occur form a closed subset of dimension $2 - 1 = 1$ in the (T, P)-plane, called a "phase diagram". It consists of curves along which two phases coexist (transitions "of order 1") stopping at isolated points: *triple points* at which three phases coexist (for water, $T = 273.16°K = 0,01°C$ and $P = 0.006\ atm$) and *critical points* at which two phases become indiscernible (transitions "of order 2", for water, $T = 647,096°K = 373,946°C$ and $P = 217,75\ atm$).

These very special values, which are far removed from easily accessible values, demonstrate the extraordinary experimental performance achieved by precise measurements of the state surface in the vicinity of these points, and in particular the critical point. At the critical point the difference of volume between liquid and gaseous states vanishes and the fluid possesses very special properties such as "critical opalescence": the fluctuations of density become *macroscopic* and the two phases mix up in a *fractal*, milky way. Beyond the critical point there is only *one* phase, in a "supercritical" state. Thus, in the external space (T, P), there are paths going *around* the critical point and en-

[10] We will see later that the ideal gas constant is $R = N_A k = 8,314\ J.mol^{-1}.K^{-1}$, where k is Boltzmann's constant, J is the joule (unit of energy), *mol* is the mole (unit of matter), and K is the degree Kelvin (unit of temperature).

abling to pass *continuously* (without any phase transition) from one phase to the other.

The construction of phase diagrams for different substances was a whole world in the 19th century. In particular, the geometry of the state surfaces in the vicinity of the critical point was studied very carefully (see Figure 2.1).

It was shown (see Figure 2.2) that, near the critical point (T_c, P_c, V_c), plane sections of the state surface, namely the *isotherms* (T = constant), the *isobars* (P = constant) and the *isochores* (V = constant), show that the projection $(T, P, V) \longrightarrow (T, P)$, restricted to the state surface Σ of equation $\Phi(T, P, V) = 0$, has a "cusp" singularity at the critical point.

More precisely, according to the variational principle, $\Phi(T, P, V) = 0$ is the surface of extrema of the thermodynamic potential Π considered, i.e.

$$\Phi(T, P, V) = \partial_V \Pi(T, P, V).$$

Thus, near (T_c, P_c, V_c), the "fold points" of this restricted projection, consisting of those (T, P, V) such that $\partial_V \Phi(T, P, V) = 0$ for some V, describe a curve C whose tangent at (T_c, P_c, V_c) is "vertical" and therefore where the second derivative $\partial^2_{V^2} \Phi(T, P, V) = 0$.

The projection of this curve onto the (T, P)-plane is the *apparent contour* of the state surface Σ in the V-direction and forms a cuspidal curve. The critical point (T_c, P_c, V_c) projects onto the cusp point defined by the equations $\partial_V \Phi(T, P, V) = 0$ (point on the surface) and $\partial^2_{V^2} \Phi(T, P, V) = 0$ (vertical tangent). In singularity theory, this is called a "cusp" singularity. If (T, P) lies inside this cuspidal curve, there are (locally) three points of the state surface projecting to (T, P): two local minima of the map $\Pi_{T,P} : V \longrightarrow \Pi(T, P, V)$ with a maximum in between. Each minimum corresponds to a phase (liquid or vapour) and the system is in the phase corresponding to the *absolute* minimum. When the two minima are at the same height, there is a conflict of equilibria between the two phases, and they *coexist* but with different volumes. The line along which the corresponding minimal values of Π are equal separates them neatly.

This *local* geometry is characteristic of a surface that is qualitatively described by a cubic equation in V. Figure 2.1 above shows the geometry of such a surface with V as the vertical axis, P as the horizontal axis, and T as the perspective axis. The following are shown

(i) the critical isotherm with the critical point denoted γ,

(ii) the liquefaction and vaporization curves,

(iii) the spinodal curve, which is the limit of phase transitions delayed by exceptional experimental conditions (supersaturated vapor and underpressurized liquid). These curves meet at the critical point with a vertical tangent and disappear.

Figure 2.2 above shows the isotherms, isobars, and isochores on the state surface near the critical point. One can appreciate the extraordinary number of measurements (with all the necessary sophistication of the measuring devices) that such a geometry synthesizes.

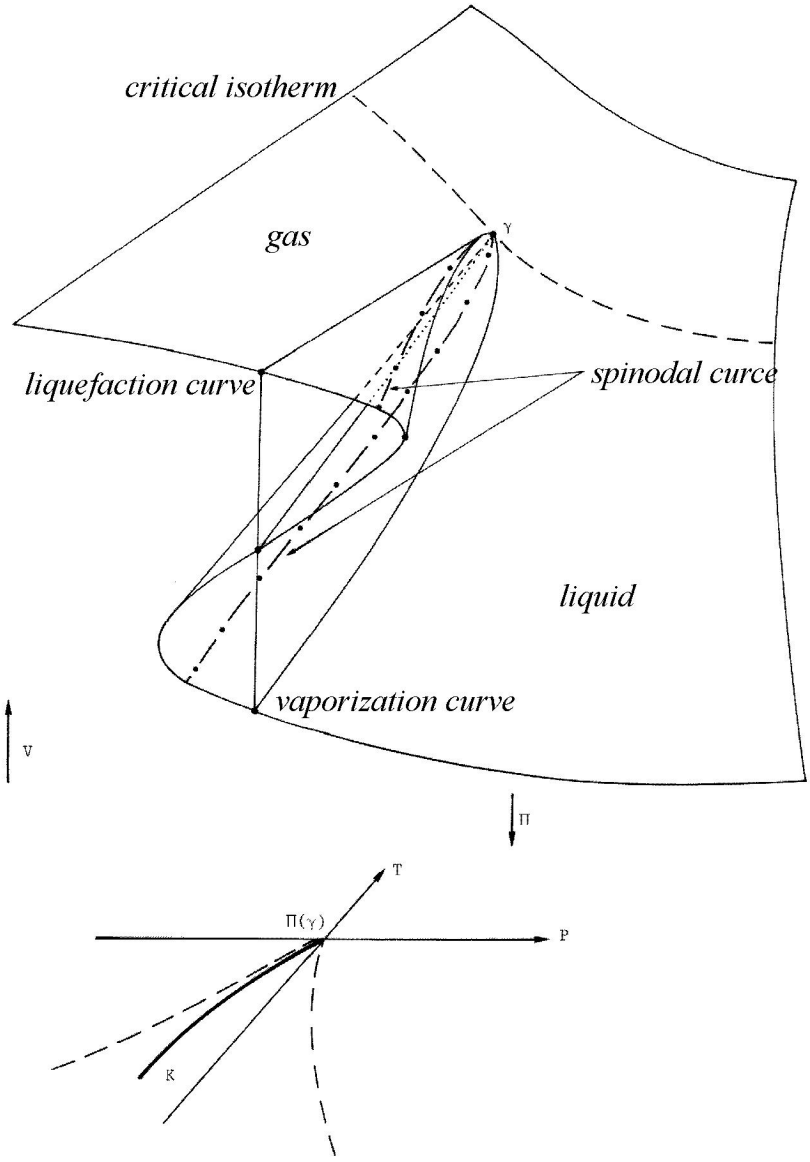


Figure 2.1: The state surface in the vicinity of the critical point (Petitot 1982).

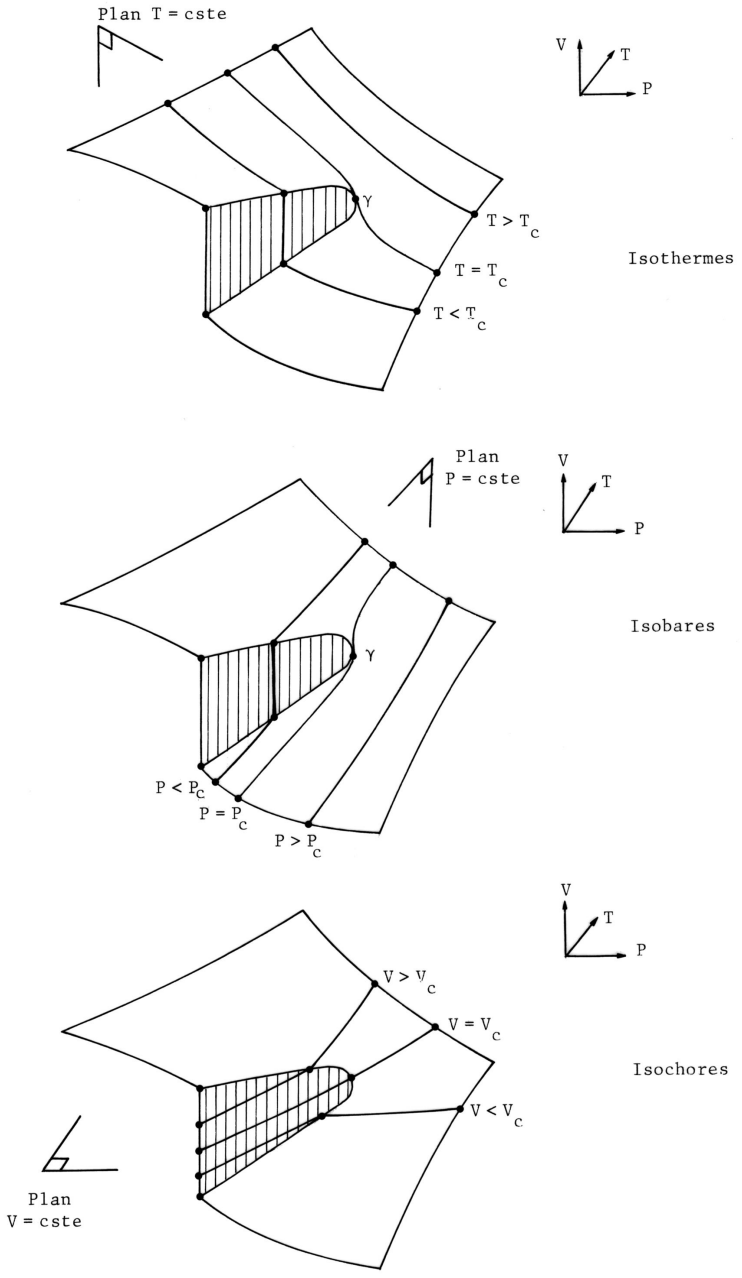


Figure 2.2: Isotherms, isobars, and isochores on the state surface near the critical point (Petitot 1982).

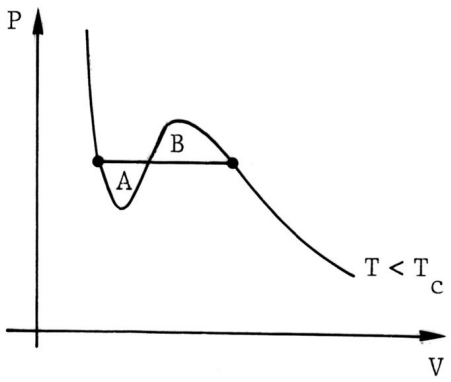


Figure 2.3: Maxwell's convention (Petitot 1982).

Figure 2.3 shows an isotherm with $T < T_c$. Mathematically, it is a cubic curve, but its physical interpretation involves what is known as the *Maxwell convention*: the physical isotherm is the "stepped" isotherm whose step is the segment P = constant connecting the volume V_L of the liquid phase to the volume V_G of the gas phase in such a way that areas A and B are equal. $V_G > V_L$ if $T < T_c$ but at the critical point where $T = Tc$ we have $V_G = V_L$.

This situation encountered by classical thermodynamics for specific "critical" phenomena such as phase transitions is in fact *universal* for systems whose state depends on an internal variable x and two external variables u and v, and minimizes a potential $f(x; u, v)$ which is a well of degree 4. The extrema (minima and maxima) of f are solutions to the equation $\partial_x f(x; u, v) = 0$, which is a cubic equation in x. Such systems are extremely numerous in nature in different fields. Their state transitions represent typical cases of what René Thom called "catastrophes," (Thom 1972, Petitot 1982).

2.1.5 The van der Waals equation (1873) (MH with a touch of meso-CMI)

A turning point in classical thermodynamics and one of its greatest successes was the explanation of phase transitions in fluids (liquid = L, gas = G) in 1873 by Dutch physicist Johannes Diderik van der Waals (1837-1923). This remarkable achievement (van der Waals 1875), which had immense consequences, earned him the Nobel Prize in 1910.

Initially, the aim was to modify the ideal gas law to produce a version that more closely reflected the behavior of real gases. To do this, van der Waals used hypotheses about the microscopic (atomic, molecular) structure of the fluid, even though this structure, assuming it existed, was unobservable. This was

therefore a turning point which, in the context of the epistemological investigation of this article, falls within the scope of CMI. However, this is a *weak* and purely *intuitive* form of CMI, situated at a *meso* level rather than a truly micro level.

The model of an ideal gas is based on the assumption that the gas consists of particles

(i) that are point-like (with volume = 0),

(ii) that do not interact with each other, and

(iii) constitute in a sense a *strictly atomistic* medium.

Here we can clearly see the specificity of "hypotheses" in physics: it is not a question of whether (i) and/or (ii) correspond to an ontological reality, but whether they allow us to correctly reconstruct the observed empirical reality through computational synthesis.

To model *real* gases, Van der Waals makes (i) more realistic by assuming that particles have a certain volume and therefore occupy part of the volume V. V is then replaced by $(V - nb)$, where n is the number of moles and b is the minimum volume of one mole of gas. To make (ii) more realistic, he assumes that there are attractive interactions between particles due to forces (since called van der Waals forces). These forces decrease the pressure by an average factor of an^2/V^2, where a is a "cohesion" term. The equation $PV = nRT$ for ideal gases is thus modified to

$$(P + an^2/V^2)(V - nb) = nRT.$$

This yields the state equation

$$\Phi(T, P, V) = P - \frac{nRT}{V - nb} + \frac{an^2}{V^2} = 0,$$

where a, b are positive.

Rewriting the state equation as a polynomial equation of degree 3 in V, one sees that the map $(T, P, V) \longrightarrow (T, P)$, restricted to the state surface $\Phi(T, P, V) = 0$, has a cusp singularity at (T_c, P_c, V_c) where $\partial_V \Phi(T, P, V) = \partial^2_{V^2} \Phi(T, P, V) = 0$. By calculating the derivatives with respect to V, we immediately find that the critical point is given by

$$(T_c, P_c, V_c) = (\frac{8a}{27bR}, \frac{a}{27b^2}, 3nb).$$

We therefore have $a = \frac{27R^2T_c^2}{64P_c}$, $b = \frac{RT_c}{8P_c}$/ Considering the "reduced" state variables, $T_r = T/T_c$, $P_r = P/P_c$, $V_r = V/V_c$, the equation becomes

$$(P_r + \frac{3}{V_r^2})(V_r - \frac{1}{3}) = \frac{8}{3}T_r,$$

which is *universal* in the neighborhood of the critical points of fluids. This famous "law of corresponding states" is of crucial importance because it shows that the critical properties of complex systems such as thermodynamic systems are largely *independent* of the fine structure of the substrates and therefore

constitutively *multi-realizable*. This universality demonstrates the autonomy of the macro level. Thermodynamics proves how much CMI is not reductionism.

2.2 Microstatistical thermodynamics (CMI)

2.2.1 The revolution of statistical physics. Gas kinetics

Let us now turn to the revolution of statistical physics, which represents the first breakthrough of CMI in physics.

Since thermodynamic variables do not refer directly to atomic or molecular concepts, any *semantic* reductionism must be eliminated from the outset. However, this does not preclude *causal* reductionism that attempts to explain them through underlying micro processes. The whole problem is whether or not a CMI-type approach allows for the computational synthesis of empirical observations of classical thermodynamics.

Here, we will focus on scientific hypotheses that are experimentally plausible, even if the atoms whose existence is assumed remain unobservable. As early as 1738, Daniel Bernoulli introduced the kinetic theory of gases in his *Hydrodynamica* (Bernoulli 1738), which already anticipated what van der Waals would later do. Then atomistic hypotheses were inferred *abductively* from experimental observations. The molecular structure of water (H_2O) has been discovered in the late 18-th century. This revolution in chemistry triggered the celebrated controversy between Joseph Priestley and Antoine Laurent de Lavoisier (1779). We should also mention John Dalton's work in the first third of the 19-th century who, following Joseph Louis Proust's studies on the masses of chemical compounds in chemical reactions, observed that they could be explained by integer proportions (with small numbers) of the constituents. We should also mention Robert Brown, who in 1827 published his observations under the microscope on the chaotic movement of particles in pollen grains and attributed it to the collisions of these particles with invisible atoms.

After these breakthroughs, it was natural to ask for a micro-level interpretation of thermodynamical variables in terms of local behavior (vibrations) and interactions (impacts, collisions) between molecules. Already Lavoisier had introduced the hypothesis that heat was due to some molecular agitation. But it has been the genius of Ludwig Boltzmann (1871), Willard Gibbs (1902) and Louis Bachelier (1901), followed by Albert Einstein (1905), Paul Langevin (1908), Jean Perrin (1909), Norbert Wiener (1923) and Paul Lévy (1933),.to achieve this revolutionary reinterpretation of thermodynamics in terms of statistical physics. The interpretation of T as the mean kinetic energy of molecules is one of the major achievements in the history of science. The micro-dynamics is incredibly complex (molecular chaos), but some simple macro "mean field" "order parameters" emerge.

The molecular hypothesis was obviously very controversial and gave rise to heated debate, particularly between Boltzmann, Ernst Mach, and Ernst Zermelo. In a tragic twist of fate, Boltzmann died by hanging himself in 1906,

feeling persecuted. Just after Albert Einstein had proven in 1905, with his molecular interpretation of Brownian motion, that the hypothesis was correct !

In short, in physics, a field where atoms were completely unobservable (at least until the discovery of X-ray diffraction by atoms in a crystal) and where only macroscopic properties were observable and measurable, and where MH therefore seemed self-evident, bold scientists attempted to make CMI computationally operational and succeeded masterfully. Let us repeat once again that, by a strange irony of the history of science, in the social sciences, where, on the contrary, micro-individuals are the most easily observable and where therefore CMI should be self-evident as a hypothesis to be tested, many efforts are being made to disqualify it for philosophical reasons.

2.2.2 Statistical computations

The whole theoretical question was to move from the micro to the macro, whereas the concepts, laws, and rules governing actions and processes at both levels are, as we have already said several times, heterogeneous and even incommensurable. A molecule in motion has velocity but no temperature or pressure, and conversely, a gas at rest has temperature and pressure but no velocity. Descriptive vocabularies are heterogeneous, and semantic reductionism therefore makes no sense.

Statistical thermodynamics has been able to overcome this aporia through a triple strategy

(i) The use of statistics, which allows to move from individual micro properties to average collective properties. For example, from a CMI perspective, to say that a gas is at rest is to say that the average velocity of the particles is zero, and we can impose this constraint on probability laws.

(ii) The fact that, if the number N of individuals at the micro level is very large, we can move to the thermodynamic limit $N \to \infty$ and the formulas are then greatly simplified. The number of atoms in a mole is Avogadro's number $N_A = 6,022 \times 10^{23}$.

(iii) The existence of quantities that can have meaning at *both* levels. The prime example is energy, which, as we have said, is a general equivalent in physics and exists in countless forms. A molecule in motion has kinetic energy and a gas at rest has internal energy. These are two different modes, but they are both energies and we can consider *converting* one into the other.

Since N is very large, we cannot describe the micro states and their evolution exactly. Paradoxically, we introduce the micro level of individual particles while putting their specific *identity* in parentheses. We seek to understand how the laws of probability governing the properties of these elementary micro units allow us to reconstruct macro observations, the collective observables being average values. This is therefore purely a *methodological* strategy, a CMI strategy.

Let's try to explain in simple terms how statistical thermodynamics computations are performed, for example for a gas. We consider a volume of N molecules of gas with mass $M = Nm$ (m is the mass of a molecule) and we

focus on a set of specifically *mechanical* micro properties that can be expressed in the vocabulary of classical mechanics:

(i) the molecules have trajectories with velocities v_i (3-component vectors) and therefore a translational kinetic energy $E_{c,i} = (1/2)m|v_i|^2$, ($|v_i|$ is the module of v_i);

(ii) if they are not point-like (if they are, for example, small spheres), they can have rotations and therefore rotational kinetic energy;

(iii) if they have an internal structure, they can vibrate or deform and have corresponding energies;

(iv) they can interact with each other through short-range forces such as the van der Waals forces mentioned above;

(v) they can collide with each other with more or less elastic collisions; etc.

So, if we dispose of general computational procedures for moving from the micro to the macro, we can then explicitly correlate thermodynamical properties with microscopic properties. The basic example remains the van der Waals equation as a modification of the ideal gas law.

The basic statistical procedure is as follows. We define microstates ω that are represented by points in a very large state space $\mathcal{E}$. For example, for each particle (atom, molecule, etc.), we will have at least three position coordinates and three velocity components, and therefore at least $6N$ dimensions for $\mathcal{E}$. Since infinite measurement precision is impossible in physics and there is always some uncertainty, $\mathcal{E}$ is decomposed (discretized) appropriately into a very large number Γ of small cells δ_c indexed by an integer $c = 1, ..., \Gamma$.[11] We then consider the "accessible" microstates ω, i.e., those that can be realized by the micro laws within the constraints characteristic of the situation under consideration (the gas has a fixed volume or a fixed temperature, etc.). We can then *count* these "accessible" microstates ω, which gives us a (huge) number Ω. Each microstate will be characterized by the distribution of the N particles in the δ_c cells of $\mathcal{E}$. If n_c is the number of particles in δ_c, we obviously have $\sum_c n_c = N$ (remember that the symbol $\sum_c$ denotes the sum from $c = 1$ to $c = \Gamma$).

The statistical procedure consists first of counting the microstates, but making the fundamental assumption that the particles are *indistinguishable*, i.e., identical in nature even though they have different mechanical states (position, velocity, etc.). This means that the particles have kinematic and dynamic states but no *internal* states.[12] Two microstates that differ only in the permutation of particles are therefore identifiable and have the same "configuration." They define a macrostate that is an *equivalence class* of microstates. This provides a precise "bridge assumption" between the micro and macro levels.

The equivalence relationship of indistinguishability of micro units accounts for the fact that the macro level, even if it is micro-based, is phenomenologically macro and has no possibility of individuating these units.[13] As we noted in the

[11] After the advent of quantum mechanics, the size of δ_c naturally led to the introduction of Planck's constant h, which controls Heisenberg's uncertainty relations.

[12] This is obviously a major difference from CMI systems whose micro units are agents (even ultra-elementary ones).

[13] Even today, atoms remain very difficult to observe due to the intrinsic limitations imposed

previous chapter, the situation is the opposite of that in the social sciences.

A macro-state therefore comprises a certain number of macro-equivalent micro-states and is therefore massively *multi-realizable*. This number can be computed by combinatorial counting (even if the computation is sometimes very difficult to perform) and provides probability laws that allow us to move to the macro thermodynamical level.

It is easy to see how combinatorial counting will work. The number of different microstates contained in a configuration varies with the configuration. For example, if there are two cells, A and B, and two particles, the configuration where both particles are in A (or B) contains only one state. However, if the configuration includes one particle in A and the other in B, it includes two microstates that differ in the permutation of the particles. We must therefore count how the N particles can be distributed among the Γ cells, where N is comparable to Avogadro's number N_A and Γ is the decomposition of the *bounded*[14] volume of $\mathcal{E}$ into small cells.

The techniques of counting in combinatorial analysis have been known since ancient times and have been systematically formalized since the 17-th century (Pascal, Fermat, and many other mathematicians): permutations, arrangements, combinations, etc. If, for example, the N particles are distributed in the Γ cells with n_c indistinguishable particles in cell δ_c, their number of permutations is the quotient of $N!$ divided by the product of the $n_c!$, where the factorial function $n! = 1 \times 2 \times ... \times (n-1) \times n$.

The main problem then is to move from the *counting* of microstates to *probabilities* so that the macrostate occupied by the system can be defined as the *most probable* one. It is a delicate point. For couting we dispose of links connecting equivalence classes of "accessible" microstates to "possible" (massively multi-realizable) macrostates. However, in order to understand the emergence of real (and not just possible) macrostates that are empirically observed, we need to find *selection* criteria, because these empirical macrostates are very specific and have very specific dynamics (a drop of ink diffuses in water and never the other way around). We therefore need a *synthetic a priori* principle for selecting and actualizing the real macrostate from among the virtual possibilities, just as the variational principle of least action in mechanics allows us to select real trajectories from among possible trajectories. This principle must be justified and argued for by criteria of plausibility and operability in terms of the fit between the models and empirical data.

---

by quantum mechanics on observation. However, scanning tunneling microscopes (STM), a nanotechnology invented in 1981 by Gerd Binnig and Heinrich Rohrer (Nobel Prize winners in 1986) and based precisely on quantum effects, make it possible to observe and move atoms.

[14]The positions are limited by the boundaries of the system, the velocities are limited by physical constraints of mass, etc.

2.2.3 Thermodynamic irreversibility VS mechanical reversibility

As we have already mentioned, the justification of the statistical approach raised difficult problems and therefore generated much debate. The most crucial issue concerned the apparent contradiction between the *reversibility* of mechanical equations at the micro level and the *irreversibility* at the macro thermodynamical levels expressed by the second principle (entropy growth). It seems impossible to deduce from reversible evolution equations, however complex they may be, a state variable that is generally strictly increasing, such as entropy.

In fact, it is a question of understanding how myriads (of the order of Avogadro's number) of small particles colliding with each other in the vacuum and satisfying Newtonian kinetic equations can, at the macro level, generate *continuous* media that satisfy fluid dynamics laws such as those of Euler or Navier-Stokes. Such an emergence occurs through the *Boltzmann equation*, which statistically describes an intermediate meso level. The equation seems simple but is extremely difficult to solve.

We will summarize the true scientific epic that led to the resolution of this problem, known as *Hilbert's sixth problem* because it is the sixth of the 23 problems that Hilbert proposed at the International Congress of Mathematicians in Paris in 1900 as major problems to be solved during the 20th century (Hilbert 1901). It focuses on solving Boltzmann's equation and was finally partially solved in 2015 after 150 years of intensive work.

But before embarking on this adventure, we will summarize the basics of Gibbs' statistical "ensembles" (Gibbs 1902).

2.2.4 The "micro-canonical" case (purely CMI)

Indeed, despite the theoretical difficulties, statistical thermodynamics has succeeded in computing, from microstates, the state variables and thermodynamic potentials

(i) in the case of isolated systems at a given energy equilibrium E,

(ii) in the case of systems that can exchange energy with the outside while being maintained at a fixed temperature T by means of a thermostat,

(iii) in the case of systems of type (ii) that can also exchange matter with the outside.

In his 1902 treatise, Gibbs refers to these three cases as (i) "microcanonical ensemble," (ii) "canonical ensemble," and (iii) "grand canonical ensemble," respectively.

Let's start with the "microcanonical" case dealing with isolated systems: E, N, V are constant but not T. The fundamental postulate is as follows: in an *isolated* system at *equilibrium*, all accessible microstates are *equiprobable*. This is an extremely strong postulate, which is validated a posteriori by its consequences insofar as it allows a good computational synthesis of macro thermodynamic phenomena. We can refer to Newton's "hypotheses non fingo" regarding the postulate of instantaneous gravitational forces at a distance: there is no

need to justify it by means of metaphysical "hypotheses" that go beyond the phenomena.[15]. We are now faced with macrostates that are equivalence classes of microstates (for the equivalence relation of indistinguishability of units) that are all equiprobable.

It was later realized that the principle of equiprobability was related to the informational principle of entropy maximization, i.e., a variational principle. The development of information theory around 1946 by Claude Shannon (1916-2001), a Bell engineer who formalized the constraints of any information transmission process (Shannon 1948), followed by John von Neumann's explanation of its analogy with physical entropy, and finally the unified theory of thermodynamic, physical, statistical, and informational entropy by Bayesian inference specialist Edwin Thompson Jaynes in 1957 (Jaynes 1957), considerably clarified the postulate of the "microcanonical" case. If we have Ω entities ω_i $(i = 1, ..., \Omega)$, here the microstates, whose occurrences in a random process are of probability p_i, the entropy is given by $S = -k \sum_i \ p_i log(p_i)$ where k is Boltzmann's constant[16]. S is always ≥ 0 because the p_i are ≤ 1 since they are probabilities and therefore the $log(p_i)$ are ≤ 0. This quantity is *maximal* when all the p_i are equal (equiprobability). In this case, all the p_i are equal to $1/\Omega$, and we therefore have $S = -k\Omega plog(p)$, i.e.

$$S = -k\Omega(1/\Omega)log(1/\Omega) = klog(\Omega).$$

The consequence of this equiprobability postulate is that the macrostate containing the *largest* number of equivalent microstates is the *most* probable. The number of states is thus transformed into probability. And since computations show that probabilities have an extremely sharp peak for large N, at the thermodynamic limit $N \to \infty$ the most probable macrostate becomes certain. If N is finite but large, then there are small statistical fluctuations around the state of equilibrium.

The fundamental postulate makes it possible to *prove* the second principle of thermodynamics, which is an extraordinary achievement. Indeed, for independent systems, the number Ω is *multiplicative*. It can be made into an additive macro-extensive variable by taking logarithms. We thus define a statistical version of entropy by setting $S = klog(\Omega)$. The function S is an increasing function of Ω, and since the most probable macrostate is the one for which Ω is maximal, it is the one with maximum entropy.

In the form $S = klog(W)$, this formula was engraved on Boltzmann's grave in Vienna.

---

[15] But we can imagine that one day it will be validated by advances in theory, as Einstein's did it for Newtonian gravity with general relativity.

[16] The value of k is $k = 1,38 \ x \ 10^{-23} \ J.K^{-1}$ where J is the joule (unit of energy) and K is the degree Kelvin. Let's remember that the constant R for ideal gases is $R = N_A k = 8,314 \ J; mol^{-1}.K^{-1}$.

2.2.5 The "canonical" case and Gibbs' formula (combination of CMI and MH)

The "canonical" case (ii), where T, N, and V are constant but energy E is not, is a triumph of CMI in physics. It is extremely interesting for the history and epistemology of CMI, since a state variable (temperature) is kept constant by the surrounding environment. The micro system is therefore not isolated. It is constrained by an *external global macro* constraint that acts causally on it in a *top-down* manner. We can therefore see that, from the outset, physical CMI went much further than a theory of bottom-up micro $\rightarrow$ macro emergence (which was already remarkable) and introduced top-down macro $\rightarrow$ micro causalities! To claim that MI is incompatible with such constraints is historically inaccurate.

We have therefore to understand how the constraint T = constant can influence the probability of microstates. We consider the total system $\mathcal{T} = \mathcal{S} + \mathcal{R}$ composed by the system under consideration $\mathcal{S}$ and the thermostat $\mathcal{R}$ as a closed system in equilibrium whose total energy is $E_{\mathcal{T}}$. The "micro-canonical" strategy can be applied to it. There are $\Omega_{\mathcal{T}}(E_{\mathcal{T}})$ micro states of $\mathcal{T}$ and they are all *equally probable*. Let ω then be a microstate of $\mathcal{S}$ alone with energy E_ω. This energy, which is now explicitly defined at the micro level (for example, as kinetic energy), corresponds to the *internal* energy U of the corresponding macrostate. The task is to find the probability $P(\omega)$ of ω.

The thermostat $\mathcal{R}$ is in a micro-state of energy $E_{\mathcal{R}} = E_{\mathcal{T}} - E_\omega$ and since, in the "micro-canonical" case, the probability is essentially the number Ω, the probability $P(\omega)$ of ω is the quotient $\Omega_{\mathcal{R}}(E_{\mathcal{T}} - E_\omega)/\Omega_{\mathcal{T}}(E_{\mathcal{T}})$. We then note that (i) the denominator is independent of ω and (ii) E_ω is very small compared to $E_{\mathcal{T}}$. We can therefore perform a first-order expansion with respect to E_ω. However, we have seen that entropy is given by $S = klog(\Omega)$, i.e. $\Omega = exp(S/k)$, and on the other hand that $\partial_E S = 1/T$. We can therefore write the first order expansion, with T constant,

$$S_{\mathcal{R}}(E_{\mathcal{T}} - E_\omega) \sim S_{\mathcal{R}}(E_{\mathcal{T}}) - (E_\omega \partial_{E_\omega} S) = S_{\mathcal{R}}(E_{\mathcal{T}}) - E_\omega/T.$$

Applying the formula $\Omega = exp(S/k)$, we find

$$\Omega_{\mathcal{R}}(E_{\mathcal{T}} - E_\omega) = exp(S_{\mathcal{R}}(E_{\mathcal{T}} - E_\omega)/k) \sim exp((S_{\mathcal{R}}(E_{\mathcal{T}}) - E_\omega/T)/k)$$

and therefore, since the exponential function transforms sums into products,

$$\Omega_{\mathcal{R}}(E_{\mathcal{T}} - E_\omega) \sim exp\left(\frac{S_{\mathcal{R}}(E_{\mathcal{T}})}{k}\right)\ exp\left(-\frac{E_\omega}{kT}\right).$$

This shows that $\Omega_{\mathcal{R}}(E_{\mathcal{T}} - E_\omega)$ is, up to the factor $exp\left(\frac{S_{\mathcal{R}}(E_{\mathcal{T}})}{k}\right)$, of the form

$$exp\left(-\frac{E_\omega}{kT}\right)$$

and that, consequently, the probability $P(\omega)$ of ω is

$$\frac{1}{Z}\ exp\left(-\frac{E_\omega}{kT}\right)$$

where Z is the factor ensuring that the sum of the $P(\omega)$ is equal to 1, as required by the fact that these are probabilities. Traditionally, one denotes the inverse $\frac{1}{kT}$ by β, which gives the exponentials

$$exp(-\beta E_\omega).$$

So, using four ingredients:

(i) a universal formula for differential calculus (first-order Taylor formula) applied to thermodynamical functions (assumed to be differentiable);

(ii) a property of the exponential function;

(iii) the fundamental formula $S = klog(\Omega)$ linking the macro variable of entropy to the number of micro states in the microcanonical ensemble;

(iv) a partial differential relation linking energy, entropy, and temperature, we get a remarkable formula concerning closed systems (N, V constant) whose environment (thermostat) maintains a constant temperature. This formula is the basis of statistical thermodynamics. It gives the probability of microstates as a function of their energy and the external constraint T. The function Z, which is the sum of the terms $exp\left(-\frac{E_\omega}{kT}\right)$ over all accessible microstates, is called the partition function of the system $\mathcal{S}$. Generally very difficult to compute as a sum of exponentials, it contains all the micro information and provides the formulas for converting micro properties into macro properties. It could be called the *basic operator* of CMI in physics.

For example, the mean energy of the system $\mathcal{S}$ is by definition

$$\langle E\rangle = \sum_\omega P(\omega)E_\omega = \frac{1}{Z}\sum_\omega E_\omega exp(-\beta E_\omega).$$

But, since $E_\omega exp(-\beta E_\omega) = -\partial_\beta exp(-\beta E_\omega)$, we have

$$\langle E\rangle = -\frac{1}{Z}\partial_\beta\left(\sum_\omega exp(-\beta E_\omega)\right) = -\frac{1}{Z}\partial_\beta Z,$$

i.e., the wonderful formula

$$\langle E\rangle = -\partial_\beta log(Z).$$

The variance of E can be calculated in the same way.

There are many other formulas of this kind that allow to recover the classical thermodynamics of the system from its partition function by computational synthesis. For example, since $P(\omega)$ depends only on E_ω, the probability that a microstate has energy E is $P(E) = (1/Z)\Omega(E)exp(-\beta E)$ where $\Omega(E)$ is the number of microstates with energy E. Therefore, $log(P(E)) = -log(Z) + log(\Omega(E)) - \beta E$. But, according to Boltzmann's formula, $log(\Omega(E)) = S(E)/k$, and since $\beta = 1/(kT)$, we have

$$log(P(E)) = -\beta(E - TS(E)) - log(Z).$$

Since T is fixed, the most probable energy E is the minimum of $E - TS(E)$, which is equal to $\langle E\rangle - TS(\langle E\rangle)$. We thus recover the *free energy* $F = E - TS$, with the mechanical energy E now representing the internal energy U of macro thermodynamics.

We also recover the principle that the classical thermodynamic equilibrium state *minimizes* the free energy F. And since the averages are extremely sharp, $\langle E\rangle$ has a probability ˜1, so $log(P(\langle E\rangle)) \sim 0$ and therefore

$$F \sim -kTlog(Z)$$

in the "canonical" situation. This remarkable and fundamental new formula shows that the statistical approach allows us to compute explicitly, with a very good approximation, the thermodynamic potential "free energy" introduced by classical thermodynamicists as a hypothesis in order to be able to use variational principles of minimization. Here again, this is a remarkable achievement by the CMI.

2.2.6 The perfect gas

These formulas make it easy to compute the properties of an ideal gas using only micro mechanical assumptions. The ideal gas approximation means that the N particles are all identical (indistinguishable), have mass m, no volume, no internal structure (no rotation, no vibration), no interactions, and move freely (independence) within volume V. Their energy is purely kinetic.

If a real gas is sufficiently diluted, the distances between the particles will be much greater than the range of the interaction forces, and the particles can therefore be considered as free particles. Kinetic energy will then dominate, as in ideal gases. We can thus see how micro physical hypotheses can be integrated into statistical computations and generate macro properties that can be tested by macro thermodynamic experimental means. This is the *methodological* core of CMI in thermodynamics.

For an ideal gas, it is easy to compute the partition function. Since for independent systems the partition function Z is multiplicative and the particles are independent, Z is $Z = zN/(N!)$ where z is the individual partition function of a particle and where the factorial $N!$ expresses indiscernibility. Let us assume that we are in the "canonical" case (a thermostat maintains the gas at a constant T). We discretize the velocities (for example, by taking Planck's constant h as the base scale), look at the kinetic energy levels, and apply the canonical probability law $exp(-\beta E_\omega)$. Isotropy implies that z is the cube of the z_1 corresponding to a single velocity component. Simple calculations then lead to the formula $z = (V/h^3)(2\pi mkT)^{3/2} = V/A^3$ and therefore to

$$Z = (1/N!)(V/h^3)N(2\pi mkT)^{3N/2}.$$

We deduce the "canonical" free energy $F = -kTlog(Z)$ using Stirling's formula, which states that for very large N (which is indeed the case here), the factorial $N!$ is equivalent to $(2\pi N)^{1/2}(N/e)^N$. We have

$$F = -NkT[(V/N) + (3/2)(2\pi mkT/h^2) + 1].$$

And since $P = \partial_V F = NkT/V$, we recover the ideal gas equation $PV = nRT$ (n being the number of moles) with the micro interpretation of the universal constant R: $R = kN_A$ (N_A being Avogadro's number).

This derivation of an empirical macro law by means of a statistical computation based on mechanical physical principles unrelated to the phenomenality and conceptuality of classical thermodynamics is truly impressive.

2.2.7 The van der Waals equation

Let us return in this context to the van der Waals equation, which introduces corrections to the ideal gas equation using macro terms that are apparently micro-based but in fact only meso-based, as they do not explicitly describe any underlying effective micro-mechanisms.

The N particles have a certain size a and are more concentrated in the volume V than in the case of an ideal gas. If v is their volume, the actual available volume is therefore of the form $V - Nv$ and their volume density $\mu = N/V$ is greater. Collisions can be computed using elementary geometry, giving an "effective cross section" for collisions. If the particles have a radius a and a diameter $d = 2a$, they will collide if the distance between their centers becomes $= 2a$, which creates an exclusion sphere with a radius d around each particle. The average number of collisions will depend on the volume density $\mu = N/V$. We then compute the mean free path and the number of collisions per unit time.

What greatly complicates the situation is that the *position* of the particles becomes relevant, which was not the case for ideal gases. The properties of the system therefore begin to depend on V in a non-trivial way. In fact, computing the partition function that takes into account the interactions between particles becomes very difficult, which is why physicists started by making an approximation known as the *mean field* approximation.

The idea is particularly interesting for the epistemology of CMI. It introduces a dialectic between *micro particles* and *macro fields*, i.e., a particle-field interaction such as that found in the case of a heavy particle in a gravitational field or an electric particle in an electromagnetic field. It is thus based on a *mixture* of CMI and MH and consists of treating particles at the micro level while treating only on average the influence of all other particles on each of them. This greatly simplifies the problem because positions are no longer relevant variables.

Mean field approximations are very effective. To go further and treat interactions in truly micro terms, remarkable new mathematical advances known as "renormalization groups" were needed. We will return to this later when discussing magnetic systems.

It should also be noted that such approximations combining CMI and MH are common in the social sciences. For example, in economics, prices are a mean field for individual actors seeking to optimize their utility. Similarly, in the studies on public opinion, the media function as mean fields for the formation of individual opinions.

The mean field translates into an interaction potential $u(r)$ where $r > d$ is the distance to the center of the particle. Since there are $(N/V)4\pi r2dr$ particles at a distance between r and $r + dr$, the interaction energy is the integral from $r = d$ to $r = \infty$ of the differential $u(r)(N/V)4\pi r^2 dr$. Physical considerations lead to use potentials $u(r)$ which are powers of (d/r), resulting in interaction energies of the form $-C(N/V)$ where C contains a power of d. Applying Boltzmann's law $exp(-\beta E_\omega/2)$ (where the factor $1/2$ takes into account the fact that an interaction is common to two particles), we find as an individual partition function $z = (1/A^3)(V - Nv)exp(-\beta E/2)$, where E is now the sum of the kinetic energy and the interaction energy. Computations analogous to those for ideal gas then show that

$$P = NkT/(V - Nv) - wN^2/V^2$$

where the second term comes from the interaction with the mean field. This gives us the classic van der Waals equation

$$P = nRT/(V - nb) - an^2/V^2 = 0$$

with $R = kN_A$, $N = nN_A$, $Nv = nb$, $wN^2 = an^2$ i.e. $w = a/N_A^2$ (N_A being Avogadro's number).

Thus, not only do we recover the van der Waals equation on a purely micro basis, and therefore the possibility of first-order phase transitions and second-order critical points, but we also see that the cubic V-shape of the equation of state comes from the mean field approximation and that the theory therefore needs to be considerably refined if we want to make it exact. Here again, the CMI is fundamental, since it allows us to *modulate* the aspects of micro behavior that will be retained in a given circumstance.

All these methods can be generalized and applied to very different situations, such as in the nonlinear thermodynamics of open systems out of equilibrium, where self-organized structures can emerge without contradicting the second law of entropy growth, since this only applies to equilibrium (see the work of Peter Glansdorff and Ilya Prigogine, Nobel Prize 1977, on "dissipative structures"). Let us also mention the analysis of strange media such as plasmas (gases of ionized particles and electrons at very high temperatures). This is a vast universe of natural phenomena that have become explainable in terms of statistical thermodynamics and therefore CMI.

That said, there is still much to understand, and CMI needs to be explored beyond mean field theories, which are a mixture of CMI and MH. At the molecular level, a liquid $\rightarrow$ gaz phase transition is an incredibly dramatic revolution: all the links giving to the liquid its cohesiveness are suddenly broken and the collective state becomes a dispersed state of "atomic" independent units. The

sophisticated formalisms used to model these "catastrophes" have been developed particularly for magnetic phase transitions, and we will discuss them later.

But first, we will revisit Boltzmann's equation and Hilbert's sixth problem, the recent solution of which represents a triumph for CMI in physics.

2.2.8 The Boltzmann equation

The Boltzmann equation is an explicit evolution equation describing how a system evolves toward a state of equilibrium. It is a highly nonlinear partial differential equation describing the *density* $f(x, v, t)$ of particles at position x and velocity v at time t in a dilute gas of spherical particles interacting only through collisions. This situation is therefore purely CMI.

It is essential to note from the outset that we are at the *meso* level of a continuous and differentiable density where, in an infinitesimal cell $dx\ dv$, there are a very large number of microparticles, and we follow the evolution of this meso-level density for a *single* particle. We will return to this point several times.

Apart from collisions, the particles describe purely kinetic uniform rectilinear motions[17], and the density function $f(x, v, t)$ is *constant* along these trajectories of constant velocity v and acceleration $= 0$.

We therefore have $f(x, v, t) = \ f(x - vt, v, 0)$, and f satisfies the partial differential equation

$$\partial_t f + v \cdot grad_x(f) = 0,$$

where $v \cdot w = \sum_i v_i w_i$ is the scalar product of the vectors v and w with respective components v_i and w_i and where $grad_x$ is the gradient of f. More precisely,

$$df = grad_x(f) \cdot dx + grad_v(f) \cdot dv + (\partial_t f)dt = 0$$

along the trajectories under consideration. If $v =$ constant and $x = vt$, then $dv = 0$ and $dx = dt$. Hence $grad_x(f) \cdot vdt + (\partial_t f)dt = 0$, and we can factor out dt. This is the first (purely kinetic) part of the equation.

However, *collisions* must be taken into account, and it is from this point that several models can be developed. In the simplest cases, we introduce binary, elastic, and micro-reversible collisions based on Boltzmann's fundamental "molecular chaos" hypothesis (*Stosszahlansatz*), which states that the velocities of two incident particles are decorrelated, that is, *statistically independent*. The hypothesis is expressed by the fact that the probability of having the incoming velocities (v_1, w_1) before a collision corresponds to the product $f(x_1, v_1, t)f(x_1, w_1,\ t)$. We will return to this factorization property frequently.

The spherical symmetry of the particles implies that the velocities remain coplanar. Let v_1 and w_1 be the incoming velocities of the two particles upon collision, and v_2 and w_2 their outgoing velocities. The assumptions imply

(i) the conservation of momentum $v_1 + w_1 = v_2 + w_2$,

[17] The external gravitational field is neglected.

(ii) the conservation of kinetic energy $|v_1|^2 + |w_1|^2 = |v_2|^2 + |w_2|^2$ (where $|v|$ is the norm of v).

In the inertial frame of the center of mass, we have the velocities V_1,W_1 and V_2,W_2 with $V_1 = (1/2)(v_1 - w_1)$ and $W_1 = -V_1$, i.e. $V_1 + W_1 = 0$. After the collision, we have $V_2 = |V1|n$ and $W_2 = -|W_1|n$, where n is a unit vector encoding the angle of deflection of the trajectories produced by the collision. We have $V_2 + W_2 = 0$ (conservation of momentum) and

$$|V_2|^2 + |W_2|^2 = (|V_1|^2 + |W_1|^2)|n|^2 = |V_1|^2 + |W_1|^2$$

since $|n|^2 = 1$ (conservation of kinetic energy). Returning to the initial inertial frame, we obtain

$$\begin{cases} v_2 = v_1 - (\sigma \cdot (v_1 - w_1))\sigma \\ w_2 = w_1 + (\sigma \cdot (v_1 - w_1))\sigma \end{cases}$$

where σ is the unit vector of the line segment connecting the two particle centers.[18] We obtain the same formulas by swapping the indices. Indeed,

$$\sigma \cdot (v_2 - w_2) = \sigma \cdot (v_1 - w_1) - 2(\sigma \cdot (v_1 - w_1))(\sigma \cdot \sigma) = -\sigma \cdot (v_1 - w_1)$$

because $\sigma \cdot \sigma = 1$ since σ is a unit vector. Therefore

$$v_1 = v_2 + (\sigma \cdot (v_1 - w_1))\sigma = v_2 - (\sigma \cdot (v_2 - w_2))\sigma$$

and idem for w_1.

We can easily verify by adding the two equations that $v_1 + w_1 = v_2 + w_2$ (conservation of momentum) and also the conservation of kinetic energy. Indeed,

$$\begin{gathered} |v_2|^2 + |w_2|^2 = |v_1|^2 + (\sigma \cdot (v_1 - w_1))^2|\sigma|^2 - 2(v_1 \cdot \sigma)(\sigma \cdot (v_1 - w_1)) + \\ |w_1|^2 + (\sigma \cdot (v_1 - w_1))^2|\sigma|^2 + 2(w_1 \cdot \sigma)(\sigma \cdot (v_1 - w_1) = \\ |v_1|^2 + |w_1|^2 + (\sigma \cdot (v_1 - w_1))^2|\sigma|^2 - 2((v_1 - w_1) \cdot \sigma)(\sigma \cdot (v_1 - w_1)) = \\ |v_1|^2 + |w_1|^2 + (\sigma \cdot (v_1 - w_1))^2|\sigma|^2 - 2(\sigma \cdot (v_1 - w_1))^2 = \\ |v_1|^2 + |w_1|^2 \end{gathered}$$

since σ is a unit vector.

Boltzmann then introduced a fundamentally nonlinear quadratic collision operator[19], traditionally denoted *Q(f,f)*. Hence his equation of evolution (EqB) toward thermodynamic equilibrium:

$$\partial_t f + v.grad_x(f) = Q(f, f).$$

This striking equation was the first equation describing the evolution of a *probability* distribution, which was revolutionary for its time.

[18] At the moment of collision, the length of the segment is exactly twice the radius, i.e., the diameter d of the spherical particles. For point-like particles, $d = 0$.

[19] Quadratic because collisions involve (generically) only two particles

The collision term $Q(f, f)$ between pairs of particles is highly problematic. Using the scattering parameter σ (encoding the angle of deviation of the trajectories due to the collision), we can write $Q(f, f)$ as an integral over w_2 and over σ of a product

$$B(v_2 - w_2, \sigma)(f(v_1)f(w_1) - \ f(v_2)f(w_2))$$

where B is the Boltzmann "collision kernel."[20] If σ is the unit vector of the segment (of length d) connecting the centers of the two particles at the moment of impact, $B(v_2 - w_2, \sigma)$ is essentially the scalar product $(v_2 - w_2) \cdot \sigma$. We restrict the integration to $(v_2 - w_2) \cdot \sigma \geq 0$, denoted $((v_2 - w_2) \cdot \sigma)_+$, to ensure that the incoming particles are indeed pre-collisional. The kernel thus measures the statistical distribution of post-collisional velocities derived from the pre-collisional velocities.

Note that the first factor $B(v_2 - w_2, \sigma)$ of the integrand concerns the outgoing particles, and that the order of the particles in the second factor

$$(f(v_1)f(w_1) - \ f(v_2)f(w_2))$$

is "incoming – outgoing".

Since we integrate over w_2 and over σ, and since v_1 and w_1 are functions of v_2 and w_2,$Q(f, f)$ is a function of v_2 (and of x and t). Hereafter, we will often denote the variable v_2 (the outgoing velocity of particle 1) by v. We are thus interested in the post-collision history of a particle possessing a pre-collision history (which is, a priori, very complicated).

The collision term

$$Q(f, f) = \int_{(w_2, \sigma)} ((v - w_2) \cdot \sigma)_+ (f(v_1)f(w_1) - \ f(v)f(w_2))dw_2 d\sigma$$

expresses the probability that v is the outgoing velocity of a particle with incoming velocity v_1 that collided at x at time t with a deflection σ with a particle having incoming velocity w_1 and outgoing velocity w_2. The first term (incoming) with $f(v_1)f(w_1)$ is a "gain" term since the collision will generate a particle with velocity v. The second term (outgoing) with $- f(v)f(w_2)$ is a "loss" term since the collision will, on the contrary, remove a particle with velocity v.

As noted by Gallagher (Gallagher et al. 2013):

> "The Boltzmann collision integral is a quadratic operator that is purely local in the position and time variables, meaning that it acts as a convolution in the v variable, but as a pointwise multiplication in the t and x variables." (p. 14)

We will now examine the implications of Newton's conservation laws for the EqB. If $\varphi(v)$ is a function of v, let $\langle\varphi\rangle$ denote the integral $\int_v \varphi(v)dv$. A lemma related to the symmetries of the situation states that we have

[20] $f(v_i)$ obviously means $f(x, v, t)$ with $v = v_i$, etc.

$$\langle \varphi Q \rangle = \int_v \varphi(v) Q(f,f) dv =$$
$$\frac{1}{4} \int_{(v,w_2,\sigma)} B(v - w_2, \sigma)(f(v_1)f(w_1) - f(v)f(w_2))$$
$$(\varphi(v) + \varphi(w_2) - \varphi(v_1) - \varphi(w_1)) dv dw_2 d\sigma$$

It can then be easily verified that if a function $\varphi(v)$ is a "collisional invariant," i.e., a linear combination of 1, the components of v, and $|v|^2$, then $\int_v Q(f,f)\varphi(v)dv = 0$. Indeed,

(i) for $\varphi(v) = 1$, $\varphi(v) + \varphi(w_2) - \varphi(v_1) - \varphi(w_1) = 1 + 1 - 1 - 1 = 0$ is the conservation of mass and therefore $\langle Q \rangle = 0$;

(ii) for $\varphi(v) = v$, $\varphi(v) + \varphi(w_2) - \varphi(v_1) - \varphi(w_1) = v + w_2 - v_1 - w_1$ is the conservation of momentum and therefore $\langle vQ \rangle = 0$;

(iii) for $\varphi(v) = |v|^2$, $\varphi(v)+\varphi(w_2)-\varphi(v_1)-\varphi(w_1) = |v|^2+|w_2|^2-|v_1|^2-|w_1|^2 = 0$ is the conservation of kinetic energy and therefore $\langle |v|^2 Q \rangle = 0$.[21]

But if the density function $f(x,v,t)$ satisfies EqB $\partial_t f + v \cdot grad x(f) = Q(f,f)$, we immediately deduce three equations for the averages from the three Newtonian conservation laws. To write them simply, let ∇_x denote the formal vector of partial derivatives $(\partial/\partial a, \partial/\partial b, \partial/\partial c)$ where (a,b,c) are the coordinates of x. If f is a function of x, $\nabla_x f = grad_x(f) = (\partial f/\partial a, \partial f/\partial b, \partial f/\partial c)$. If v is a vector field with components (v_a, v_b, v_c), the scalar product $\nabla_x \cdot v$ is the divergence $div_x(v) = \partial v_a/\partial a + \partial v_b/\partial b + \partial v_c/\partial c$. If $m = u \otimes v = (u_i u_j)$ is a matrix (where i and j range from a to c), the vector product $\nabla_x \cdot m$ is calculated using the formula $w \cdot (u \otimes v) = (u \otimes v \cdot w) = u(v \cdot w)$. Indeed, for each component u_i of u, we have effectively $u_i(\sum_j v_j w_j) = \sum_j w_j(u_i v_j)$.

We also make use of the possibility of interchanging the integrals with respect to v (which define the averages) and the partial derivatives with respect to the coordinates of x and t. Noting, for example, that the components of the vector $\langle vf \rangle$ are the $\langle v_i f \rangle$ and integrating EqB with respect to v, we obtain a first equation:

$$\partial_t \langle f \rangle + \nabla_x \cdot \langle vf \rangle = \langle Q \rangle = 0.$$

For the other equations, we note that $\partial_t(vf) = v\partial_t f$ since the variable v is independent of t and that, more generally, the partial derivatives D with respect to t and to the components of x of a product $g(v)f(x,v,t)$ give $D(g(v)f(x,v,t)) = g(v)Df(x,v,t)$ since $Dg(v) = 0$.[22] If we multiply EqB by v and integrate it with respect to v, we obtain the second equation[23]:

$$\partial_t \langle vf \rangle + \nabla_x \cdot \langle v \otimes vf \rangle = \langle vQ \rangle = 0.$$

[21] Recall that v is v_2.

[22] Recall the Leibniz formulas for the differentiation of products, which are of the form $d(fg) = fdg + gdf$.

[23] $v \otimes v$ is the matrix of the $v_i v_j$.

And we obtain the third equation in a similar manner:

$$\partial_t \left\langle \frac{1}{2}|v|^2 f \right\rangle + \nabla_x \cdot \left\langle v\ \frac{1}{2}|v|^2 f \right\rangle = \left\langle \frac{1}{2}|v|^2 Q \right\rangle = 0.$$

Hence the system of equations

$$\begin{cases} \partial_t \langle f \rangle + \nabla_x \cdot \langle vf \rangle = \langle Q \rangle = 0 \\ \partial_t \langle vf \rangle + \nabla_x \cdot \langle v \otimes vf \rangle = \langle vQ \rangle = 0 \\ \partial_t \left\langle \frac{1}{2}|v|^2 f \right\rangle + \nabla_x \cdot \left\langle v\ \frac{1}{2}|v|^2 f \right\rangle = \left\langle \frac{1}{2}|v|^2 Q \right\rangle = 0 \end{cases}$$

It should be noted – and this is essential (see Bardos 1991) – that these equations are *not* "closed." We need $\langle vf \rangle$ for $\langle f \rangle$, $\langle v \otimes vf \rangle$ for $\langle vf \rangle$, and $\left\langle v\ \frac{1}{2}|v|^2 f \right\rangle$ for $\left\langle \frac{1}{2}|v|^2 f \right\rangle$.

Boltzmann also introduced *entropy* (actually its opposite) in the form of the quantity

$$H = \int_{(x,v)} f\ log(f) dx dv$$

and showed, under the very strong assumption of "ergodic hypothesis", that H decreases irreversibly as the system evolves because its time derivative $dH/dt = \int_{(x,v)} Q(f,f) log(f) dx dv$ is ≤ 0. This is the famous *H-Theorem*. Since f is a solution to the Boltzmann equation, it implies

$$\partial_t \langle f\ log(f) \rangle + \nabla_x \cdot \langle vf\ log(f) \rangle = \langle Q\ log f \rangle \leq 0.$$

Indeed, consider $\langle log(f) Q(f,f) \rangle = \int_v log(f(v)) Q(f(v), f(v)) dv$. If we apply the lemma above to the function $log(f(v))$, we obtain a factor $log(f(v)) + log(f(w_2)) - log(f(v_1)) - log(f(w_1))$, that is, $log(f(v)f(w_2)) - log(f(v_1)f(w_1))$, which goes from index 2 to index 1. However, the other factor coming from Q is $f(v_1)f(w_1) - f(v)f(w_2)$, which, on the contrary, goes from index 1 to index 2. And since the logarithm is an increasing function, the product of the factors is necessarily ≤ 0. Integration with respect to x preserves these relations.

Furthermore, for the partial derivatives D with respect to t and to the coordinates of x, we have[24]

$$D(f\ log(f)) = (Df) log(f) + \ f(Df)/f = (Df) log(f) + Df).$$

If EqB_L is the left-hand side of EqB, we have $\mathrm{EqB}_L(f\ log f) = \mathrm{EqB}_L(f)(1 + log(f))$. Since f satisfies $\mathrm{EqB}(f)$, we have $\mathrm{EqB}_L(f\ log f) = Q(f,f)(1 + log(f))$. When we take the averages, $\langle Q \rangle = 0$ and we indeed have

$$\partial_t \langle f\ log(f) \rangle + \nabla_x \cdot \langle vf\ log(f) \rangle = \langle Q\ log f \rangle .$$

Entropy can become constant, but this requires that $log(f)$ be a *collisional* invariant and thus a linear combination of $1, v, |v|^2$. It can easily be shown that this is possible only if the density function $f(x,v,t)$ is a *Maxwell distribution*

[24] Recall that the derivative of the logarithm of a function is $log(f)' = f'/f$.

M. In this case, $Q(M, M) = 0$, which means that collisions do not alter the density and that the system is in statistical equilibrium. We will specify below the parameters of M related to a macroscopic hydrodynamic interpretation of the averages.

This paradoxical irreversible behavior of a mechanical system – admittedly extremely complex but nevertheless perfectly reversible – sparked serious debates at the time, such as those raised by the physicist (specialist of molecular structures) Josef Loschmidt. Of course, one can point out that the EqB deals with evolutions of densities, and thus loses information compared to the exact trajectories of the particles, and that consequently one cannot go backward. But this is very difficult to formulate rigorously.

Another argument was based on a theorem by Poincaré (the "return theorem") stating that a microstate (which necessarily evolves within a finite volume with bounded velocities) generally returns, after a certain time, like any non-dissipative mechanical system, as close as one wishes to its initial state (Poincaré 1908). Ernst Zermelo challenged this theorem, to which Boltzmann rightly replied that the return time is incommensurably greater than the age of the universe and therefore makes absolutely no physical sense.

Later, progress was made in understanding the paradox with the discovery of deterministic chaos in the 1970s, a major breakthrough that clarified what Boltzmann had anticipated with his "ergodic hypothesis." A system of colliding particles is a typical example of a chaotic system "sensitive to initial conditions" that is mathematically deterministic and at the same time physically unpredictable. The slightest variation, however minute, in the initial conditions causes trajectories that are initially very close to diverge completely after a certain time. So if we reverse the velocities of a microstate, it will go back in time for an extremely short time with H increasing, then begin to follow a trajectory with H decreasing again.

But in fact, the real problem was to actually prove Boltzmann's results. This was a veritable epic, whose technical mathematical challenge was conceived, within a broader theoretical framework, as the explanation for the emergence of the equations of macroscopic fluid dynamics from the chaotic Newtonian kinetics of small spherical microparticles colliding elastically.[25]

As we have emphasized, this is the sixth of the problems that Hilbert proposed in 1900 in his famous list of crucial mathematical problems to be solved during the 20th century. It is clear from the very wording of the problem that it falls exactly within the epistemology of the CMI.

In Hilbert's formulation, this sixth problem concerned the "mathematical treatment of axioms in physics" and included, regarding the kinetic theory of gases (Hilbert 1901),

> "the problem of developing mathematically the limiting processes (...) which lead from the atomistic view to the laws of motion of continua."

[25]The model of hard spheres with elastic collisions is the simplest. But there are more subtle models with "short-range" interaction potentials.

Let us provide now an overview.

2.2.9 Hilbert's Sixth Problem. Part I: Meso → Macro

There are two major problems related to the Boltzmann equation (EqB) in Hilbert's Sixth Problem.

1. On the one hand, the study of rigorous procedures for finding explicit solutions to EqB based on micro-Newtonian models of colliding particles. This is obviously an incredibly difficult problem since we have a gigantic number N (of the order of N_A) of particles of diameter d colliding chaotically and elastically.[26] Since the collisions induce velocity discontinuities, all the system's descriptors are fundamentally chaotically irregular functions. And since the transition from the discrete micro-particle scale to the continuous macro-scale is first described by thermodynamic-kinetic limits as $N \to \infty$ and $d \to 0$, we see that these will encounter formidable mathematical obstacles pertaining to functional analysis.

2. On the other hand, when the EqB is used as a mesoscale intermediate to arrive at hydrodynamic equations, methods of *averaging* ("mean field" techniques) are applied. This involves another type of limit. The articulation of these two types of limits (thermodynamic and hydrodynamic) constitutes the main challenge.

We see that we are therefore dealing with *three* levels.

(i) a micro level of N particles of diameter d colliding with one another (Newtonian mechanics),

(ii) via a kinetic limit, a meso level (EqB, statistical physics, and kinetic theory of gases),

(iii) via a hydrodynamic limit, a macro level (Euler's and Navier-Stokes' equations in fluid dynamics).

EqB thus corresponds to a meso level coordinating two physically *completely different* levels of reality governed by completely different axioms, principles, and laws. This is the problem that puzzled Hilbert: the coordination of *two* axiomatic systems describing the same physical reality viewed from the "individual" micro perspective or the "collective" macro perspective. We can see how the essence of this problem is indeed that of the CMI.

The intermediate meso level of the EqB, in particular, allows us to understand how the reversible Newtonian kinetics of a corpuscular matter agitating in empty space can generate, at the macro level of a continuous matter filling-in its entire space, typically irreversible phenomena as diffusion phenomena.

2.2.9.1 Mesoscale → macroscale: the hydrodynamic limit.

Let's start by providing some background on the hydrodynamic limit – that is, the transition from the Boltzmann equation to the equations of fluid dynamics.

[26] As we mentioned above, more complicated models can introduce inelastic collisions, or even a non-spherical geometry of the particles, as well as interaction potentials rather than collisions. But the "simple" spherical-elastic case is already so complicated that it is more than sufficient for our epistemological purposes.

The basic idea is quite simple. We will consider very singular limits (which is what makes the problem difficult) of cases where $N \to \infty, d \to 0$, and where the mean free path ℓ between two collisions $\to 0$. To arrive at rigorous results and exact solutions to the EqB, we will obviously need to control, through assumptions, how these limits coordinate, and this is not self-evident. One approach is to consider asymptotic expansions in terms of a small scale parameter, which, as Harold Grad (Grad 1958) states

> "enables us to encompass the transition to macroscopic continuum flow."

Formally, we correlate the mesoscopic "averages" seen above with *local macroscopic* variables – that is, those depending on spatial position x and time t – and use the equations satisfied by the former under Newtonian micro-conservation laws to derive hydrodynamic equations for the latter.[27]

Let $\rho(x,t)$ be the "average" $\langle f \rangle = \int_v f(x,v,t)dv$. This is a macroscopic density, the fluid volumic mass. Similarly, let u denote the "average" macro velocity $u(x,t) = \frac{1}{\rho}\int_v vf(x,v,t)dv$ defined by $\rho u = \langle vf \rangle$. This is the macroscopic velocity of the local flow of the fluid, and ρu is a macroscopic momentum.[28] Note that, while (i) in vf the variable v is an independent variable, (ii) only f depends on x and t, and therefore (iii) $\nabla_x \cdot (vf) = v \cdot \nabla_x(f)$, the same is not at all true for ρu, for which we have $\nabla_x(\rho u) = u \cdot \nabla_x(\rho) + \rho \nabla_x(u)$.

Finally, let us define the macroscopic temperature $T(x,t)$ using the "average" energy[29]

$$E = \left\langle \frac{1}{2}|v|^2 f \right\rangle = \rho\left(\frac{1}{2}|u|^2 + 3\frac{1}{2}T\right).$$

This formula can be explained as follows. The first term $\frac{1}{2}\rho|u|^2$ is the macro kinetic energy of the moving fluid, and the second term $3\frac{1}{2}\rho T$ is the interpretation in terms of macro temperature of the micro "internal kinetic energy" of the particles. We also introduce as a local macroscopic variable the thermodynamic pressure $p = \rho T = \frac{1}{3}\int_v |v-u|^2 f dv$, with the temperature T given by $T = \frac{1}{3}\int_v |v-u|^2 dv$.

In 1912, Hilbert introduced a method for solving EqB consisting in expanding f into a series in a small parameter. In the years 1916–1917, Sydney Chapman and David Enskog showed, using this method, how one could obtain hydrodynamic equations from the EqB (see Grad 1958). To give a very brief idea of this result, let us consider the simplest case, that of the 0th-order term of such expansions. This amounts to considering *locally Maxwellian* distributions, that

[27] It is important to understand, let us repeat, that the local infinitesimal macro scale remains a macro scale relative to the true micro scale. It is in fact a meso scale, and this is why the three micro-meso-macro scales are absolutely necessary.

[28] Recall that, classically, the momentum of a body of mass m and velocity V is mV.

[29] $3\frac{1}{2}$ is read as 3/2 and not 3.5. There is a $\frac{1}{2}$ per degree of freedom, and there are 3 degrees of freedom for velocity v.

is, Maxwellian in v but locally dependent (in the sense of a macro-local) on x and t.

We take the three basic equations expressing the micro-Newtonian conservation laws:

$$\begin{cases} \partial_t \langle f \rangle + \nabla_x \cdot \langle vf \rangle = \langle Q \rangle = 0 \\ \partial_t \langle vf \rangle + \nabla_x \cdot \langle v \otimes vf \rangle = \langle vQ \rangle = 0 \\ \partial_t \left\langle \frac{1}{2}|v|^2 f \right\rangle + \nabla_x \cdot \left\langle v \ \frac{1}{2}|v|^2 f \right\rangle = \left\langle \frac{1}{2}|v|^2 Q \right\rangle = 0 \end{cases}$$

and interpret them in terms of the new macro variables. We obtain:

$$\begin{cases} \partial_t \rho + \nabla_x \cdot (\rho u) = 0 \\ \partial_t (\rho u + \nabla_x \cdot (\rho u \otimes u) + \nabla_x p = 0 \\ \partial_t (\rho (\frac{1}{2}|u|^2 \ + \ 3\frac{1}{2}T)) + \nabla_x \cdot (\rho u (\frac{1}{2}|u|^2 + 3\frac{1}{2}T)) + \nabla_x \cdot (pu) = 0 \end{cases}$$

Recall that $\rho T = p$ by definition and that therefore $\rho u T = pu$.

The first equation is obvious since by definition $\langle f \rangle = \rho$ and $\langle vf \rangle = \rho u$.

The second equation is obtained by multiplying EqB by v and simply using the fact that $v = u + (v - u)$. The average $\langle v \otimes vf \rangle$ yields the tensor $\rho u \otimes u$ and other terms (in particular the tensor $\langle (v-u) \otimes (v-u) f \rangle$), which reduce (for locally Maxwellian distributions) to the term $p = \rho T$.

Note that if we take the first equation into account, this second equation can also be written as

$$\rho \partial_t u + \rho (\nabla_x \cdot u) u + \nabla_x (\rho T) = 0.$$

Indeed,

$$\nabla_x \cdot (\rho u \otimes u) = (\nabla_x \rho \cdot u) u + \rho (\nabla_x \cdot u) u + \rho (u \cdot \nabla_x) u.$$

On the other hand, $\partial_t (\rho u) = \rho \partial_t u + u \partial_t \rho$. By summing and using $\nabla_x \cdot (\rho u) = \rho (\nabla_x \cdot u) + u$, we see that we obtain the second form of the second equation plus u times the first equation. Hence the equality.

Similarly, we obtain the third equation by multiplying EqB by $\frac{1}{2}|v|^2$, again using $v = u + (v - u)$ and the fact that $|v - u|^2 = |v|^2 - 2v \cdot u + |u|^2$.

We then observe that the macroscopic hydrodynamic equations obtained in this way are exactly the *Euler equations* for a compressible fluid. We thus see how a macroscopic level, with its *own* laws, emerges through a mesoscopic level from a microscopic level governed by entirely different laws. It is difficult to imagine a more spectacular example of CMI. And at the same time, we see to what extent the modeling of this type of emergence results from the computational generativity of mathematics.

In terms of the macroscopic hydrodynamic variables, the locally Maxwellian solutions $M(x, v, t)$ of EqB corresponding to a local thermodynamic equilibrium state of constant entropy[30] are given by the formula

[30] $log(M)$ is then a collisional invariant and $Q(M, M) = 0$.

$$M(x,v,t) = \frac{\rho(x,t)}{(2\pi T(x,t))^{\frac{3}{2}}} exp\left(-\frac{1}{2}\frac{|v-u(x,t)|^2}{T(x,t)}\right).$$

It can be easily verified by hand or using a software that we indeed have $\langle M\rangle = \rho$, $\langle vM\rangle = \rho u$, $\left\langle \frac{1}{2}|v|^2 M\right\rangle = \rho(\frac{1}{2}|u|^2 + 3\frac{1}{2}T)$, $\langle Q\ log(M)\rangle = 0$.

We cannot summarize here how, by introducing perturbative expansions of the first order (and no longer just of order 0), we arrive at the Navier-Stokes equations, which are much more complicated than the Euler equations. But the basic idea is, as we have said, to introduce a small parameter ε that is (cf. Grad 1958, p.18)

> "a scale factor for the mean free path or mean collision time,"

to rescale the EqB into the form $\partial_t f + v\cdot\nabla_x f = \frac{1}{\varepsilon}Q(f,f)$ and to seek solutions of the form $f = f_0 + \varepsilon f_1 + \varepsilon^2 f_2 + ...$ (starting from a locally Maxwellian distribution f_0). We insert the series into EqB and equate the corresponding terms sharing the same power of ε on both sides of the equation. This allows us to compute the terms f_n using algorithms for solving integrals and differential equations that involve only *macro* hydrodynamic variables.

The computations quickly become very complicated as cross-collisional terms $Q(f_0, f_1)$ and higher-order terms $Q(f_1, f_1)$, etc., appear. This computational complexity is not surprising as it must be understood that all the well-known irreversible and singular phenomena of macroscopic fluid dynamics – diffusion, discontinuities, shock waves, turbulence, etc. – phenomena that make absolutely *no sense* at the Newtonian micro level, must nevertheless be recoverable from the Boltzmannian meso level. Once again, it is difficult to imagine a more exemplary illustration of CMI's methodological theses.

2.2.10 Hilbert's Sixth Problem. Part II: Micro →Meso

The second aspect of Hilbert's Sixth Problem concerns the mathematically rigorous transition from the Newtonian micro-scale to the Boltzmannian meso-scale. It requires limit processes of extreme complexity and faces formidable difficulties pertaining to functional analysis. Let us discuss it briefly, as it represents a pinnacle of theoretical physics and has opened up a computational universe. It was recently solved by Yu Deng, Zaher Hani, and Xiao Ma (Deng et al. 2025a, 2025b).

For an introduction to the problem, the reader may consult Cedric Villani's excellent review "A review of mathematical topics in collisional kinetic theory" (Villani 2002), as well as that by Isabelle Gallagher, Laure Saint-Raymond, and Benjamin Texier, "From Newton to Boltzmann: Hard Spheres and Short-Range Potentials" (Gallagher et al. 2013).

The micro-to-meso limit problem was first partially resolved, but only for *very short* times (shorter than the average time between two successive collisions of the same particle) by Oscar Lanford in 1975 (see Uffink-Valente 2010). Subsequently, important theorems were proven, for example, by Fields Medalists

Pierre-Louis Lions and Cédric Villani, as well as by, for instance, Laure Saint-Raymond in *Hydrodynamic limits of the Boltzmann Equation* (Saint-Raymond 2009), Marshall Slemrod, Alexander Gorban, and Ilya Karlin. To take things further, it was necessary to solve the problems for long times, which Deng, Hani, and Ma achieved. Their fundamental paper was the subject of a Bourbaki seminar (Gallagher 2026). Some results on the long-time limit had already been obtained, for example, by Laure Saint-Raymond, Isabelle Gallagher, Thierry Bodineau, or Sergio Simonella, but on distributions close to the Maxwellian equilibrium (see Bodineau et al. 2023).

To successfully carry out a program for solving Hilbert's problem, one must first control the ratio between several quantities in order to compute the limits at stake. A first important coefficient is $\alpha = Nd^2$. When a particle travels a distance λ between two collisions, it generates a tube whose volume is of the order of λd^2. Thus, α allows us to estimate, from the mean free path, the volume occupied by the motion of particles between collisions. Keeping α constant, for example $\alpha = 1$, is called the *Boltzmann-Grad limit*. This is equivalent to the choice of a *scale* and corresponds to a dilute gas where the mean free path of the particles is "large" and where the particles move freely in a vacuum most of the time.

When we take the limit $\alpha \to \infty$ (i.e., when we let the mean free path of the particles tend to 0), we reach an increasingly high density which, in the limit, becomes *continuous*. This is the "continuous hydrodynamic limit" we discussed earlier. It may be convenient to write $\alpha = 1/\varepsilon$, since then ε becomes a small parameter that tends toward 0. As we have seen, we use the scaled Boltzmann equation $\partial_t f + v \cdot \nabla_x f = \frac{1}{\varepsilon} Q(f, f)$. For very short times, we obtain macroscopic descriptors (ρ, u, T) that satisfy the Euler equations. But the limit process depends on the fundamental assumption of molecular chaos (*Stosszahlansatz*) introduced by Boltzmann, namely that of asymptotic independence of the individual particles.

We will now see how one can move from the Newtonian micro-scale to the Boltzmannian meso-scale in the Boltzmann-Grad limit.

2.2.10.1 Lanford's Theorem

A pivotal step in the progress of all these issues was Oscar Lanford's 1975 demonstration that the EqB is, for *short* times, a limit of Newton's equations for a gas of particles. We will attempt to summarize this difficult and technical result by drawing on the texts (Lanford 1975, 1976) and (Gallagher et al. 2013). The introductory and conceptual text (Lanford 1976) is particularly interesting to us because Lanford emphasizes the *exemplarity* of this result for the *general* understanding of transitions from the micro scales to the macro scales. His methodological reflections perfectly illustrate the CMI methodology, even though, of course, he does not refer to debates on the MI in the social sciences since he is discussing physics. To quote Lanford 1976 p. 117:

> "The conceptual foundations of the Boltzmann equation seem to me

> to merit careful study not so much for their own sake as because the Boltzmann equation is a prototype of a mathematical construct central to the theory of time-dependent phenomena in large systems – a reduced or "macroscopic" description taking into account only partial information about the underlying microscopic state but nevertheless undergoing an autonomous time development."

The descriptors[31] of the micro states (the particles' x, v, t) are not macro-accessible, but we can still derive macro-accessible descriptors from them. In the exact equations at the micro level, these two types of descriptors of completely different scales are of course coupled, but it is not self-evident that the evolution of the macro descriptors is *autonomous* and corresponds to an *emergent* level of reality governed by *sui generis* principles and laws. Hilbert's sixth problem is precisely to determine whether, under certain conditions, such an autonomous and sui generis evolution exists. We see that this is exactly the problem of the CMI. In the social sciences, the transition from the interactive individual level to the collective level is not so much a matter of whether one can produce statistics, averages, polls, etc., but rather of whether *autonomous* and *sui generis* holistic social laws can, under certain conditions, emerge from interactions between agents endowed with individual faculties. In short, can Durkheim be derived from Hayek?

The idea of this "strong" transition from the micro to the macro – that is, emergence – is remarkably formulated by Lanford 1976 on p. 118:

> "As the number of particles increases to infinity, so the microscopic description becomes more and more intricate; one can hope that the effect of the non-macroscopic coordinates on the macroscopic ones becomes, statistically, a function of the macroscopic coordinates themselves, so that in the limit the macroscopic coordinates determine their own time evolution." (emphasis added)

In other words, in the limit, the bottom-up *causal* action of the micro level on the macro level becomes statistically describable in a *macro* manner. This is the key point of any CMI-type approach.

Lanford's theorem relies on limit theorems that hold for a dilute gas (the Boltzmann-Grad limit, B-G) and for short times that are fractions of the mean free path. It has been a turning point and has introduced new methods. It was leaving two problems open: that of the continuous hydrodynamic limit mentioned above and that of the existence of long-time solutions, which we will discuss below. For now, let us summarize the basic idea behind his proof.

The EqB considers a continuous distribution $f(x, v, t)$ in the phase space $\Lambda \times \mathbb{R}^3$ of one particle, where Λ is a region of the vector space $\mathbb{R}^3$ representing positions (the box confining the gas) and the second $\mathbb{R}^3$ is the velocity vector space. This $f$actually approximately describes a microconfiguration of a very large number N of particles in kinetic states $z_i = (x_i, v_i)$ (i from 1 to N). This

[31] Lanford refers to "coordinates".

configuration can be viewed as a subset Z of $\Lambda \times \mathbb{R}^3$ or as a point $(\Lambda \times \mathbb{R}^3)^N$. It can also be viewed as a normalized sum of Dirac delta distributions δ_i assigning mass 1 to the points z_i: $d_Z = \frac{1}{N}\sum_i \delta_i$. The measure $f(x,v,t)dxdv$ locally describes, at a mesoscopic level, this exact micro-distribution d_Z on $\Lambda \times \mathbb{R}^3$. The irreversible EqB approximates the reversible Newtonian evolution of d_Z at this level.

Lanford emphasizes the meso scale as a localized macro level where the infinitesimal cells (dx, dv) remain very large relative to the particles' micro scale:

> "Intuitively, the cells are to be taken small enough to be treated as infinitesimal from the macroscopic point of view but big enough so that each of them contains a large number of particles." (Lanford 1976, p. 121)

The question then becomes how to rigorously define, in mathematical terms, what it means for d_Z to be "approximable" by f. To do this, we consider a functional space of *probability measures* (containing the f) on the one-particle phase space $\Lambda \times \mathbb{R}^3$, restricting ourselves to configurations Z that satisfy a finiteness constraint on their total kinetic energy. This functional space is then endowed with a topology stating that, given a sequence $Z[k]$ of configurations[32], the sequence $d_{Z[k]}$ converges to f

1. if the average kinetic energy of $Z[k]$ converges to $\frac{1}{2}\int |v|^2 \ fdv$ and
2. if for every small parallelepiped cell Δ in $\Lambda \times \mathbb{R}^3$ the fraction of the z_i in $Z[k]$ that lie in Δ converges to $\int_\Delta \ fdxdv$.

Given initial conditions $d_{Z(0)}$ and $f(0)$ at time $t = 0$ that are "close," one might then ask whether, with $Z(t)$ being the solution to Newton's equations EqN for Z and $f(t)$ being the solution to EqB for f, $d_{Z(t)}$ and $f(t)$ remain "close." [33] But this is certainly completely false, as it would imply a property of *stability* of the trajectories, whereas we know that the collisional systems under consideration are sensitive to initial conditions ($Z(t)$ is an example of "deterministic chaos"). Moreover, let us repeat it once again: EqB is irreversible, whereas EqN is reversible.

But on the other hand, as Lanford clearly explains, one might think that if t is *sufficiently small*, $d_{Z(t)}$ will *likely* be close to $f(t)$. But to prove this, we must introduce probability measures on the d_Z, which are *themselves* probability distributions. It is easy to imagine that this is technical and difficult. We will introduce a measure μ on $(\Lambda \times \mathbb{R}^3)^N$ concentrated on the Z whose d_Z are close to f, and then show that this measure remains concentrated on the $Z(t)$ whose $d_{Z(t)}$ are close to $f(t)$ if t is sufficiently small. To do this, we will use a *micro-canonical* type statistic based, as we saw above, on a principle of *equiprobability* of the micro-states under consideration.

[32] *Note on notations.* In what follows, when entities depend on multiple parameters, to avoid a cascade of subscripts and superscripts, we will sometimes indicate the parameters in brackets []. Here, $Z[k]$ is a sequence Z_k, but the notation $Z[k]$ avoids nested subscripts in $d_{Z[k]}$.

[33] To simplify the notation, we highlight the relevant t coordinate while leaving the other coordinates (x and v) implicit.

Hence what Lanford calls his "pre-theorem." Given a solution $f(t)$ to the EqB, for every neighborhood V of $f(t)$ there exists a neighborhood W of $f(0)$ such that the set of Z with d_Z in W and $d_{Z(t)}$ in V is a fraction close to 1 of the Z with d_Z in W. In other words, among the configurations Z with d_Z close to f, almost all have a *short time* evolution $d_{Z(t)}$ close to $f(t)$.[34]

But obviously, we must justify the microcanonical equiprobability hypothesis, which is linked to Boltzmann's molecular chaos hypothesis (*Stosszahlansatz*) regarding the asymptotic independence of the various particles. This is a problem of *a priori* choice of microscopic probabilities that Lanford 1976 formulates on p. 127 as:

> "the problem of a priori microscopic probabilities, of deciding which probability measure on the microscopic phase space to use to describe a given macroscopic situation."

Lanford then *reverses* the perspective, starting from general measures μ on $(\Lambda \times \mathbb{R}^3)^N$ and seeking the conditions on μ under which μ will assign a high probability to the $Z(0)$ close to $f(0)$ such that $Z(t)$ remain close to $f(t)$. And it is only in a second step that he asks whether the micro-canonical μ satisfy these conditions. He starts with a solution $f(t)$ to the EqB and with sequences of configurations Z_k where the number of particles N_k and their diameter d_k are variable and satisfy the B-G limit: $N_k \to \infty$, $d_k \to 0$ and $N_k d_k^2 \to 1$ as $k \to \infty$. He then considers a sequence of measures μ_k and technically defines the fact that the sequence μ_k approximates f in the sense that it concentrates, as $d_k \to 0$, on the set of microstates arbitrarily close to f.

If Δ is a parallelepiped cell in $\Lambda \times \mathbb{R}^3$, let $F_\Delta(Z)$ be the fraction of the z_i in Z that lie in Δ, that is, $F_\Delta(Z) = \frac{1}{N} \sum_{i=1,\ldots,N} C_\Delta(z_i)$ where C_Δ is the characteristic function of Δ. The sequence μ_k is an approximating sequence of f if, for all Δ, the probability, in the sense of μ_k, that the fraction $F_\Delta(Z)$ is close to $\int_\Delta f dx dv$ converges to 1 as $d_k \to 0$. The question then becomes: under what conditions, if $\mu_k(0)$ is initially an approximating sequence of $f(0)$, does $\mu_k(t)$ remain an approximating sequence of $f(t)$ for a sufficiently short period of time?

Lanford first assumes that the μ are measures of the form

$$d\mu = \mu(Z)dZ = \mu(z_1, \ldots, z_N)dz_1 \ldots dz_N$$

with functions $\mu(Z)$ continuous and symmetric with respect to the z_i, symmetric because the particles are indistinguishable.[35] Since μ is a probability measure, $\int d\mu = \int \mu(Z)dZ = 1$.

He then considers the *correlation functions*, which provide progressively more precise information about μ. These are called the *marginal* distributions of

[34] Hereafter, to simplify the language and notation, we will simplify "Z with d_Z such that..." to "Z such that...".

[35] Technically the measure μ is "absolutely continuous" with respect to the Lebesgue measure $dz_1 \ldots dz_N$ on $(\Lambda \times \mathbb{R}^3)^N$.

the distribution μ and involve selecting s particles and integrating $d\mu$ over the remaining ones. Their explicit definition is

$$\mu[s](z_1,\dots,z_s) = \frac{N!}{(N-s)!\ Ns}\int \mu(z_1,\dots,z_N)dz_{s+1}\dots dz_N.$$

The combinatorial coefficient $N!/((N-s)!\ N^s)$ is related to the binomial coefficients $N!/((N-s)!s!)$, which count the ways of choosing s objects from N. It tends to 1 as $N\to\infty$. By definition,

$$\mu[1](z_1) = \int \mu(z_1,\dots,z_N)dz_2\dots dz_N$$

and similarly

$$\mu[2](z1,z_2) = \frac{N-1}{N}\int \mu(z_1,\dots,z_N)dz_3\dots dz_N.$$

Let us then return to the $F_\Delta(Z) = \frac{1}{N}\sum_{i=1,\dots,N} C_\Delta(z_i)$ and compute $\int F_\Delta(Z)d\mu$.[36] We find the sum, normalized by N, of the

$$\int_\Delta dz_i \int \mu(z_1,\dots,z_N)dz_1\dots[dz_i]\dots dz_N$$

where $[dz_i]$ means that we do not integrate over z_i. But by rearranging the variables and using the fact that the function $\mu(Z)$ is symmetric, we see that the N terms of the sum are equal and all equal to $\int_\Delta dz_1 \int \mu(z_1,\dots,z_N)dz_2\dots dz_N$. We therefore have

$$\int F_\Delta(Z)d\mu = \int_\Delta dz_1 \int \mu(z_1,\dots,z_N)dz_2\dots dz_N = \int_\Delta \mu[1](z_1)dz_1.$$

Similarly, we find

$$\int (F_\Delta(Z))^2 d\mu = \int_{\Delta\times\Delta} \mu[2](z_1,z_2)dz_1dz_2 + \frac{1}{N}\int_\Delta \mu[1](z_1)dz_1$$

and so on.

We see, then, that the conditions for a sequence of μ_k to approximate a density $f = f_0$ are

(i) that the limit of $\mu_k[1](z_1)dz_1$ as $d_k\to 0$ is $f_0(z_1)dz_1$,

(ii) that the limit of $\mu_k[2](z1,z_2)dz_1dz_2$ is $f_0(z_1)f_0(z_2)dz_1dz_2$, and so on.

It is this property of *factorization* of the limits of the correlation functions into products of $f_0(z_i)$ that is crucial. It expresses molecular chaos.

Lanford's theorem can then be formulated as follows (recalling that equilibrium states are characterized by Maxwellian distributions). If, for the norms defining the topologies under consideration, $f = f_0$ is bounded by a Maxwellian,

[36] We will work out a few computations in detail, as they help to identify where the difficulties lie.

if μ_k is an approximating sequence of f whose correlation functions are bounded by Maxwellians[37], and if the above limits for the $\mu_k[s]$ (molecular chaos) are uniform in N, Z and s on the compacts subsets of $(\Lambda \times \mathbb{R}^3)^s$, then, in the B-G limit and for short times (where "short" is a fraction of the mean free time definable in terms of the parameters of the Maxwellian distributions), we have $\lim \mu_k[1](t) = f(t)$and $\lim \mu_k[s](t) = f(t)^{\otimes s}$ (where $\otimes$ denotes the product). In other words, molecular chaos (the asymptotic independence of particle motions) is *preserved* for short times at the B-G limit. This is the crucial point because when $d \neq 0$, it cannot be preserved.

Lanford's theorem essentially states that if the particles are statistically independent (molecular chaos) in the initial state, the collisions that cause this independence to be lost do not, however, affect – at the B-G limit and over a "short" time – the *global* dynamics of the system, which continues to behave *as if* there were "global independence."

This theorem, with its explicit quantitative estimations of the bounds, is difficult to prove, and to do so, Lanford had to introduce groundbreaking methods. Its computational structure makes it possible to overcome the paradox of an irreversible Boltzmannian limit of a reversible Newtonian dynamics. For, as Lanford 1976 emphasizes (p. 133), it shows that

> "the decrease of Boltzmann's H-function[38] is a consequence of classical mechanics (and the proper choice of initial conditions)."

This example illustrates wonderfully what the emergence of a micro → macro level of reality means within a CMI framework. We saw earlier how the laws of hydrodynamics emerge from a Boltzmannian meso level. We now see how this meso level itself emerges from a Newtonian micro level.

The solution to the paradox of irreversibility from reversibility is as follows. If, at a "short" time, we reverse the velocities, take the state $(x, -v)$ as the new initial state, and go backward in time to $t = 0$, we obtain a $f_t^*(x, -v)$ that is *not at all* $f_0(x, -v)$. This is Boltzmann irreversibility. In contrast, Newtonian dynamics are reversible. There is no contradiction because the factorization condition necessary for a sequence $_{\mu k}$ to be an approximating sequence of f is *no longer* satisfied if the velocities are reversed. There are regions in the phase spaces $(\Lambda \times \mathbb{R}^3)^s$ where the correlation functions no longer converge to products.

2.2.10.2 Clarifications.

Lanford's theorem is an aspect of the *law of large numbers* showing that as the diameter $d \to 0$, the distributions $\mu[s]$ converge to the *deterministic* limit given by EqB. To clarify the details of Lanford's computations, we will refer to the text (Gallagher et al. 2013), which systematically explains the parallel between

[37] These bounds have a regularization function that prevents particles with excessively high energy and correlations that are too strong. They are expressed (with continuous f) as conditions of the form "($f\ exp(\beta|v|^2)$ is essentially bounded". Technically speaking, they state that L^∞-norms are $< \infty$.

[38] The growth of entropy, which measures irreversibility.

the Newtonian μ and the Boltzmannian f, the goal being to show that, under certain conditions, the μ converge to the f at the B-G limit.

First, note that when the diameter d of the spherical particles is nonzero, the integration domains in the phase spaces must exclude configurations with $|z_i - z_j| < d$. When it is relevant to refer to them, we will denote these domains (which are topologically closed sets) by D. And we must also correctly handle the boundary conditions at the boundary ∂D of D. They correspond to *collisions* and are written as $\mu(t, Z^{in}) = \mu(t, Z^{out})$.

Next, for given N and d, we specify the Newtonian motion of the particles in the closed set D, which are called the *Liouville* equations (EqL). They are written as

$$\partial_t \mu + \sum_i v_i \cdot \nabla_{x_i} \mu = 0$$

with the boundary conditions. The problem is to derive the EqB from it. But in this case, f will be the limit of $\mu[1]$. Hence the need to be able to derive an equation for $\mu[1]$ from the Liouville equation. However, $\mu[1]$ depends on $\mu[2]$, which in turn depends on $\mu[3]$, and so on. In other words, the evolution equation for $\mu[1]$ is not "closed", which distinguishes it from EqB. Hence the need to consider the *hierarchy* of correlation functions (the marginals) $\mu[s](t, Z_s)$ (where $Z_s = (z_1, \ldots, z_s)$). We obviously have

$$\mu[s](t,\ Z_s) = \int \mu[s+1](t, Z_s, z_{s+1}) dz_{s+1}.$$

The problem is therefore to derive an equation of a type analogous to Liouville's equation but for the $\mu[s]$. Now, the $\mu[s](t,\ Z_s)$ are integrals over the $N-s$ unselected particles. However, we *cannot* interchange the order of integrations and differentiations in such a way as to transform the integral of the Liouville equation over the Z_{N-s} into a Liouville equation over the integral with respect to the Z_{N-s}. We must introduce *boundary terms*, which introduces *collision operators*.

The hierarchy of the $\mu[s]$ and their equations has been extensively studied in statistical physics, kinetic theory of gases, and fluid mechanics by Nikolaj Bogoliubov, Max Born, Herbert Green, John Kirkwood, and Jacques Yvon; for this reason, it is referred to, in the specialized literature, as the BBGKY hierarchy.

We will use it to measure "observables" by explicitly introducing the collision operators via a fundamental theorem of functional analysis due to George Stokes (and also Mikhail Ostrogradsky, then Lord Kelvin, and later Élie Cartan). This theorem states that integrals over D of integrands that are of "derivative" type can be reduced to integrals over ∂D.[39]

Since the distributions μ and $\mu[s]$ can be highly irregular, we treat them as "distributions" in the technical sense of Laurent Schwartz. We do not need

[39] Technically, the theorem states that if ω is a differential $(n-1)$-form on D of dimension n, the integral of ω over the $(n-1)$-dimensional boundary ∂D is equal to the integral of its exterior derivative $d\omega$ (which is a differential n-form) over D.

to delve into this theory here. It suffices to know that a distribution T with respect to a (multi)variable U of domain D acts linearly and continuously on "test functions" $\varphi(U)$ (which are regular and vanish outside a bounded region) via generalized integrals $\langle T, \varphi \rangle = \int_D T(U)\varphi(U)dU$. The theory shows that by using integration by parts, we are led to define the partial derivatives $\partial_u T$ of T by

$$\langle \partial_u T, \varphi \rangle = - \langle T, \partial_u \varphi \rangle = - \int_D T(U) \partial_u \varphi(U) dU.$$

The observables considered are "test" functions $\varphi(t, Z_s)$ defined on $\mathbb{R}^+ \times D[s]$ (the positive reals $\mathbb{R}^+$ for t, $D[s]$ for Z_s), thus depending on s distinguished particles, with bounded support, symmetric with respect to the z_i of Z_s, and satisfying the boundary condition $\varphi(t, Z_s^{in}) = \varphi(t, Z_s^{out})$. We then compute the integrals[40]

$$\int_{\mathbb{R}^+ \times D} \left(\partial_t \mu + \sum_i v_i \cdot \nabla_{x_i} \mu \right) \varphi(t, Z_s) dZ dt,$$

which are all zero by Liouville's theorem.

Let us begin with the subintegral $\int_{\mathbb{R}^+} (\partial_t \mu)\varphi(t, Z_s)dt$ and integrate it by parts using Leibniz's formula $(gh)' = gh' + g'h$ for the derivative of a product. We write

$$\int_{\mathbb{R}^+} \partial_t(\mu\varphi(t, Z_s))dt = \int_{\mathbb{R}^+} (\partial_t \mu)\varphi(t, Z_s)dt + \int_{\mathbb{R}^+} \mu(\partial_t \varphi(t, Z_s))dt.$$

But since the first term is the integral of a derivative, it is equal to $\mu\varphi(t, Z_s)$ taken between the limits $t = 0$ and $t = +\infty$, that is, to $-\mu(0)\varphi(0, Z_s)$ since φ is zero at infinity.

We therefore have

$$\int_{\mathbb{R}^+ \times D} (\partial_t \mu)\varphi(t, Z_s)dZdt = - \int_D \mu(0)\varphi(0, Z_s)dZ - \int_{\mathbb{R}^+ \times D} \mu(\partial_t \varphi(t, Z_s))dZdt.$$

Note that the first term on the right reduces to

$$- \int_{D[s]} \mu[s](0)\varphi(0, Z_s)dZ_s$$

since φ depends only on Z_s, and the second reduces to

$$- \int_{\mathbb{R}^+ \times D[s]} \mu[s](t, Z_s)\partial_t \varphi(t, Z_s)dZ_s dt.$$

Let us now calculate the second subintegral $\int_D \left(\sum_i v_i \cdot \nabla_{x_i} \mu\right) \varphi(t, Z_s)dZ$. We first integrate by parts using Leibniz's rule $\nabla_{x_i}(\mu\varphi) = \nabla_{x_i}(\mu)\varphi + \mu\nabla_{x_i}(\varphi)$.

[40] In what follows, some variables are implicit. N and d are given, μ is a $\mu(t, Z_N)$, etc.

(Note that since φ depends only on Z_s, $\nabla_{x_i}(\varphi) = 0$ for $i > s$.) We therefore have $(v_i \cdot \nabla_{x_i}\mu)\varphi = v_i \cdot \nabla_{x_i}(\mu\varphi) - \mu v_i \cdot \nabla_{x_i}(\varphi)$. Next, note that since ∇_{x_i} depends only on the x_i and the v_i are independent of the x_i, we have $\mu v_i \cdot \nabla_{x_i}(\varphi) = \mu\nabla_{x_i}(v_i\varphi)$ and $v_i \cdot \nabla_{x_i}(\mu\varphi) = \nabla_{x_i}(v_i\mu\varphi)$. The latter term is the *divergence* of the vector field $\mu\varphi v_i$ and is essential. To compute $\int_D (\sum_i \nabla_{x_i}(v_i\mu))\,\varphi(t, Z_s)dZ$, which can be written in shorthand as $\int_D \nabla_X(V\mu)\varphi(t, Z_s)dZ$, we can use the so-called "Green's formula" of Stokes' theorem, which states that the integral of the divergence $\int_D \nabla_X(V\mu\varphi)$ is equal to the integral over the boundary $\partial D = B$ (equipped with its surface differential $d\beta$) of the scalar product $n \cdot (V\mu\varphi)$ of the field $V\mu\varphi$ with the (outgoing) normal vector n at ∂D. It is therefore necessary to carefully analyze these boundary integrals.

Since we are computing over sums $\sum_i$, we must distinguish each x_i and consider the portion $B(i, j)$ of the boundary B where $|x_i - x_j| = d$ and where spheres i and j collide. Since there are N particles but φ depends only on s particles, we must treat the four cases separately: $1 \le i \ne j \le s$, $s \le i \ne j \le N$, $1 \le i \le s$ and $s + 1 \le j \le N$, $1 \le j \le s$ and $s + 1 \le i \le N$.

Therefore, using integration by parts and Stokes' theorem, we end up having to compute

$$- \sum_{i=1,\dots,s} \left(\int_{\mathbb{R}^+ \times D} \mu(t, Z_N)\nabla_{x_i}(v_i\varphi(t, Z_s))dZ_N dt \right) + \sum_{1 \le i \ne j \le N} \left(\int_{\mathbb{R}^+ \times \mathbb{R}(V) \times B(i,j)} n \cdot (V\mu\varphi)dV d\beta dt \right).$$

In the first term, the sum runs only from 1 to s because $\nabla_{x_i}(v_i\varphi(t, Z_s)) = 0$ if $i > s$. And for the integrals with $i \le s$, we use the formula $\int h(Z_N)g(Z_s)dZ_N = \int h[s](Z_s)g(Z_s)dZ_s$, which defines the correlation functions, and we find

$$- \sum_{i=1,\dots,s} \left(\int_{\mathbb{R}^+ \times D[s]} \mu[s](t, Z_s)\nabla_{x_i}(v_i\varphi(t, Z_s))dZ_s dt \right).$$

In the second term, $\mathbb{R}(V)$ is the $\mathbb{R}^{3N}$ of velocities. It is this second term that breaks down into four sub-terms, which must be computed carefully. This is somewhat tedious, and we refer to (Gallagher et al. 2013). We take into account the fact that we have boundary conditions for both μ and φ, that these are symmetric, and that φ depends only on the first s variables. We show that this implies the vanishing of the first two subterms corresponding to $1 \le i \ne j \le s$ and $s \le i \ne j \le N$. We also show – this is more complicated – that for the other two terms, using $\mu[s](t, Z_s) = \int \mu[s + 1](t, Z_s, z_{s+1})dz_{s+1}$ and the symmetry of the functions, we obtain the key formula[41]

$$-(N - s)d^2 \sum_{i=1,\dots,s} \left(\begin{array}{c} \int_{\mathbb{R}^+ \times \mathbb{S} \times \mathbb{R}^3 \times D[s]} (\omega \cdot (v_{s+1} - v_i))\mu[s + 1] \\ (t, Z_s, (x_i + d\omega, v_{s+1}))\varphi(t, Z_s))\; dZ_s d\omega dv_{s+1} dt \end{array} \right).$$

[41] The long line in parentheses has been split into two lines.

In this somewhat complicated formula, $(N-s)d^2$ is a scaling factor equivalent to 1 for s not too large (B-G limit), $\mathbb{S}$ is the unit 2-sphere of the ω, ω is the unit vector along $(x_{s+1} - x_i)$ and $(x_i + d\omega, v_{s+1})$ is the additional particle z_{s+1} required by $\mu[s+1]$. The boundary conditions are expressed by the fact that, if $(x_i,\ v_i)$ and $(x_i + d\omega, v_{s+1})$ are post-collisional, then

$$\mu[s+1](t, Z_s, (x_i + d\omega, v_{s+1})) = \mu[s+1](t, Z'_s, (x_i + d\omega, v'_{s+1}))$$

where Z'_s is Z_s with $(x_i,\ v_i)$ replaced by (x_i, v'_i) and where v'_i and v'_{s+1} are the pre-collisional incoming velocities. As $d \to 0$, $x_i + d\omega$ becomes x_i and we obtain the Boltzmann hierarchy of $f[s]$.

It is the last boundary term that allows us to define a collision operator

$$(C_{s,s+1}\mu[s+1])(t, Z_s) =$$
$$-(N-s)d^2 \sum_{i=1,\dots,s} \left(\int_{\mathbb{S}\times\mathbb{R}^3} (\omega \cdot (v_{s+1} - v_i))\mu[s+1](t, Z_s, (x_i + d\omega, v_{s+1}))) d\omega dv_{s+1} \right).$$

By combining all these calculations, we see that, up to the initial condition, we obtain an evolution equation for the marginal distribution $\mu[s]$ acting on the test functions:

$$\partial_t \mu[s](t, Z_s) + \sum_{i=1,\dots,s} v_i \cdot \nabla_{x_i} \mu[s](t, Z_s) = (C_{s,s+1}\mu[s+1])(t, Z_s).$$

This is the (BBGKY) hierarchy of evolution equations for correlation functions. For $s = 1$, the equation is

$$\partial_t \mu[1](t, z) + v \cdot \nabla_x \mu[1](t, z) = (C_{s,s+1}\mu[2])(t, z)$$

which, as we can see, closely resembles EqB.

We can clarify this similarity. In the definition of the collision operator $C_{s,s+1}$, the unit vector ω runs through the entire 2-sphere $\mathbb{S}$. However, depending on the sign of the scalar product $\omega \cdot (v_{s+1} - v_i)$, we can separate the hemisphere corresponding to pre-collisional situations from the one corresponding to post-collisional situations. Thus, $C_{s,s+1} = C^+_{s,s+1} - C^-_{s,s+1}$. In $C^-_{s,s+1}$, we can replace ω with $-\omega$ and keep $\omega \cdot (v_{s+1} - v_i) \geq 0$ provided that we replace $x_i + d\omega$ with $x_i - d\omega$. We see that in the case $s = 1$, we then arrive at a collision operator absolutely analogous to Boltzmann's $Q(f, f)$, provided that we have the factorization (tensorization) $f[2] = f \otimes f$, which expresses the molecular chaos.

We thus obtain a formal parallel between the Newtonian hierarchy of the $\mu[s]$ derived from EqL and a Boltzmannian hierarchy of $f[s]$ associated with EqB. And the whole problem is to rigorously show that, in the B-G limit, the former converges to the latter.

(Gallagher et al. 2013) summarize the situation as follows. Let $T[s](t)$ be the Newton-Liouville flow of the s distinguished particles. We have

$$\mu[s](t, Z_s) = T[s](t)\mu[s](0, Z_s) + \int_{0 \leq r \leq t} T[s](t-r)(C_{s,s+1}\mu[s+1])(r, Z_s)dr$$

and similarly, for the limit $d \to 0$, with $S[s](t)$ the Boltzmann flow,

$$f[s](t, Z_s) = S[s](t)f[s](0, Z_s) + \int_{0 \leq r \leq t} S[s](t-r)(C_{s,s+1}f[s+1])(r, Z_s)dr.$$

Once these parallel hierarchies of evolution equations have been established, let us return to the main problem, which is to rigorously prove the desired convergences using uniform estimates in "good" functional spaces.

As we have already noted, the norms of the functions $g[s](Z_s)$ used are the sup over the Z_s of $g[s](Z_s)exp(\beta E)$ (where E is the average kinetic energy of the s particles and β is a thermodynamic parameter of the form $1/kT$), and we restrict ourselves to the normed space of those for which this norm is $< \infty$. If the diameter d is > 0, Z_s is obviously constrained to be in $D[s]$. For a sequence of $g[s]$, the norm considered is then the sup over $s \geq 1$ of the norms of the $g[s]$ multiplied by another Maxwellian factor $exp(ms)$, where m is a real parameter.

As we have noted, being bounded with respect to these norms has a "regularization" function that avoids particles with excessively high energy and overly strong correlations. We then show that if we start from an initial condition $\mu[s](0)$ with parameters $\beta(0)$ and $m(0)$ for the norm, then we have a unique solution $\mu[s](t)$, $\beta(t)$, $m(t)$ over a "short" time T with *uniform* estimates ensuring that the norm of the sequences $\mu[s](t)$, $\beta(t)$, $m(t)$ remains less than twice that of $\mu[s](0)$, $\beta(0)$, $m(0)$ as long as $t \leq T$.

We then define "admissible" Boltzmannian initial data, that is, data attainable via the B-G limit from Newton-Liouville initial data with Maxwellian bounds. Among these are the *chaotic* $f[s](0)$, i.e., the tensorized $f[s](0) = f(0)^{\otimes s}$ that we have already seen. As the authors (Gallagher et al. 2013) state:

> "Such initial data are particularly meaningful insofar as they will produce the Boltzmann equation." (p. 44)

They are essential since it can be shown that the admissible Boltzmann initial data form the convex hull of the chaotic data.

Lanford's theorem can then be refined to state that if the $f[s](0)$ with Maxwellian bounds form a Boltzmannian initial data which is a limit of a sequence $\mu[s](0)$, then there exists a "short" time T such that $\mu[s](t)$ converges in the B-G limit to $f[s](t)$ *uniformly* for $t \leq T$. In particular, if the sequence $f[s](0)$ is chaotic ($f[s](0) = f(0)^{\otimes s}$), then the $f[s](t)$ remain tensorized and $\mu[I](t)$ converges to the solution f of the Boltzmann equation with initial data $f(0)$. Thus, the theorem

> "describes an asymptotic form of propagation of chaos, in the sense that an initial property of independence is propagated in time, in the limit." (Gallagher et al. 2013, p. 49).

Analysis of the *computational* structure of these results allows us to understand the emergence of irreversibility as a meso-macro property depending on *sui generis* laws (such as entropy growth), which do not exist at the micro level.

> "The irreversibility inherent in the Boltzmann hierarchy appears in the limiting process as an arbitrary choice of the time direction (encoded in the distinction between pre-collisional and post-collisional particles), and more precisely as an arbitrary choice of the initial time, which is the only time for which one has complete information on the family of marginals." (p. 50)

Since this presentation of Lanford's theorem is already quite technical, we cannot explain the convergence computations in greater detail and refer interested readers to (Lanford 1975) and (Gallagher et al. 2013). The goal is to express $\mu[1](t)$ in terms of $\mu[1](0)$ using expansions based on "collision histories," which are graphs tracing backward the history of finite groups of particles.

Let us simply quote the authors (pp. 109 and 119)

> "The main difficulty lies in the fact that, in contrast to the Boltzmann situation, collisions in the BBGKY configuration are not pointwise in space (...). At each collision time tk, a small error is therefore introduced, which needs to be controlled."
> "The main difficulty is to control the possible spatial concentrations of particles, which would contradict the rarefaction assumption [B-G] and lead to an uncontrolled collision process."

2.2.10.3 The Deng, Hani, Ma theorem

But the most difficult part of Hilbert's Sixth problem remained to be solved after Oscar Lanford's breakthrough. It was necessary to prove convergence results for the *long time*, at least over the time during which the regular solution of EqB exists. This is what Yu Deng, Zaher Hani, and Xiao Ma successfully achieved (Deng et al. 2025a, 2025b).

The main difficulty was that, for long times, even if the molecular chaos hypothesis holds initially and even if it propagates over a short time according to Lanford, it is very quickly invalidated by collision cascades. And one cannot partition an interval $[0,\ T]$ into small intervals via a time layering. This does not work because the initial conditions at t_k $(k \geq 1)$ for the interval $[t_k, t_{k+1}]$ no longer satisfy the conditions of Lanford's theorem. In short, one must take into account the *history* of molecular collisions, and this is extremely difficult.

Deng, Hani, and Ma succeeded in doing so using powerful combinatorial tools. They introduced graphs, diagrams, and patterns representing the collisions that could affect "global independence" in Lanford's theorem. Then,

by cutting them up and rearranging them, and by explaining how "pathological" collisions could be eliminated, they were able to show that this "global independence" could exist as long as the regular solution to the EqB also exists.

This finally allowed them to solve Hilbert's Sixth problem. The event was unanimously praised as a tour de force.

2.2.10.4 A scientific epic.

Regarding the epistemological scope of this astonishing scientific story spanning a century and a half (1872–present), we can revisit the previously mentioned relationship between scientists and a hypothesis. When a hypothesis is plausible and legitimized by the fact that it is testable in highly simplified cases, it becomes a scientific challenge that can last for a very long time and be carried forward by several generations. Such a research program completely escapes the notion one might form of it at a purely conceptual level. The transcendence of computational synthesis over conceptual analysis is particularly evident and spectacular here.

This is the case for all CMI-type approaches to complex multiscale systems. They constitute a theoretical and computational challenge, and it is therefore irrelevant to philosophize about them purely conceptually, especially if it is to argue that micro → meso → macro emergence processes are impossible, since we have just seen, on the contrary, a remarkable physical example of them.

2.3 Magnetic media

A particularly important branch of statistical physics, which has had a huge number of applications in countless fields, including the social sciences, is that of *magnetic* fields. Let us give a few brief details about its models.

2.3.1 Phase transitions and mean fields (CMI and MH)

The interest of magnetic systems for our purposes is that they are composed of units (atoms, etc.) located at the nodes of a lattice, units that have their *own internal* magnetic moment (spin) and can interact depending on their distances while being subjected to an *external* magnetic field. The physical explanation of internal magnetic moments is not trivial. It requires a change of scale and consideration of the electronic layers of atoms. Electrons are 1/2 spin particles (fermions) and most are paired in opposite spin pairs. Those that are not paired can induce a magnetic moment in the atom.

Since the position of the atoms is fixed by the lattice, their *external* kinetics (the kinematics and dynamics of velocities and collisions) are no longer relevant, and the state of the medium arises from the configuration of the *internal* spin states and their magnetic interactions. The situation is completely opposite to that of ideal gases, where the units have no internal states and only kinetic states. Obviously, having a few internal magnetic states is rather poor

as subjective interiority[42], but that is enough to make units into "agents." And besides, , quite often, the dialectic of the political opinions of human subjects is not so much richer than that between spin up and spin down...

We can thus see that, even at the most fundamental physical level, CMI can involve an internal structure of micro units that plays a decisive role in the organization of emerging macro structures. Models of this type are applicable to all situations where units have an internal state (e.g., neural networks or political opinions shared by citizens) and behavior that depends on interactions. It is important for our purpose to emphasize that, at the physical level, they can legitimately be considered *multi-agent* models.

The scope of their possible applications is clear. While in "social physics" it is not easy to see what the equivalent of pressure or volume might be, it is much easier to understand why and how opinion models can be inspired by networks of units with an internal state. The values of the internal state correspond to opinions, the couplings to agreements/disagreements, the external field to the collective influence of "opinion makers," etc. In "cold" societies, order and consensus dominate spontaneously, while in "hot" societies, disorder, dissent, and even chaos dominate spontaneously, and the influence of external mean fields (e.g., an ideology or a threat) is necessary to induce a minimum of overall consensus. We will return to this point below.

The basic discrete models of magnetic systems are models of configurations of N spins s_i with n components that can take a discrete set of q values, spins distributed at the nodes i of a network of N sites. If we then assume that each spin interacts only with the spins of a certain number of its neighbors (short-range interaction hypothesis) and if J_{ij} is the coupling constant between spin s_i and its neighboring spin s_j,, we can immediately write the energy E – known as the Hamiltonian – of the system. For the most classic case of 1-component spins that can take two values ± 1 (up/down), we obtain what is called the *Ising model* (Ernst Ising 1900-1998, see Ising 1925)

$$E(s) = \sum_{i,j} J_{ij} s_i s_j - H \sum_i s_i$$

where H is the external magnetic field.[43] Despite their apparent simplicity, such formulas lead to partition functions $Z = \sum_s exp(-\beta E(s))$ (the "canonical" case) of extreme complexity. Even the Ising model becomes very challenging as soon as the J_{ij} are not constant and may have different signs. The free energy per site is $f = -\frac{1}{N} kTlog(Z)$ and the magnetization per site is $(\partial_H f)_T$. These magnetic media are susceptible to *phase transitions*. For example, a ferromagnetic material (i.e., permanently magnetized like iron) has its magnetic moments aligned, a microscopic property that causes magnetization to emerge as a *spontaneous macroscopic* magnetic order. However, the material loses its

[42]But not so poor, because in addition to two-state spins (Ising model), there can be multiple spin states (Potts model) and, in more complex cases (Heisenberg model), spin vectors that are three-dimensional unit vectors.

[43]In all models involving equations of this type, the J_{ij} are restricted to truly coupled units. However, they can be extended to all pairs (i, j) by setting $J_{ij} = 0$ if there is no coupling.

magnetization and becomes paramagnetic above a certain temperature – known as the "Curie point" – when thermal disorder takes precedence over magnetic order. This phase transition is a second-order transition similar to the one we encountered earlier with critical points in substances such as water.

To explain these transitions, the great physicist Lev Landau (1908-1968, Nobel Prize winner in 1962) developed a *mean field* theory. As we saw above, these mixed CMI-MH theories establish a transition between the micro and macro levels using meso-level concepts. Landau's central idea was that transitions from an ordered phase to a disordered phase could be described using an *order parameter* η, which is an extensive variable that is zero in the disordered phase and non-zero in the ordered phase (see the classic treatise *Statistical Physics* by Landau and Lifchitz, 1958). The external magnetic field H is the intensive variable conjugate to η.

The free energy F, whose minima represent the phases, depends essentially on the two controls T (temperature) and H, and we have (if the pression P and the volume V are constant) $S = -(\partial_T F)_H$ (entropy) and $\eta = (\partial_H F)_T$. For permanent magnets (ferromagnetism), in the absence of an external magnetic field H, only temperature remains as a relevant state variable with its critical value T_c at the Curie point, the value of the order parameter being $\eta = 0$ for $T > T_c$ and $\eta \neq 0$ for $T < T_c$. The thermodynamic quantities behave as they do at the critical point for gases and liquids, but with different variables: (H, η) instead of (P, V). Figure 2.4 shows the cubic geometry of the state surface.

Landau's mean field theory is based on the extremely strong assumption that at the critical point, the free energy, which is qualitatively a potential well with η as its minimum, remains *differentiable* in η. For $T > T_c$, $\eta = 0$ is the minimum, but at T_c, this minimum bifurcates towards a maximum flanked by two minima $\neq 0$ with opposite signs. It is the new minimum > 0 that gives the new $\eta \neq 0$. It can then be shown that, in the vicinity of the critical point, η varies as the square root $(T_c - T)^{1/2}$. This power law in ΔT^β with *critical exponent* β is a basic example of the analysis of the system's quantities in the vicinity of the critical point in terms of $\Delta T = T_c - T$. Other critical behaviors include magnetic susceptibility, which diverges in $1/\Delta T$ (Curie-Weiss law), or specific heat, which is discontinuous (with a jump proportional to T_c), or the relation to the external field, which varies in $H^{1/3}$.

2.3.2 Onsager's computations and scaling laws

Following these early results, the theory of critical magnetic phenomena underwent extraordinary developments, on the one hand by gathering a huge body of experimental data on very different magnetic bodies, and on the other hand by solving certain models *exactly*. This revealed dtrong discrepancies with the predictions of Landau's model and motivated a search for the cause. The conclusion was reached that the "mean field" hypothesis (that the thermodynamic potential is still differentiable near the critical point) is fundamentally wrong because it neglects *fluctuations* which become "giant" (macroscopic) at the critical point.

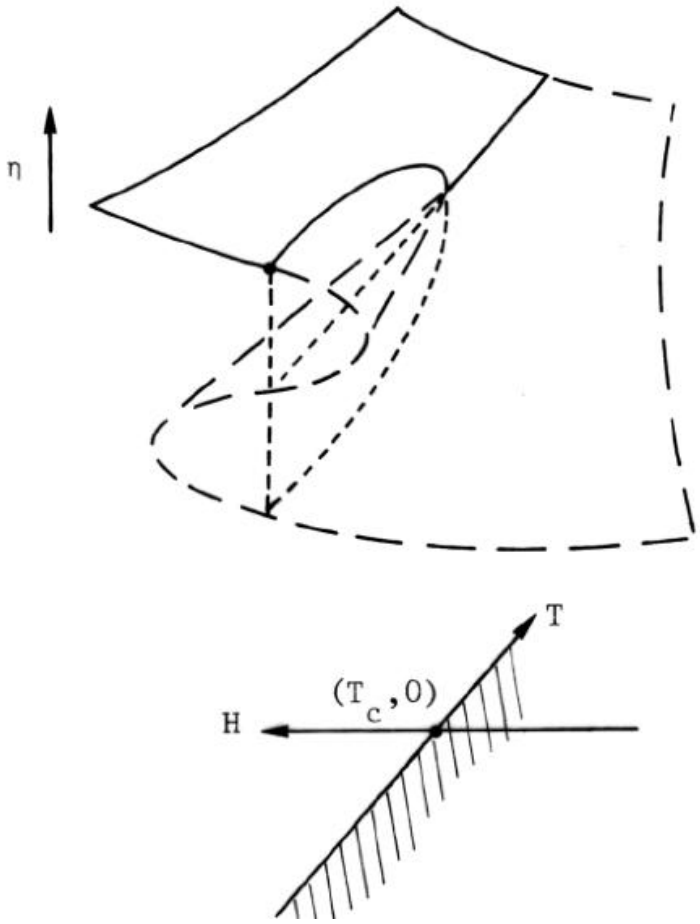


Figure 2.4: The cubic geometry of the phase transition surface for magnetic phase transitions. The external field H is ≥ 0 (Petitot 1982).

Technically speaking, Landau's approximation *cannot* account for the microscopic causes of phase transitions for the following reason. It is a mixture of CMI and MH. In a ferromagnetic body, spins always correlate with each other and tend to align locally. As a result, there are always fluctuating areas where the spins are parallel. Let ξ be their mean dimension. This is a characteristic length of the system called the *correlation length*. At the critical point, ξ *diverges*, the fluctuations become macroscopic and act as *long-range* correlations.

Particular importance was then given to the measure of the critical exponents: η as ΔT^{β}, specific heat as $\Delta T^{-\alpha}$, isothermal susceptibility as $\Delta T^{-\gamma}$, ξ as $\Delta T^{-\nu}$, H in η^{δ} , correlation function $\Gamma(r)$ in $1/r^{d-2+\eta}$ (this is not the same η as the order parameter) where d is the dimension of the system and r is the distance between two points. β and δ concern the order parameter η, α and γ concern the thermodynamics of the system, ν and η concern fluctuations. Landau's theory gives $\beta = 1/2$, $\delta = 3$, $\alpha = 0$, $\gamma = 1$, $\nu = 1/2$, $\eta = 0$ and therefore, for dimension $d = 3$, $\Gamma(r)$ is in $1/r$.

The notion of critical exponent is very interesting not only physically but also philosophically. One studies a complex system described by a certain number of descriptors and having two phases, i.e., two completely different types of macro states, in a sense two "modes" of behavior, corresponding to different external constraints and in particular to different values of a controlling parameter T. The phase transition (the "catastrophe") occurs for a critical value T_c. One then assumes that fundamental information about the *internal* microstructure of the system is provided by the *critical* behavior of the descriptors, by their *singularities*, which are expressed by power laws as a function of $\Delta T = T_c - T$. This principle could be called the "Thom principle" because René Thom gave it an 'ontological' status with his "catastrophe theory"."

However, the critical values of Landau's model have not been verified empirically, a result that opened up a new field of research. A theoretical breakthrough in this context was Lars Onsager's[44] solution of the Ising model in dimension $d = 2$ on a square lattice with interactions between nearest neighbors (Onsager 1944). The computations were highly complex but made it possible to compute the critical exponents *exactly*. Onsager found $\beta = 1/8$, $\delta = 15$, $\alpha = 0$, $\gamma = 7/4$, $\nu = 1$, $\eta = 1/4$, values very different from those of Landau.

After Onsager, more and more sophisticated observations were carried out, either experimentally or numerically on explicit models. They led to the conclusion that the critical exponents depend only on the dimension d and the dimensionality n of η (here $n = 1$) and that they are related by *universal scaling laws*. These properties of universality are very deep and express a universality of critical behaviors of the underlying processes. Their discovery constitutes a true revolution, as important as Kepler's laws in classical mechanics or Rydberg and Balmer's laws on the spectral lines of the hydrogen atom, which were at the origin of quantum mechanics. The four main scaling laws are

(i) Rushbrooke's law $\alpha + 2\beta + \gamma = 2$,

(ii) Widom's law $\gamma = \beta(\delta - 1)$,

[44] 1903-1976, Nobel Prize in Chemistry in 1968.

(iii) Fisher's law $\gamma = (2-\eta)\nu$,
(iv) Josephson's law $d\nu = 2 - \alpha$.

Let's check that they are satisfied by Landau's and Onsager's values, which are quite different from each other.

For Landau ($d = 4$):
(i) Rushbrooke's law $0 + 2(1/2) + 1 = 2$, i.e. $1 + 1 = 2$,
(ii) Widom's law $1 = (1/2)(3-1)$, i.e. $1 = 2/2$,
(iii) Fisher's law $1 = (2-0)(1/2)$, i.e. $1 = 2/2$,
(iv))Josephson's law $4(1/2) = 2 - 0$, i.e. $4/2 = 2$.

For Onsager ($d = 2$):
(i) Rushbrooke's law $0 + 2(1/8) + 7/4 = 2$, i.e.$8/4 = 2$,
(ii) Widom's law $7/4 = (1/8)(15-1)$, i.e. $7/4 = 14/8$,
(iii) Fisher's law $7/4 = (2-1/4)1$, i.e. $7/4 = 7/4$,
(iv) Josephson's law $d1 = 2 - 0$, i.e. $2 = 2$.

2.3.3 The renormalization group (purely CMI)

The explanation of scaling laws led to a major theoretical breakthrough that was another triumph for the CMI in statistical physics. The adventure began with the discovery by Benjamin Widom[45] (Widom 1965) that they can be deduced from *homogeneity* properties of the *singular* part of the potential F (free energy) and of the correlation function $\Gamma(r)$ in the neighborhood of the critical point:

$$\begin{cases} F(c\Delta T,\ c^{\beta}\eta) = c^{(2-\alpha)}F(\Delta T, \eta) \\ \Gamma(r) \sim r^{-(d-2+\eta)} g\left(\frac{r}{\xi}\right). \end{cases}$$

This means that F is a *fixed point* of the transformation $T \to cT$, $\eta \to c^{\beta}\eta$, $F \to c^{(2-\alpha)}F$. In other words, the fundamental empirical laws extracted from both experimental and computational data on a large number of magnetic systems (which is already a remarkable achievement) can be summed up as a *structural* property of functional homogeneity and, as a result, expressed in terms of fixed points of appropriate transformations in appropriate functional spaces. As Gérard Toulouse and Pierre Pfeuty (1975) pointed out, this leads to a complete rethinking of the problem of phase transitions within the framework of general theories of dynamical systems,

> "theories [which], due to their global perspective and emphasis on universal properties, constitute [for the analysis of critical phenomena] a new framework for thinking."

This new framework of thought is known as the *Renormalization Group* (RG). Its guiding principle is that of *scale invariance*. It was introduced in the second half of the 1960s by Leo Kadanoff and Benjamin Widom, then developed by Kenneth Wilson (Nobel Prize winner in 1982). As Wilson emphasized

[45] Benjamin Widom was awarded the Boltzmann Medal in 1988.

(Wilson-Kogut 1974), the fundamental physical fact manifested by phase transitions is the existence of large-scale fluctuations which, at the critical point, no longer have a characteristic length. This therefore leads to treating the problem at all scales, with two consequences. First, on a methodological level, it leads to breaking down the problem into sub-problems that do have characteristic lengths. Such a method will be iterative (recursive). Second, on a mathematical level, it leads to *self-similar* structures with internal properties of homothety that implies their *fractal* nature. Thus, the issue of a fractality imposed by very simple constraints of scale invariance has become the key to understanding critical phenomena. As Gérard Toulouse and Pierre Pfeuty (1975) also stated

> "It does not seem exaggerated to see this as a new level of physics, endowed with a system of original and powerful concepts."[46]

The method consists of using the property of scale invariance and the absence of characteristic length at the critical point to drastically reduce the number of degrees of freedom in the system.

Let us assume that the spin lattice has a mesh μ. Near the critical temperature T_c, the correlation length ξ becomes very large (macroscopic) compared to the lattice mesh. Large-scale fluctuations are therefore overwhelmingly dominant, and small-scale fluctuations can be eliminated. To do this, the spins are replaced by *blocks* of spins with a lattice mesh λ, where $\mu << \lambda << \xi$,[47] each block having a spin equal to the average of the (strongly intercorrelated) spins belonging to it. The lattice of the new network is then renormalized to make it equal to the initial lattice mesh μ. Averaging reduces the number of degrees of freedom and transforms the given problem into a problem of reduced correlation length $\xi/s < \xi$, where the parameter s is the change of scale used to renormalize the lattice mesh.

These operations are described by an operator R_s – known as the *renormalization operator* – which acts on the function F describing the system (Hamiltonian, free energy, partition function). Since $R_t R_s = R_{st}$, the operators R_s constitute a semigroup operating on the functional space $\mathfrak{F}$ of functions F. At the critical point, the correlation length ξ diverges, which means (since $\xi/s = \xi$) that F is a fixed point of the operators R_s. One can therefore apply standard methods of group action analysis in the neighbourhood of a fixed point, and the critical exponents then appear as special cases of the well-known *characteristic exponents* in this type of situation.[48]

It is thus by taking into account a new fundamental phenomenon, that of fractality induced by scale invariance properties, that specialists have begun to explain scaling laws. We can see how the convergence of the work of many em-

[46] To draw a parallel with classical mechanics, we could say that scaling laws are equivalent to Kepler's laws, Widom's homogeneity properties are equivalent to Newton's law of gravity, and the Kadanoff-Widom-Wilson renormalization group is equivalent to the Newtonian model of Keplerian systems.

[47] << means "much less than."

[48] Readers interested in more details on RG can consult our compilation (Petitot 1982).

inent physicists, mathematicians, and computer scientists has made it possible to considerably deepen the methodology and broaden the scope of the CMI.

Chapter 3

Natural self-organized complex systems

After discussing the CMI in thermodynamics and dealing with complex systems such as magnetic systems, we will now use an example to show how such models can migrate from one domain to another and reveal profound analogies in formal structures between domains whose phenomenology, conceptuality, and ontology are radically different. This transdisciplinarity is a consequence of the fact that computational syntheses operate in the inverse way to conceptual analyses.

The example is that of the transfer of magnetic models of *spin glasses* in statistical physics to *neural networks* models.

3.1 Spin glasses

Among the spin models of energy $E(s) = \sum_{i,j} J_{ij} s_i s_j - H \sum_i s_i$ ((s) is the configuration of spins s_i at sites i in the lattice), physicists have paid particular attention to materials (mainly metal alloys containing spin impurities) called "spin glasses" where the interactions are mixtures of couplings J_{ij}j of *opposite* signs, with couplings > 0 (ferromagnetic) pushing the spins to align and couplings < 0 (antiferromagnetic) pushing the spins to oppose each other.[1] These are necessarily "frustrated" systems where the units must carry out contradictory injunctions. The state function (free energy) of these highly complex systems has *many local* relative minima of low energy, and there are therefore many *metastable* states in which the system can become trapped while trying to reach the ground state corresponding to the absolute minimum.

To access global absolute minima, traditional methods such as "gradient descent" become ineffective. Sophisticated algorithms such as "simulated annealing" must be used. Noise is introduced into the system, i.e., a "computational

[1]These are, in a way, "social" systems whose units (atoms) have "opinions" (spins) that are not only different but possibly "politically" opposed.

temperature" T, and one starts from an initial state $(s)_0$ of the network at high temperature so as to have access to all the attraction basins. Then a configuration $(s)_1$ close to $(s)_0$ is chosen at random. If E decreases by $\Delta E < 0$, the transition is made. If, on the other hand, E increases by $\Delta E > 0$, the transition is made but only with probability $1/(1+exp(\Delta E/T))$, which allows to step over the thresholds. The process is then repeated until a local minimum is reached, then the temperature T is lowered and the process is repeated following a sequence T_n. It can be shown that, if the sequence T_n decreases slowly enough, this stochastic process converges when $T \to 0$, towards an *absolute* minimum on which the Gibbs distribution $\frac{1}{Z}exp(-\frac{E(s)}{T})$ is concentrated.

3.2 Neural networks and CMI in neuropsychology

A major interdisciplinary scientific breakthrough was the discovery by physicist John Hopfield[2] that the equations for neural networks previously developed by specialists such as Hugh Wilson and Jack Cowan in the early 1970s were, formally speaking, generalizations of those for spin glasses (Hopfield 1982). This revealed a powerful *formal analogy* between phenomenologically, methodologically, and ontologically heterogeneous domains, a structural isomorphism that once again demonstrated the *universality* of certain equations. This made it possible to revive the long-standing and vast program of neurophysiological explanation of psychological states, which is a fascinating micro → macro example in life sciences.

The idea of interpreting mathematically psychological structures in terms of complex neural networks means that activities of billions of elementary microneurons connected through specific complex connectomes called "functional architectures" can implement a macrolevel of emerging psychological states and processes.

This research program has become one of the main aspects of the CMI and is now ubiquitous in cognitive science. It is the issue of some of major international research programs.[3] It can be said that today, CMI dominates the understanding and explanation of cognitive phenomena with an impressive number of experimental results (particularly from neural imaging) and models. The imaging revolution has made it experimentally verifiable by turning the brain's "black box" into a more "transparent box."

As we have already pointed out, the formalization of neural networks goes back to 1940-1950 with John von Neumann, Norbert Wiener, Warren McCulloch, Walter Pitts, and the famous Macy Conferences. One of his first key ideas was to build supervised learning algorithms and "input → output" systems that

[2]For which he was awarded the Nobel Prize in 2024 with Geoffrey Hinton.

[3]As the billion euros *Human Brain Project* aiming to simulate cortical modules of the visual cortex. It uses a computational power up to a million teraflops (a teraflop is 1000 billions operations per second).

move from a micro-neural level to a macro-psychological level capable of recognizing which category a stimulus should be classified under. For example, in an input layer, we have N pixel-type receptors that detect the light intensities and specific colors of an image (a kind of retina), and we assume that the output layer is composed of n labels ("cat," "bottle," "flower," "car," etc.). Within the system, there is a functional architecture of connections, and we want (this is supervision) this architecture to enable the network to give the correct response for a certain number p of examples. The processing is bottom-up (feed-forward).

It is important to note that such a possibility is not at all self-evident, as it runs up against the barrier of large numbers. For example, the number of square black and white images with 20 pixels on each side (which are images of ridiculous simplicity) is 2^{400}, a large number of the order of 10^{120} which is immeasurably greater than the number of atoms in the universe, estimated at 10^{80} ! A learning corpus of one billion images to be correctly distributed across a thousand categories is therefore *infinitesimal* compared to the number of images that could potentially activate the input layer.

The first supervised learning models (such as the "Perceptron" introduced in 1957 by Frank Rosenblatt) were far too basic. They were improved by introducing multi-layer networks where the different layers perform intermediate categorizations following a mereological hierarchy. These are already quite efficient. They are particularly interesting for the epistemology of CMI because they multiply the micro-macro levels into a *hierarchy* of levels. They are often called *connectionist* models in reference to the constitutive role of connections. As David Rumelhart and James McClelland stated in their book *Parallel Distributed Processing* (1986), connectionist models aim to model the "microstructure of cognition." The link with thermodynamics is self-evident, since Geoffrey Hinton and Terry Sejnowski refer to them as "Boltzmann machines"!

Next, *convolutional* networks were introduced, inspired by the way biological neural networks function. For example, in the visual system (see *Elements of Neurogeometry*, Petitot 2017), we start with the retina, which is the first layer. In fact, it already contains several sublayers (five) that transmit signals from photoreceptors to ganglion cells, whose axons form the optic nerve. The lateral geniculate body is an intermediate relay before the primary visual cortex. In the first cortical area $V1$, neurons (called "simple") locally detect retinal positions and directions. They act as filters on the retinal signal J they receive. Each has a "receptive field" D, which is the small retinal area of ganglion cells to which it is connected. It also has a "receptive profile" (its transfer function) P with which it convolves the signal J restricted to its receptive field D, allowing it to detect a feature (e.g., an orientation). For each feature, there is a layer of neurons whose receptive fields D overlap and cover the retinal domain and whose receptive profile P is of the same type. This layer therefore convolves the signal J with the receptive profile P.

Long-range *intracortical* connections with a very specific functional architecture then enable these local features to be *integrated* globally and, for example, locally detected orientations to be integrated into global object contours. This results in emerging *shapes* at the macro level. There are also "blobs" that detect

local color values. Modules in post-$V1$ areas then complete contours hidden by occlusions, process textures, process information provided by shadows and other aspects of images, and then even more posterior areas perform shape recognition and categorize them. All these layers, modules, and areas are hierarchical and operate not only in a feed-forward (bottom-up) manner but also with a great deal of feedback (top-down).

The analogy between neural networks (even single-layer ones) and spin glasses triggered a new period of intense research. In the simplest possible neural network, where activity states are reduced to 1 (active) or 0 (inactive), the basic equation for spin glasses $E(s) = \sum_{i,j} J_{ij} s_i s_j - H \sum_i s_i$ is reinterpreted in the following way: $E(s) = -\sum_{i,j} w_{ij} s_i s_j + \sum_i t_i s_i$. The $s_i = 1, 0$, are the activation states of the neurons u_i occupying the sites of the network[4] and the w_{ij} are the "synaptic weights" of the connections between neuron i and neuron j: w_{ij} is > 0 if the connection is excitatory and < 0 if the connection is inhibitory. Unlike in physics, these w_{ij} are not symmetrical but *oriented* (spikes propagate in a certain direction along the axons), which makes the model much more complex.

The network "computes" as follows. Each neuron u_i receives afferent signals from its presynaptic neurons and changes its internal state based on the ratio between the sum $a_i(t) = \sum_j w_{ij} s_j(t)$ $(j \neq i)$ of these signals weighted by the synaptic weights w_{ij} (i fixed) and a certain threshold t_i specific to it. When it activates, it transmits its activation to the neurons to which it is connected. We therefore have evolutions of the type

$$s_i(t+1) = \mathfrak{h}(a_i(t) - t_i) = \mathfrak{h}\left(\sum_{j \neq i} w_{ij} s_j(t) - t_i\right)$$

where $\mathfrak{h}$ is the Heaviside step function[5]. They are related to the partial derivatives of $E(s) = -\sum_{i,j} w_{ij} s_i s_j + \sum_i t_i s_i$ with respect to s_i and are highly nonlinear due to the Heaviside step $\mathfrak{h}$. The dynamics of the network are therefore those of a system of threshold automata whose the w_{ij} and the t_i range over a control space.

What makes these neural networks much more complex than spin glasses is that, as we have said, the synaptic weights are *not* symmetric. This entails that $E(s)$ no longer has the status of an energy to be minimized and that the dynamics are much more complicated than a gradient descent attracting the state of the system towards asymptotic states of stable equilibrium which are point attractors (the minima). The attractors may have a topology and *internal* structure (attractive cycles, "strange" attractors). The system can also become chaotic by following universal "path to chaos" processes.

From the 1970s and 1980s onwards, neural network theory experienced a veritable scientific and technological explosion (and an economic one too, since the largest groups, Bell Labs, AT&T, etc., then Google, Facebook, etc., dedi-

[4] To reinforce the analogy with spin glasses, we can make the change of variables $\sigma = 2(s - 1/2)$. If $s = 1$ or 0, then $\sigma = 1$ or -1.

[5] $\mathfrak{h}x) = 1$ if $x > 0$ and $\mathfrak{h}(x) = 0$ otherwise.

cated entire laboratories to it), leading to an extraordinary transfer of computational and methodological tools from the statistical physics of spin glasses to neuropsychology. This is one of the greatest transfers of models between phenomenologically and ontologically heterogeneous fields in the history of science. The successes of CMI in neuropsychology have given rigorous status to the thesis of René Thom and Christopher Zeeman (Thom 1972b, Zeeman 1977) that mental contents are *attractors* of neurally implemented dynamical systems and that, consequently, cognitive functions can be conceived in terms of statistical physics and qualitative dynamics.

Two particularly important cognitive processes are *categorization* and *learning*. Categorization results from the partitioning of the configuration space of the network states (s) into *basins of attraction* $B(A)$ of attractors A. The basins $B(A)$ are the "categories" and the attractors A are the "prototypes". What psychologists call "typicality gradients" can then be interpreted as a kind of gradient descents (known as Lyapunov functions) in the basins $B(A)$, descents that push the system towards the attractors A.

As for learning, it can be interpreted as an *inverse* problem. The direct problem is, given the matrix (w) of synaptic weights, to find the attractors of the dynamics. In learning, on the contrary, given a priori the desired states (called "associative memories"), the aim is to find a matrix (w) that makes them attractors. Certain algorithms have been developed for this purpose, in particular the "backpropagation" algorithm, which consists of teaching the system to correct its errors. In the case of symmetric synaptic weights where the attractors are the energy minima, one of the most widely used learning rules is Hebb's law, which states that the connections used to perform a task are strengthened, while unused connections are weakened. If the desired attractors are states $(s)_k$, $k = 1, \ldots, p$, one takes $w_{ij} = \sum_k (s_i)_k (s_j)_k$. This rule works well as long as the ratio p/N of associative memories to the number of units in the network remains below a critical value $c = 0.144$. Above this limit a *phase transition* known as "forgetting catastrophe" occurs, causing the system to fail to categorize correctly. Daniel Amit's classic work *Modeling Brain Function* (Amit 1989), devoted to "attractor neural networks," provides numerous details on this situation.

Other types of learning have been studied in detail beyond supervised learning: reinforcement learning (only through trial and error, without supervision), deep learning, etc.

It should be added that, since the 1990s, revolutionary new techniques of in vivo optical imaging have provided precise access to the *geometry* of the *connectivity* of the functional architectures of certain cortical areas, particularly those of the primary visual areas. This geometry is very rich mathematically and can be integrated into neural networks. Such *neurogeometric* approaches have made it possible to explain a great many perceptual phenomena that had long been identified but could not be explained, in particular some well known laws of Gestalt theory (see *Elements of Neurogeometry*, Petitot 2017, 2024).

Let us highlight a few epistemological aspects of the transfer of spin glass models to neural network models.

(i) Spin glasses are interesting for neurophysics because they have a large number of local minima that can therefore model many macro states (s) of systems of micro units u_i with internal states s_i interacting by means of couplings w_{ij}.

(ii) With asymmetric couplings, attractors can be more complicated than minima and have an "internal" structure that can model a macro "content" of macro states (s).

(iii) When generalized, spin glass models are used in a field very different from the physics of magnetism, and *new* questions relevant to this new field will be posed to the model. For example, for neurocognitive networks, these will be psychological questions about categorization and learning.

The massive and operational development of neural networks is still in its infancy, but it has already significantly transformed the epistemological debate on CMI. Indeed, it is the basis of artificial intelligence (AI) technologies, which are now involved in all areas of human knowledge and activity (see Le Cun et al. 2015). As they represent an undeniable success of CMI, the question is no longer whether it is possible to explain certain cognitive faculties and human actions in terms of CMI, since its possibility is now a *factum rationis*. On the other hand, we can of course question the fundamental differences that exist between AI techniques, which are certainly very powerful but highly specialized, and human intelligence, which involves a great deal of background knowledge, sensitivity to context, heuristics, and rules of common sense. We may also wonder how to judge a world in which all human knowledge and activities are duplicated by artificial intelligence.

3.3 Sandpiles, sanddunes and self-organized criticality

Until now, the physical models we have discussed concerned the emergence of a macrophysical level from a microphysical level. They therefore went from the physical to the physical. But there are other types of situations where the emerging structure is no longer physical in nature but *geometric* and *morphological*.

A typical example is provided by the geometry and the morphological dynamics of sandpiles and sandunes. Here, the micro elementary entities are grains of sand locally interacting by rolling against each other. The micro dynamics is very complicated and the physics of dry cohesionless granular materials is a very active discipline. The macro geometric shape emerges from a *fractal* entanglement of myriads of micro *avalanches* of different scales, and geometric macro parameters such as the slope of the dune or pile (called the "angle of repose") can be explicitly computed using non-linear equations. It is a perfect example of a complex *spontaneous order*. The remarkable fact is that this geometric slope is a critical value and the complex multiscale system of sand grains has the strange property of *stabilizing* on its *critical* state, while criticality seems to

be the opposite of stability. This phenomenon called "self-organized criticality" was discovered in 1987 by Per Bak, Kurt Wiesenfeld and Chao Tang (Bak et al. 1987). Once again, there is no possible semantic reductionism: the geometrical concepts used to adequately describe the morphology of dunes have nothing to do with fractal micro-avalanches of rolling sand grains.

When wind is added as an "external field," as is the case for dunes in the desert, the morphology of the dune acts as an obstacle and induces all the well-known phenomena of front-side overpressure, rear-side underpressure, and turbulence on the air flow. It therefore acts in a top-down manner on granular microdynamics and generates complex morphologies analyzed by specialists such as Stéphane Douady (Douady-Hersen 2011).

3.4 Towards a social physics of collective behaviors

The profound advances in statistical physics have led to a reassessment and equally profound advances in the transfer of models to *social science* fields such as economic equilibrium, financial markets, dynamics of political opinion, decision-making, in short, all the themes of "social physics" that have long been inspired by probability and thermodynamics.

Social physics is an old tradition. It can be traced back to Hobbes or, if you prefer, to the first discussions on probability (Pascal and Fermat). Daniel Bernouilli (1700-1782) invented the concept of utility. Then the Marquis de Condorcet (1743-1793), the Comte de Saint Simon (1760-1825), Adolphe Quetelet (1796-1874), and Auguste Comte (1798-1857) began to theorize a "political arithmetic" using probabilities to evaluate, for example, voting systems. Then, based on statistical thermodynamics, great theorists such as Vilfredo Pareto and others we have already mentioned developed models that became influential.

Social physics is typical of CMI. By definition, it depends on transfers of physical models. The limitations of its classical models are obvious. For example, in classical economic equilibrium theory, the axioms of rationality of agents maximizing their utility are unrealistic. The models do not take into account the fact that, in general, the optimum is unattainable and fragile. Similarly, in finance, classical models of stock market movements and financial risks assume *Gaussian* fluctuations. These are "mean field" models that do not take into account rare but large-scale events such as stock market crashes. To explore them further, we need to move on to "renormalization group" approaches.

Numerous research programs have been developed. We mentioned them earlier: the Santa Fe Institute, the MIT Media Lab (see Pentland 2014), and the Paris Institute of Complex Systems. They have profoundly transformed the status of the social sciences toward what David Lazer has called "computational social sciences." From our perspective, this is effectively a shift from conceptual analysis to computational synthesis.

For example, in *econophysics*, Marc Mézard[6] and Jean-Philippe Bouchaud[7] (Bouchaud-Mézard 2000, Bouchaud 2023) have constructed multi-agent economic models composed of N agents with different (variable) statuses (employers, employees, unemployed) and distributing a certain money supply M according to salary ranges, hiring, firing, and consumption rules, etc. These models incorporate well-known empirical laws such as Pareto's law for income distribution and Zipf's law for company size (see above section 1.4). They can be used to test growth or recession processes, wage or unemployment dynamics, and many other economic and social processes.

In *sociophysics*, physicist Serge Galam began using the physics of magnetic systems to study certain social phenomena in the early 1980s. It is interesting to see how a physicist thought about things at the time. In (Galam et al. 1982[8]), after acknowledging, of course, that the social sciences are immensely far removed from the physical sciences, he remarks:

> "There is, however, one aspect of human behavior which is directly related to physics; namely collective behavior." (p. 2)

In social systems, there are dimensions and factors that are *only* related to collective behaviors characteristic of complex macro-general systems involving the interaction of a large number of micro-elementary units. This is what gives the CMI its transdisciplinary relevance in these fields.

For example, suppose we have a system of N agents u_i that can be in an activity state s_i ranging from -1 to $+1$ and reflecting a choice. The mean activity is $M = \frac{1}{N}\sum_i s_i$. If the agents do not interact and do not have their own preferences – in other words, if the system is "atomistic" – the individual probabilities are symmetrical between the two possible states: $p(s_i) = \frac{1}{2}$ for $s_i = \pm 1$. Since individuals are independent, the probability of a configuration $s = (s_i)$ is $\left(\frac{1}{2}\right)^N$. All configurations are equiprobable, and therefore the mean of M is $\langle M \rangle = 0$. The system is macroscopically neutral; there is no collective tendency towards a preferred choice.

The agreement and disagreement between two agents u_i and u_j can be interpreted by the product $s_i s_j$. Indeed, it is equal to $+1$ if the two agents are in the same state and to -1 if the two agents are in different states. Let us then assume that the agents are linked by mutual *interactions* that can be described by couplings J_{ij}. In many situations, one can therefore assume that the term to be optimized is of the multiplicative type $J_{ij}s_i s_j$. which will introduce a variational principle of conflict minimization. Suppose J is constant and > 0. Minimization leads to a dynamic alignment of states, which breaks the symmetry of the neutral macro state $\langle M \rangle = 0$ and generates the two optimal macro states $M = +1$ and $M = -1$, which are extremely polarized behaviors.

Let us further assume that the context can be described by an external field H, a kind of "social pressure" influencing the state of the agents, for example

[6] Lars Onsager Prize, 2016.
[7] Lars Onsager Prize, 2024.
[8] See also Galam-Moscovici 1991.

(it's the easiest), in a linear manner Hs_i. Depending on the sign of H, this will bias the optimal collective state toward one of the choices. As a result, we will end up with a "free energy" to be minimized of the type $E(s) = \sum_{i,j} J_{ij} s_i s_j - H \sum_i s_i$ and we will be able to apply physical formalisms of the phase transition type.

It should be noted that the "atomistic" case without interactions simply corresponds to the special case of couplings $J_{ij} = 0$. For CMI, "atomism" is merely a choice of parameters and it is therefore irrelevant to try to make it a philosophical issue.

Many social phenomena can be modeled using this type of formalism. The computqations are complicated, but they can be simplified to reveal qualitative structures. For example, we can assume that the J_{ij} are > 0 and all equal, and introduce a *mean field* hypothesis that "externalizes" the interactions between agents by modifying the external field H. The idea is as follows. For the "canonical" set, we saw above that the probability law $P = (p_s)$ of the configurations $s = (s_i)$ is given by $p_s = \frac{1}{Z_P} exp(-\beta E(s))$ with the partition function $Z_P = \sum_s exp(-\beta E(s))$ ($\beta = \frac{1}{kT}$), where the temperature T is a "computational temperature" that

> "measures the potential response of an individual to some influence due to the social atmosphere he is embedded in" (Galam et al. 1982, p. 6).

We will seek a linear probability law Q in the s_i of the form

$$q_s = \frac{1}{Z_Q} exp(-\sum_i a_i s_i).$$

Self-consistency computations show that the best approximation is given by a_i of the form $a_i = \beta(H + \sum_j J \langle s_j \rangle)$, where $\langle s_j \rangle$ is the average of the states of the agents connected to u_i. This gives us a model in which interactions between agents are replaced by the addition of a mean field $h_i = \sum_j J \langle s_j \rangle$ to the external field H.[9] If we calculate the free energy F, we obtain a function of the mean activity $M = \frac{1}{N} \sum_i s_i$ which depends on the controls H, $K = T/J$. M operates as an order parameter of the system and we meet again a situation analogous to that of phase transitions. There exists a critical value K_c of K. For $K < K_c$ there are two ordered phases, one corresponding to a minimum $0 < M \leq 1$, the other to $-1 < M \leq 0$. For $K = K_c$ these two phases merge, and for $K > K_c$ there is only one disordered phase, with $M = 0$.

Serge Galam has also studied *opinion dynamics* in elections and referendums and modeled voting outcomes in hierarchical systems (Galam 2008). He imagines what "active" spin systems might look like, where spins would be "activated" by internal dynamics and interact in "multi-layered" structures. The idea is that there are several levels of groups, and each group at a certain level

[9] There is indeed a self-consistency problem that complicates the computations, because the action of h_i on si depends on the average of the s_i.

seeks to reach consensus on a choice (opinion alignment mechanism) and passes the result on to a group at the next level. To resolve indecision, external factors (shared beliefs, values, ideologies) are used, which act as external fields, and a bias is introduced in favor of maintaining the statu quo for individuals who aren't motivated enough. It is then shown that for a change of opinion (a reform) to take place, the initial state must include *much more* than 50% of agents who are a priori favorable to the reform.

A host of other sociophysical models are available. Interested readers may consult, for example, Joshua Epstein's treatise *Generative Social Science: Studies in Agent-Based Computational Modeling* (Epstein 2008), where they will find numerous models of strikes, civil violence, and rebellions, as well as cooperation. This title is very interesting because it is, in a way, a *Manifesto* for computational syntheses of emerging macro properties using the algorithmic *generativity* of mathematics. It also identifies Agent-Based Modeling as an aspect of CMI and attempts to identify highly transdisciplinary "core generative mechanisms."

Numerous studies have also been devoted to the problems of *segregation* dynamics among human groups. Among a host of examples, interested readers can consult (Zakine et al. 2023). This Sakoda-Schelling model of urban segregation shows that small individual preferences (micro-motivations) regarding choice of residence can lead to *un-intended* complete macro-level segregation. "Bridging the gap from *micro to macro*" involves collective states analogous to thermodynamic phases, with phase transitions and mean-field approximations.

In addition to these sociophysical models, the irresistible development of internet networks for email, file sharing, e-commerce, blogs, medical records, social media, and Twitter or X has created a new universe of complex systems with massive connectivity, in which all interactions and transactions are saved on servers and thus become *accessible* for computation thanks to powerful processing and exploration algorithms (big data, data mining). Now, we can see *empirically* that these systems, with their specific characteristics, exhibit many properties typical of complex systems in physics, in particular *scaling laws* and *critical exponents*. This shows that there are universal laws governing the propagation of information in this type of connectivity, with fractal information avalanches and critical processes giving rise to macro collective structures. The CMI approach is therefore both empirically observable thanks to powerful data acquisition methods and modelable using well-known models in the physics of critical phenomena.

Specialists have shown that scaling laws can easily be explained in this case based on two assumptions:

(i) the system is open and can grow by recruiting new sites,

(ii) new sites preferentially connect to sites that are already highly connected (preferential attachment).

For example, in a pioneering article published in 1999, Albert László Barabási and Reka Albert found that in several very different mega-networks, the probability $P(k)$ that a site in the network is connected to k other sites satisfies a

Pareto, Zipf, Gibrat[10] power law $P(k) = Cexp(-\gamma)$ where γ is a critical exponent of magnitude between 2 and 4 (Barabási-Albert 1999). From the "statistical physics" of these large micro-connectivities, we can extract the macro *social structure* of the communities that are their source. We can reconstruct the socio-semantic dynamics and hierarchical clustering at work. The new methodologies and technologies used for this purpose open up a new techno-scientific and economic dimension of the CMI.

3.5 Operativity and efficiency of CMI

From classical thermodynamics to neural networks and sociophysical models, we have looked at several examples where CMI is not only operational but is by far the most efficient theory.

These examples share some common epistemological features:

(1) The two (or more) micro / macro levels are phenomenologically heterogeneous and conceptually autonomous. Their relation is a relation of emergence.

(2) But they are not causally autonomous. The macro-level "supervenes" on the micro-level. Moreover, causality operates more bottom-up from the micro to the macro level and top-down reverse causality is not always present but can be induced by external factors. The slope of the sandpile does not act on the physics of the sand grains but act on the air flow. In life sciences or social sciences top-down causalities become essential.

(3) The heterogeneity of the different levels results from the immense number of elementary micro units. The local interactions between these units must be iterated (in a sequential or parallel mode) and the iteration of elementary rules can reach any possible complexity.[11]

(4) The relation of emergence between the levels results from their incommensurability.

(5) An emerging macro structure can be rather simple (a temperature, a slope). But its emerging simplicity encompasses an immeasurable micro complexity.

(6) In physical situations, the system is frequently modelled by a large system of differential equations and the emerging properties are represented by *order parameters* which have a statistical content (as in "mean field" theories).

(7) The conceptual independence of the emerging structures is *mathematically justified* by proving that the order parameters are essentially *independent* of the fine grained underlying microphysics (see the renormalization group). Without such proofs of *invariance* with respect to *multi-realizability*, emergence would not have a strong scientific content.

So, to summarize, CMI posits that high level phenomena, structures and processes can be reconstructed and computationnally synthesized, as far as their scientific explanation is concerned, from underlying lower level phenomena,

[10] See section 1.4.

[11] It is a theorem for networks of automata.

structures and processes. But we have seen that CMI is by no means reductionnist in the "eliminativist" or "epiphenomenal" sense and is perfectly compatible with emergence, "supervenience" or "functionalism". Functionalism means that macro structures having a *functional* role can exist only if they are materially implemented in an underlying material substrate, but are at the same time, as functionally meaningful structures, largely independent of the fine grained material structure they are implemented in. Beside physical examples, another paradigmatic example is the opposition software/hardware in computer sciences (see philosophers like Hilary Putnam, Jerry Fodor, Zenon Pylyshyn, etc.).

Complex systems having different levels of reality whose emergence at different scales can be proved, are countless in nature: critical phenomena, percolation, self-organized criticality, reaction-diffusion systems, far from equilibrium dissipative structures, turbulence, cellular automata, neural networks, but also in ethology, economy and social sciences : ant colonies, swarms, stock markets, Pareto law for incomes, etc. The emerging levels are not epiphenomenal but provedly autonomous. Yet, they are not interpretable in terms of a holist realism positing their irreducible reality. They are causally reducible but not semantically reducible.

Chapter 4

From ethology to Game Theory

In the previous chapters, we briefly explained the basic formalisms of statistical physics and how they can naturally be applied to certain types of other complex systems, from neural networks to sociophysical systems.

In this chapter let us progress towards game theory through biology, psychology and ethology.

4.1 Hayek's Sensory order

We discussed above the neural networks that have become ubiquitous in neuropsychology and AI. It is relevant to recall in our context that in fact, prior to their "spin glass" physicalist formalization, as acknowledged by as important neuroscientists as Gerald Edelman (Nobel 1972) or Joaquin Fuster, the great precursor was Friedrich von Hayek in *The Sensory Order*. This unique masterpiece of CMI has been analyzed by Bruce Caldwell, Barry Smith, and one of us (Francesco Di Iorio) who all emphasized the importance of this early contribution of Hayek to the neurophysiology of psychology.[1] It goes back to the 1920s and was taken up in the 1952 book (Hayek 1952). It is essential in Hayek's intellectual journey because it was for Hayek a key example of a *complex spontaneous order*: neural cells are connected into complex, organized and specific networks which act as communicational infrastructures for complex fluxes of neural spikes ("impulses" said Hayek) and encode learning and memory. Hayek understood very early on, *well before* all the physical and mathematical advances we have discussed, which began in the 1970s, that we have to take this micro neural organization as a *basis* for explaining our macro psychological sensory-motor experiences. The specificity of what are now called "neural functional architectures" introduces a gap between, on the one hand,

[1] We refer to Part 1, Chapter 1 for information on Hayek.

the sensorial micro transduction of the stimuli by peripheral receptive organs such as the retina, and, on the other hand, the perceived scenes endowed with their Gestalt patterns processed by the central cortical brain areas.

In *The Sensory Order After 25 Years* (Hayek 1982), Hayek's summary is very akin to the contemporary concept of functional architecture:

> "Mind thus becomes to me a continuous stream of impulses, the significance of each and every contribution of which is determined by the place in the pattern of channels through which they flow within the pattern of all available channels." (p. 291)

These initial reflections played a role of model for the further works of Hayek in economy and, after he had come into contact in the 1950's with the transdisciplinary and polymathematical sciences of complex self-organized systems (Norbert Wiener, Ludwig von Bertalanffy, John von Neumann, Warren Weaver, see above), he wrote his 1952 book.

In the Hayek's archives managed by Bruce Caldwell at the Hoover Institute of Stanford, there are interesting letters with James Gibson, the future author of *The Ecological Approach to Visual Perception* (1979), who has been "much impressed" by *The Sensory Order* (24 February 1954 letter) and wanted to invite Hayek at Cornell for a Conference on "Fundamental Problems of Perception". There are also interesting remarks concerning Heinrich Klüver (1897-1979), the great neuropsychologist at Chicago and a privileged interlocutor of Hayek in psychology and neurology. Klüver was a student of Max Wertheimer and introduced Gestalt theory in the US. He arrived at Chicago in 1928 and joined the "Neurology Club" (Karl Lashley, Percival Bailey, A. Earl Walker, Ralph Gerard, Stephen Polyak, Charles Judson Herrick, and Roy Grinker). His works on vision are now well modelled using Wilson-Cowan-Hopfield equations and tools of Neurogeometry (Petitot 2008, see above).

4.2 Ethology: flocks, herds, schools, swarms

Other fundamental examples pf CMI in life sciences are found in *ethology*. In the last decades, a great deal of research and modelling have been dedicated to the collective behavior of large communities of animals: flocks of birds, herds, fish schools, insect swarms, etc. Figure 4.1 shows an example of a flock of starlings. A similar image was featured on the poster for a symposium held by the Paris Academy of Sciences on October 11, 2022, in honor of physicist Giorgio Parisi, winner of the 2021 Nobel Prize

> "for his discovery of the interaction between disorder and fluctuations in physical systems, from the atomic to the planetary scale".

These very complex global geometric patterns can result from very simple local rules, as simple as

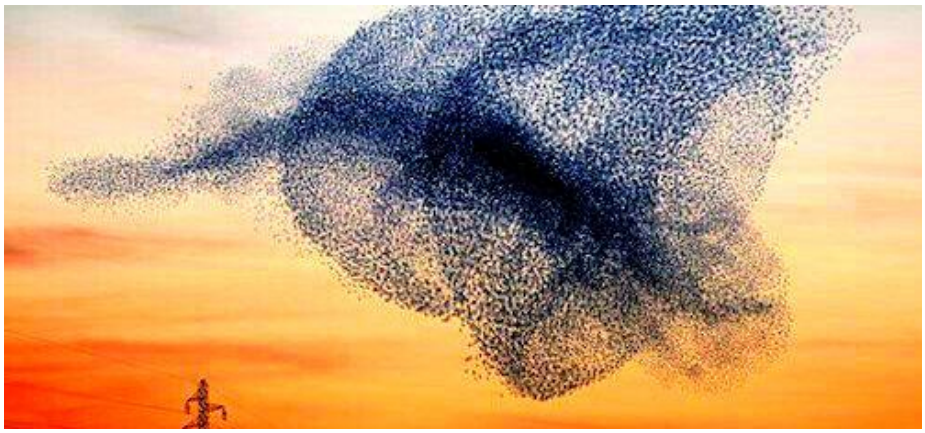


Figure 4.1: An exemple of a flock of starlings.

(i) move in the direction of your neighbors,
(ii) move to the center of the group,
(iii) if another animal is coming too close, move away from it.
Here too, computational synthesis must be highly *generative* and involve powerful computing capabilities to reconstruct *complex global geometries* from *elementary local motions*

What is particularly interesting about these collective movements is that the individuals are no longer particles but evolved (albeit rudimentary) organisms resulting from biological evolution, true agents endowed with sensory-motor faculties and cognitive abilities (albeit limited) and therefore capable of decisions and actions. As they are active and dissipate energy in order to act, they form thermodynamically *out-of-equilibrium* systems and can therefore structure themselves in violation of the second law of thermodynamics.

The overall form of their collective movements is completely unprogrammed, *unpredictable*, and *emergent*. Without any leader, without any external field, and completely beyond the faculties of individuals, it is a *global* macro *un*-intentional consequence of micro-intentional actions governed by *local* laws. It therefore already represents a *social* phenomenon that is a *spontaneous order* completely *transcending* individual capabilities. But it is inexplicable in MH terms because there are no high-level laws that can govern its geometry. CMI is therefore the most reasonable explanatory option.

We cannot overemphasize the fact that the emergence of macro properties that are *un*-intentional consequences of intentional micro actions begins at the most basic level of ethology.

On the computational level, the three rules (i) local alignment, (ii) group attraction, (iii) short-range repulsion, were implemented in particular by Tamás Vicsek and colleagues in a maximally simplified statistical physics model that echoes spin glasses (Vicsek et al. 1995, Ginelli 2016).

We assume that we have N individuals who are always active, moving at speed $v_i = v_0 s_i$ where s_i is the direction (a vector of modulus 1) and v_0 a characteristic modulus that is identical for all individuals. Each individual u_i communicates locally with its neighbors in a neighborhood V_i and interacts with them through attractive cohesive forces and short-range repulsive forces, creating a personal space around each individual. These forces depend on the distance $|r_j - r_i|$ where the r_i are the positions. The positions and velocities linked by $\Delta r_i = v_i \Delta t$ (t is time) therefore *both* play a role in the model, whereas in spin glass models, spins are internal variables that do not affect positions. This is an important difference.

The system is thermalized by adding angular noise using a stochastic rotation R_η of the directions s_i. This yields equations of the type

$$s_i(t + \Delta t) = R_\eta \left(dir \left(a \sum_j s_j(t) + b \sum_j f_{ij}(t) \right) \right)$$

where $dir(v)$ is the direction of a vector v and a and b are two parameters. In the absence of noise, the directions align and the system is in an ordered phase with collective motion. When the noise increases, a phase transition occurs, a disordered phase appears, and the trajectories of individuals become random and independent. The order parameter of the system is the mean direction $\sigma = \langle s_j \rangle = \frac{1}{N} \sum_j s_j$. The phase transition that occurs when the noise decreases is complex and depends on the two parameters (a, b). Areas of alignment are formed, move in complicated ways, and come together and break apart before a global alignment emerges. This results in a *phase diagram*.

As with the physical systems mentioned above regarding the renormalization group, there are critical exponents and universality properties that mean that the emerging macro structures are largely independent of the specific details of the micro individuals.

Vicsek's model, which plays a role similar to that of the Ising model for magnetic media, is two-dimensional and has no cohesion or repulsion terms. It only takes into account local alignment and noise. It has enabled detailed study of the phase transition and has been further developed by numerous authors, such as Hughes Chaté and his colleagues (Chaté et al. 2008). In two dimensions, the directions si are reduced to angles, and in the disordered phase, the order parameter is the average of randomly varying angles, which is known to be of the order $(1/N)^{1/2}$. The critical value of the noise is of the order of $\rho^{1/2}$, where ρ is the density N/V of the N individuals occupying the volume V.

Unlike what happens with magnetic media, the transition is first order. This is because motion changes the *positions* of individuals and therefore local densities. The formation of high-density domains increases the number of neighbors and induces *positive feedback* on alignment. In this case, there are also top-down causalities from the emerging macro level to the underlying micro level. This leads to the formation of ordered alignment domains in a context that remains globally disordered. This phenomenon of phase separation is very spectacular.

4.3 Collective intelligence

But the CMI models for active agents go far beyond the modelling of mere collective motions. They progressively led to the key idea of multi-agent collective, distributed, and decentralized *intelligence* as a new paradigm for complex problem solving, more precisely for solving problems that individual agents are unable to solve. What is now called "swarm intelligence", "distributed artificial intelligence", "ant colony optimization algorithms", etc. The idea is that an algorithm solving a complex problem can be implemented in a distributed way into the collective intelligence of a network of micro elementary agents sharing very *limited* (bounded) cognitive ressources and interacting through very elementary rules. The collective intelligence is *incommensurable* with the individual intelligences and *emerge* from their *global* cooperation.

4.3.1 Social insects

In nature, the most striking examples of collective intelligence are provided by social insects: bee hives, wasp nests, ant nests, termite mounds. This leads us towards politics since social insects are considered since Aristotle as "political animals" (*zoôn politikon*) because they cooperate to produce public goods. It is the key point. In his *Politics*, Aristotle considers that there are essentially two kinds of political species: humans and social insects. A lot of empirical observations and comparisons led him to a behavioral taxonomy. Animals are either solitary (felines, spiders) or social. Social species can live in small groups (primates, wolf packs) or in large groups. In the later case, they can have a gregarious behavior (herds, flocks) or a "political" one. And Aristotle claimed: "man is by nature a political animal", but he is not the only political species. What is true is that man is the only "rational" political species.

What is especially interesting in social insects is that:

(i) The collective intelligence is incommensurable with the individual intelligence of the agents.

(ii) The "political" performances (i.e. collective production of public goods) result from simple but extremely efficient ethological rules selected by evolution.

(iii) The collectivity constructs global architectures, called "architectures without architects" as bees honeycombs with their miraculous hexagonal tesselations, or termites immense mounds (they would be extended up to ten kms at human scale) with their pillars, external walls, galleries, cellars, channels, ridges, spiral conducts for ventilation and cooling, valves, brood chambers, etc. (see figure 4.2).

4.3.2 Stigmergy

Highly sophisticated buildings such as termite mounds are essential for the species to survive. They result from elementary spatial moves marked by pheromones. The process was called *stigmergy* in 1959 by Pierre-Paul Grassé

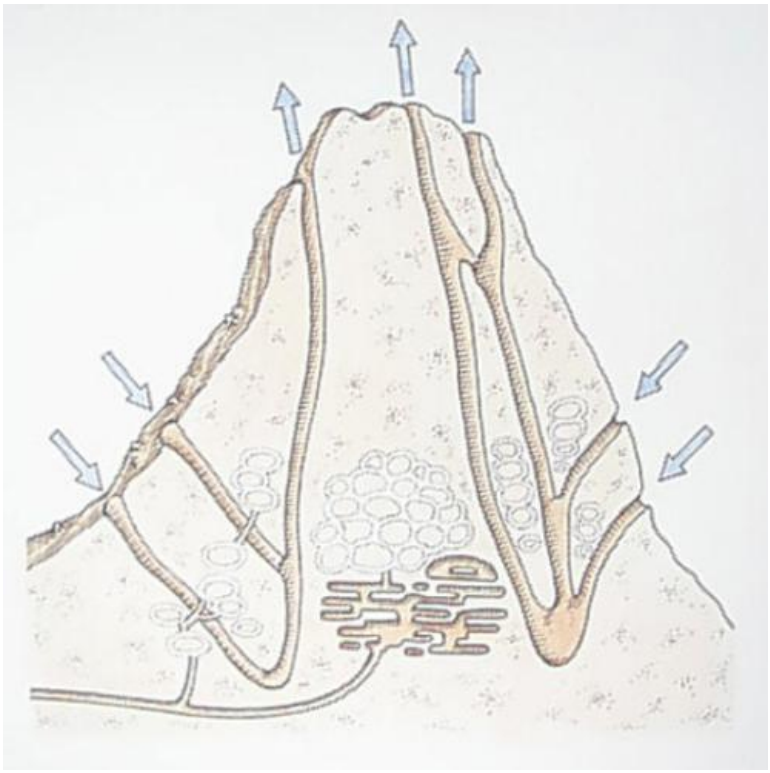


Figure 4.2: Termites' mound.

(Grassé 1982-1984). Stigmergy is the key concept of the theories of swarm intelligence. In constructing a mound, initially, each individual termite rolls a mudball, invests it with pheromones and deposits it randomly in some place. But the pheromones are *attractive* and act as *signs* (hence the term of "stigmergy": stigma = sign, ergon = action). So, many mudballs become deposited in the same place. This positive feedback produces high pillars that collapse beyond a certain height, a catastrophe which triggers the *iterated* building of new pillars, etc. This spontaneous emergence of coherent activity builds gradually, without any centralized planning and control, seemingly intelligent structures.

Their mathematical models are highly non trivial (see, e.g., works by Eric Bonabeau, Jean-Louis Denebourg, Guy Theraulaz 1998). It must be emphasized that, in contrast with the physical examples, the emerging macro architectures exerts here a strong *top-down* "imerging" causality upon the micro individuals.

4.3.3 Eusociality

This very particular paradigm of sociality "invented" by evolution is called eusociality in ethology, a neologism coined in 1966 by Suzanne Batra. Eusociality (from Greek "eu" (good) and "social"), is defined as the highest level of organization of animal sociality and characterized by cooperative brood care, overlapping generations within a colony of adults, and a division of labor into reproductive and non-reproductive groups. The queen and reproductive males monopolize reproduction. Division of labor creates specialized behavioral groups (castes) within the society. Soldiers and workers are sterile and specialized in brood, foraging, defense, maintaining food and resources.

At the evolutionary level, in a strict neo-Darwinian gene-centered view based on the "selfish gene" principle (individuals maximize the fitness of their genes), eusociality seems rather paradoxical. How organisms can increase the transmission of their genes if they become sterile and work for their close relatives? How can we imagine genetic bases for cooperation and altruist behavior? The sociobiologist Bill Hamilton explained (1964) that it could be the case if a non-reproductive (sterile) individual share more genes with a close relative that it could do with its offsprings ("inclusive fitness" principle, Hamilton 1964). And it is precisely the case with many species of social insects which are *haplodiploid*: males are haploid (i.e. have a single set of chromosomes from the mother) while females are diploid (i.e. have two sets of chromosomes from the mother and the father). So sisters can share up to 75% of their genomes, that is more than the 50% they can share with their offsprings. For a critique of Hamilton's inclusive fitness, see (Nowak, Tarnita, Wilson 2010).

4.3.4 Cognitive VS Eusocial complexification

Eusociality is the opposite of the dominant social organisation found in superior species of mammals, that of small "communities" which are like extended family groups. In that later case, the communal links are *accessible* to the experience of the agents. Social organization is "cognitively commensurable and accessible"

and we can speak of a sort of social reflexivity. For an eusocial structure it is not the case. Social organization is not community-based and is *cognitively inaccessible*.

The key point is that there exists (at least) two evolutionary ways for complexifying intelligence. On the one hand a "vertical" *cognitive* complexification of the individual intelligence, and, on the other hand, a "horizontal" *eusocial* complexification leading to a "swarm intelligence" with two levels micro-macro. Either individual intelligence increases but the social groups remain small communities, and it is the case for evolution leading towards primates and *Homo Sapiens*; or individual intelligence remains very limited but the groups increase drastically and become eusocial, which enables a distributed collective intelligence to emerge.

4.3.5 Mandeville's Fable of the Bees

The "political" character of social insects and its analogy of structure with modern human open societies has a long history and was recurring in the modern period. One of its best known occurrences is *The Fable of the Bees* of Bernard Mandeville (1670-1733) which, according to Hayek, "asked the right question".[2] Mandeville was from a family of liberal progessive physicians of Rotterdam who emigrate in England following conflicts with the Orange Party and Calvinists. His apologue "*The Grumbling Hive: or, Knaves Turn'd Honest*" (1705), later extended and called "*The Fable of the Bees; or, Private Vices, Publick Benefits*" (1714, 1723, 1729) had a major impact (Mandeville 1705 / 1714). At that time, "private vices" meant self-interest, that is what we now call choice rationality and maximization of utility.

The thesis is that a hive functions properly only when each individual bee, each with its very limited representational resources, does what it has to do in the framework of division of labor and follows strict innate rules without bothering about collective advantages or disadvantages. They are the *interactions* of the bees according to efficient selected rules, and not "moral" virtues, which produce the collective wealth, benefits and public goods of the hive. Individual bees don't have the cognitive ressources enabling them to have and share any representation of the hive. In other words, Mandeville's fable on bees as Aristotelian "political animals" concerns the social value of selfish behavior in complex societies where myriads of individuals cooperate through labor division.

Mandeville made explicit in a provocative way the conflict arising at his time between the new born economical liberalism and the traditional Christian ethics.[3] Since the beginning of the XVII-th century, some philosophers already tried to explain that "enlightened" self-love could be socially positive. This political, and in fact *theological*, thesis can previously be found in Blaise Pascal and Pierre Nicole (1625-1695, a Jansenist of Port-Royal, see his *Essais de morale*,

[2] For Mandeville, see Part I, chapter 1.

[3] For a discussion of Mandeville's position within the spectrum of conceptions of human nature in economic theories, see, for example, (Birner 2016).

Nicole 1671): society should be based upon "enlightened" self-love rather than upon charity (see, e.g., Faccarello 2006).

Mandeville inspired Adam Smith and his concept of "invisible hand" as a mechanism ensuring the "Wealth of Nations". This first formulation of self-organized spontaneous order, operates at the metaphysical, theological and political levels as a "ruse of reason". As claimed Pierre Nicole, to make selfish interests *cooperate* in an *un*-intended way to the benefit of public welfare is "the secret plan of God", "the hidden order of God". "There is no need of virtuous individuals to get a virtuous society". Selfish individuals are able, without knowing it and willing it, to do "an admirable thing": the more the persons aim at their own interests, the more they become interdependent, and the more they compose "a superior reality *able to transcend each bet*" (where "transcendence" means here incommensurability and emergence).Thus, God's "ruse of reason" is a "ruse of passion" for humans.

Later, we find the same kind of "ruse of reason" in Kant (1784, *Idea for a Universal History with a Cosmopolitan Purpose*) to solve the political antinomy of the "asocial sociality" (*ungesellige Geselligkeit*), which is the key concept of Kant's anthropology and philosophy of history. The question is no longer to speculate on the original pre-political nature of humans as hostile wolves (Hobbes) or kind lambs (Rousseau). The problem is that human is a social species while his individual nature (ambition, domination, cupidity) is anti-social. But, by a "ruse of reason", Nature uses this Hobbesian pathology to compel humans to accept collective *rules of law*.

Mandeville's originality lies in reversing the problem of the *un*-intended consequences of human actions. In general, this is invoked through the adage that "the road to (social) hell is paved with (individual) good intentions." The provocation consists in reversing the terms and saying that cultural evolution (God's secret plan) has produced mechanisms whereby certain unintended consequences of selfish individual actions can be socially positive. In a way, "(social) heaven is paved with (individual) bad intentions."

The Fable of the Bees triggered a tremendous controversy (in particular with Bishop Berkeley) and was condemned in 1724 for its "diabolic attempts against religion". Even today many people judge outrageous to introduce a principle of *moral inversion* between micro-social individual intentions and global emerging macro-social properties, to posit that intentionally selfish individuals ("rational" in the sense of the theories of rational choice) governed by their private and local self-interest can, by means of their interactions, generate, in an unintended way, a global social order propitious to public interest.

Mandeville was concerned by the economy of large cities and societies, by wealth rather than by virtue. He explained that a "sympathy fusion" between individuals is not necessary to reach a spontaneous "harmony" of interests. This said, it is not yet clear if, in Mandeville, the "invisible-hand" paradox was solved by a "hidden Providence" as it was previously the case in Pierre Nicole, or already by a *natural* self-organizing mechanism as it will be in Adam Smith, or by an *artificial* legislation as it will be in physiocrats as Helvétius or utilitarians as Bentham.

In his *Lecture on a master mind: Dr Bernard Mandeville*, given at the British Academy on March 23-th 1966 (see Hayek 1978), Hayek explains that Mandeville was not a moralist but an excellent physician and psychologist-psychiatrist, and that the question is not to know if the "harmony" is natural (phusis and cosmos) or artificial (nomos and taxis), but to understand that it is an emerging *unintended* spontaneous order. Mandeville already raised the issue of incommensurability of social structures with respect to reason, and of the selection of "good" social rules by cultural evolution.

4.3.6 Eusociality in humans

The evaluation of Mandeville's thought must be contextualized with respect to the different conceptions of social order trying to solve the Kantian "asocial sociality" in modern open societies. Three conceptions of social order have long dominated political philosophy, but none of them has attempted to explain the emerging phenomena of self-organization and spontaneous order in modern societies.

1. The paradigm of hierarchical order and absolute power theorized from the Renaissance by Machiavelli (1469-1527), then Bodin (1529-1596) and Hobbes (1588-1679). It is in reaction to this form of absolutism that many demands arose for tolerance and human rights, from Grotius (1583-1645), Bayle (1647-1706) and Locke (1632-1704) to Kant (1724-1804), Humboldt (1767-1835) and Constant (1767-1830). The "vertical" orders were challenged by more "horizontal" and "democratic" conceptions advocating like Mandeville a sort of eusociality.

2. The rational "constructivist" (in Hayek's sense) paradigm positing that an efficient order can be computed, planned, and applied.

3. The conservative paradigm of natural order, which champions a form of organicist holism and accuses individualism for "atomizing" society and destroying "natural communities" (family, corporations, churches, etc.). For example, for Saint-Simon (1760-1825, *De la physiologie appliquée à l'amélioration des institutions sociales*, 1813) and Auguste Comte (1798-1857, *Système de politique positive*, 1851-1854) holism was a sort of "organicism", a "physiological" conception of the social reality opposing "mechanistic atomism".

However, these three paradigms do not take into account a major evolutionary characteristic of societies that emerged with *urban* civilizations.

Homo Sapiens is a broadminded primate who evolved on the "vertical" axis of cognitive evolution. Our ancestors began to live in families, small groups, clans, hordes, tribes where there existed a cognitive commensurability between individuals and the collectivity. As we have already noted above, a kind of social reflexivity was then possible and there was no eusocial organization. A decentralized non coercive coordination was possible because each member could know the behavior of the others.

But after the prehistoric sedentarization of "hunter-gatherers", the advent of agriculture and farming, with the apparition of great urban civilizations in Mesopotamia, Egypt, Carthage, Rome, China, large cities, new techniques

(writing, arithmetic, geometry), currency, land and sea lines of communication, trade exchanges, etc., a fundamental break occured and an "horizontal" *eusocial-like* complexity emerged, complexity which was alien to the evolutive line of *Homo Sapiens*.

This dimension is well explained by many authors, e.g. Daniel Andler in his *La silhouette de l'humain* (2016).The difficulty lies in articulating cognitive science and social theories from an evolutionary perspective; in understanding the striking *hyper-sociality* of *Homo Sapiens*, with its division of labor, cooperation between agents enabling group performance, "altruism," and other behaviors more closely resembling those of social insects than those of primates. All the classic debates between MI and MH resurface in order to understand the evolutionary dialectic between the cognitive and the social. From a phylogenetic perspective, the problem is very difficult, and hypotheses have been proposed of a "social brain" resulting from a *co-evolution* of mind / sociality leading to a *double* complexity: a cognitive complexity improving on that of primates, coupled with a social complexity similar to that of social insects. *Homo Sapiens* would have become "pro-social" and eusocial. We can cite in particular the work of Richard Byrne, Andrew Whiten, Nicholas Humphrey, and Robin Dunbar, all of which has been further explored by Daniel Andler.

The expression "eusocial-like" uses here the term "eusocial" in a cultural sense and not in a biological sense. When Edward Wilson, the sociobiologist of Harvard specialist of ants (see *The Ants*, written with Bert Hölldobler (Hölldobler-Wilson 1990)) and founding father of *Sociobiology* (1975) wrote *The Social Conquest of Earth: humans are eusocial apes* (Wilson 2012), after having published with Martin Nowak and Corina Tarnita *The evolution of eusociality* (Nowak et al. 2010), he triggered a live debate with sociobiologists specialist of eusocial altruism such as Herbert Gintis and Bill Hamilton, or biologists and ethologists as Richard Dawkins (the author of *The Selfish Gene*, 1976), and also cognitivist psychologists such as Steven Pinker. But, in much the same way as Hayek explained that the moral debate about Mandeville was irrelevant, the biological debate about Wilson is irrelevant. Of course, *Homo Sapiens* is not an eusocial biological species. Human eusociality results from a *cultural* (and not biological) evolution.

The problem is the *mismatch* of this eusocial cultural evolution with our primate brain. Our biological inheritance is not adapted to *global* social coordination. The later, requires the introduction and the acceptation of eusocial impersonal, "objective", and external rules.

4.3.7 Ober's analysis of ancient greek poleis (cities-states)

Before concluding this second part with remarks on the CMI in cooperative game theory, one of us (Jean Petitot) would like to mention a remarkable example of historical and cultural reflection on our civilization that draws profound analogies with what we have discussed so far.

Since the time we read Thucydides' *History of the Peloponnesian War*, we

considered that Hayekian "catallaxy"[4] had been invented by Athenians of the Classical Ancient Greece. In Thucydides' chapter III, the discourse of Corinthians at Sparta[5], is striking. They describe Athenians as "addicted to innovation", "adventurous beyond their power", and "daring beyond their judgment". And insofar as, they said, "it is the law as in art, so in politics, that improvements ever prevail",

> "it happens that the vast experience of Athens has carried her further than you [Sparta] on the path of innovation". (Thucydides 431 BC).[6]

During the academic year 2013-2014, we was invited at Stanford and our office was next to that of Josiah Ober, one of the main specialists of Ancient Greece and of the poleis having emerged from the Dark Ages which followed the fall of the Mycenaean civilization. We had a great deal of discussion on Athens. He was writing a book, *The Rise and Fall of Classical Greece*, which was published later (Ober 2015). Based on the compilation of massive and detailed quantitative data, it is entirely devoted to the thesis that the "efflorescence" of classical Greece, its "miracle", its unique and exceptional wealth, its creation of public goods by a *rational* collective cooperation and not by coercion, are due to the fact that the Greek system of poleis (1035 city-states) constituted an "acentered" distributed network[7] of competitive-cooperative units associated inside leagues and federations swarming in colonies, who succeeded in inventing a political "horizontal" democratic solution to the *decentralized large scale global cooperation enigma*.

This original solution, radically opposed to the Ancient autocratic empires of Mesopotamia, Persia, or Carthage, was invented by Solon and above all,

[4]The spontaneous order generated by the interactions of individual behaviors (e.g., economic) in a collective system (e.g., a market).

[5]The Congress of the Peloponnesian Confederacy at Lacedaemon.

[6]More precisely, Corinthians describe "the great contrast between the two national characters" between Athens and Sparta as follows. "The Athenians are addicted to innovation, and their designs are characterized by swiftness alike in conception and execution. (...) They are adventurous beyond their power, and daring beyond their judgment, and in danger they are sanguine. (...) Further, there is promptitude on their side (...); they are never at home (...): for they hope by their absence to extend their acquisitions (...). They are swift to follow up a success, and slow to recoil from a reverse. Their bodies they spend ungrudgingly in their country's cause; their intellect they jealously husband to be employed in her service. A scheme unexecuted is with them a positive loss, a successful enterprise a comparative failure. The deficiency created by the miscarriage of an undertaking is soon filled up by fresh hopes; for they alone are enabled to call a thing hoped for a thing got, by the speed with which they act upon their resolutions. Thus they toil on in trouble and danger all the days of their life, with little opportunity for enjoying, being ever engaged in getting: their only idea of a holiday is to do what the occasion demands, and to them laborious occupation is less of a misfortune than the peace of a quiet life. To describe their character in a word, one might truly say that they were born into the world to take no rest themselves and to give none to others.

"(...) It is the law as in art, so in politics, that improvements ever prevail; and though fixed usages may be best for undisturbed communities, constant necessities of action must be accompanied by the constant improvement of methods. Thus it happens that the vast experience of Athens has carried her further than you on the path of innovation."

[7]For an introduction to acentered distributed networks see e.g. (Rosenstiehl-Petitot 1974).

against the tyranny of Hippias son of Peisistratos by Clisthene[8], the Alcmeonide grand-uncle of Pericles. Clisthene[9] relied upon the demos to change the political institutions, to reorganize land use planning and to introduce *isonomy* (equality of law). The new rules were based on completely new conceptions of citizenship, civil rights, and just, impersonal and equitable rules of law. They elicited in the new middle classes an avalanche of unique competences and specializations, competing exchanges, technological innovations, investments and risk taking.

To understand how such particular political conditions can *motivate* rational individual agents to cooperate, Ober uses widely works of institutional economy as those by Douglas North (Washington University, Nobel in 1993), Daron Acemoglu (MIT) or James Robinson (Harvard), and game models of cooperative choice as those by Robert Axelrod.

His book is highly original, because, even if the technical term of "eusociality" is not used, the analogy with social insects as political animals in Aristotle's sense is the core of his investigation. Ober refers recurrently to the work of his Stanford colleague Deborah Gordon on ants. Gordon focuses on the sophisticated exchanges of *information* between individuals in ant colonies (see Gordon 2010).[10]

For Ober, Greek "miracle" rests on the creation of "good" political rules enabling the emergence of a collective intelligence coordinating distributed highly specialized knowledge. In Ancient classical Greece, knowledge, technologies, economy, competences were highly specialized, but not politics. The power was shared between a collective self-government of "amateurs". It is the contrary of what Plato defended in *The Republic*.

4.4 Evolutionary game theory

One of the mysteries of eusociality is that of *cooperation*. Unilateral altruism (as opposed to reciprocal altruism) is evolutionarily paradoxical because natural selection is "selfish" and it seems that in situations of basic cooperation, defection is the winning strategy. To understand it, we must think of it as a macro property emerging in complex societies of agents interacting collectively and capable of generating hierarchical *group* effects at several scales. Once again, the CMI paradigm seems to be a good approach. It allows us to formalize the spontaneous self-organized order that is Adam Smith's famous "invisible hand" (Petitot 2006).

This is what shows *evolutionary game theory*, e.g., with the celebrated works of Robert Axelrod (see e.g. Axelrod et al. 1998). In the celebrated Prisonner's Dilemma (see e.g. Poundstone 1993), one starts with the study of Nash equilibria for a simple cooperation game such as the one shot prisoner dilemma

[8] Kleisthenes

[9] After being exiled by Hippias, he came back to Athens in 510 BC and eliminated his rival Isagoras in 508.

[10] It is well known for bees since Karl von Frisch in the 1920s, but it is much more recent for ants. High rates of information exchanges are necessary for eusociality.

between two rational players. Both players can play two moves: cooperation (c) or defection (d). The payoff matrix is such that $T > R > P > S$, with $T = (d, c)$ = Temptation, $S = (c, d) =$ Sucker, $R = (c, c)$ = Reward, $P = (d, d)$ = Punishment. In that case, cooperation is irrational and individual rationality leads to a poor defective (lose, lose) strategy, even if a collective rationality would have selected a good cooperative (win, win) strategy.

But the situation changes drastically when the game becomes more complex. First, one can iterate the game while keeping track of the moves played, which implies that defection can then be punished and cooperation rewarded. In that case, one can introduce more elaborate individual strategies such as unconditional cooperation, unconditional defection, "tit for tat" (start with cooperation, then play what the other player played at the previous move), "vindictive" (start with cooperation and defect for ever as soon as the other player defects, defection being punished as an irreversible betrayal), etc. Moreover games can be *evolutionary* games, where polymorphic populations of individuals use different strategies and define new generations using the scores in a generalized competition: strategies with good scores increase their number of representatives while those with bad scores progressively vanish (see e.g. Samuelson 1997, Weibull 1996). In these models, agents are considered as "phenotypes" expressing "genotypes" identified with strategies, and simple "micro" strategies influence complex "macro" population dynamics.

Evolutionary game theory is more realist than the classical one based on individual choice rationality. It substitutes a collective selective scheme to a variational calculus which would be impossible to perform by real agents. Moreover, it enables to understand the dynamics that drive agents towards global equilibria.

Simulations and computational syntheses prove then that anti-cooperative strategies can be eliminated, and that cooperation can win and become stable. The best strategies are nicely cooperation-supporting, rapidly reacting to defections ("retaliatory"), rapidly forgiving, and simple ("clear", without wiles). The best known is the "tit for tat" (TFT) strategy. But such cooperative strategies are *fragile* with respect to *mutations*. Indeed, as far as, in a TFT environment, unconditionally cooperative mutants exhibits exactly the same cooperative behavior as TFT agents, they can therefore substitute themselves progressively and "silently" for TFT, without any observable effect. But then "bad" (unconditionally defective) mutants can destabilize, invade and destroy the system. So, to be "retaliatory" is a necessary condition for being *collectively stable*.

Moreover, one can, as did Karl Sigmund, Martin Nowak and Robert May, spatialize these evolutionary games by introducing local neighborhood relations between the agents (see Hofbauer-Sigmund 1998, Nowak-May 1993, also Dupuy-Torre 1999). Simulations show that the transition between non-cooperative global states towards emerging cooperative global states has the status of a *critical phenomenon*, exactly as phase transitions in physics.

Let us give Nowak and May's example of systems defined on a square network with 8 neighbors by a payoff matrix with $T = b > R = 1 > P = 0 = S = 0$. So b is the parameter of the system. To define a dynamic, one defines the score of

each agent as the sum of its gain and of the gains of its neighbors and compares the scores. The rule is that each agent adopts the strategy of its neighbor (including itself) that gets the best result.

Let's take an example with a random initial state consisting of half cooperators and half defectors. For b low, e.g. $b = 1.5$, defectors start out winning, but soon the cooperating cells gain the upper hand and the global state c dominates rapidely. For high b, $b > 2$, the global state d dominates rapidely. But for b belonging to a *critical* interval there exists a critical transition $c \rightarrow d$, with multi-scale nested clusters of c and d (Cao-Hwa 1999).

Figure 4.3 displays an example where moves are coded with colors: (c then c) = blue; (d then d) = red; (c then d) = yellow; (d then c) = green. The value of $b = 1.85$ belongs to the critical interval. We see that (d, d) begins to dominate, next that (c, c) begins to reconquer ground by expanding from nuclei having resisted the initial extermination, but that, contrary to the first example $b = 1.5$, multi-scale nested clusters of c and d appear and expand: inside a blue island expanding in a red sea appears an expanding red lagoon, inside which emerges a smaller blue island, etc.

In figure 4.4 this critical dynamics is very apparent on the evolution curves. The curves (c, c) and (d, d) present, besides small oscillating fluctuations, large scale oscillations.

Many social phenomena can be modeled in terms of dynamic evolutionary game models, including ethical issues and the basis of social contracts (see e.g. Binmore 1994).

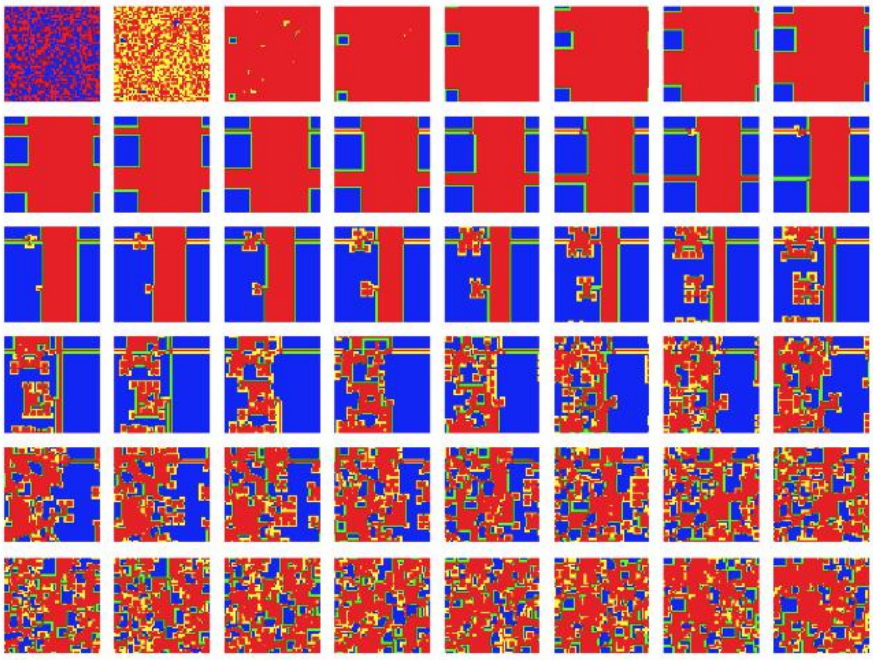


Figure 4.3: An example of Nowak and May's dynamic of cooperation. The pixels represent agents. Moves are coded with colors: c then c = blue; d then d = red; c then d = yellow; d then c = green. The value of the parameter $b = 1.85$ belongs to the critical interval. The initial state is a fifty-fifty random distribution of cooperators and defectors.

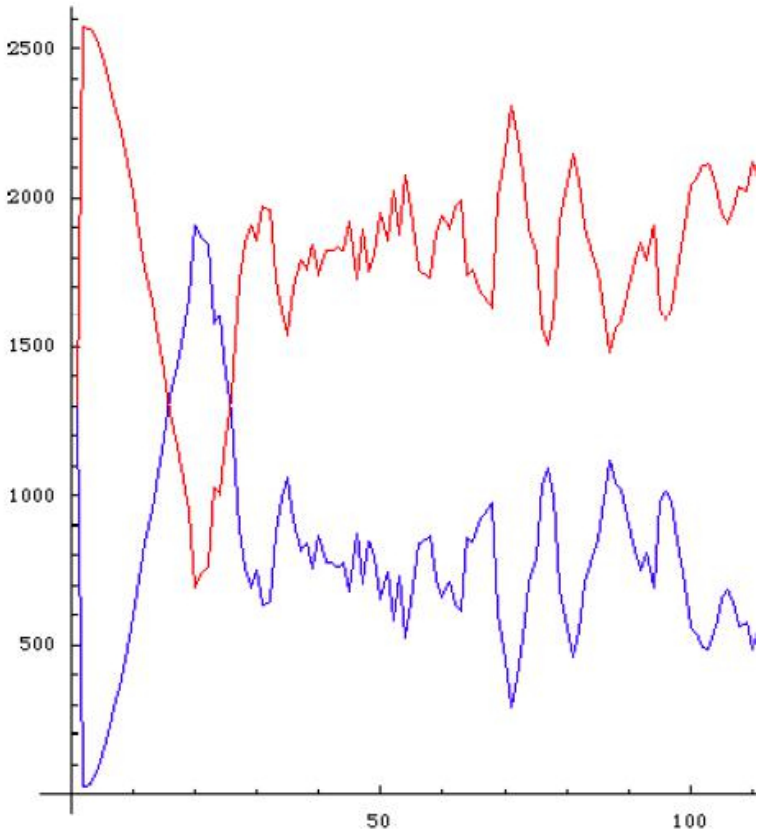


Figure 4.4: The evolution curves of figure 4.3. The curves (c, c) and $(d,\ d)$ present, besides small oscillating flutuations, large scale oscillations.

Chapter 5

Conclusion of Part II

In this second part, we have tried to illustrate, within the framework of complex systems theory, examples of CMI in the natural sciences and formalized social sciences. Moving from a micro level to a macro level, often via a meso level, we have highlighted the fundamentally *computational* nature of the processes involved. These transitions always constitute a major challenge for computational synthesis because the levels in question are conceptually, semantically, and phenomenologically different and are ruled by different laws.

We have traced the path of CMI "from physics to politics."

After a series of epistemological remarks on the recurring debates raised by CMI and after providing a few examples of transdisciplinary models as Pareto's law, we began with physics. And first with the most spectacular example of CMI, that of *thermodynamics*. After reviewing the laws and principles of classical macroscopic thermodynamics, we summarized the extraordinary history of its microscopic (molecular) interpretation in terms of *statistical physics* within the kinetic theory of gases developed by Ludwig Boltzmann and Willard Gibbs. We outlined certain computations because they demonstrate how this shift in level – which seems a priori phenomenologically and ontologically impossible – is nonetheless *mathematically* possible due to the computational *generativity* of mathematics.

This challenge corresponds to the 6th of the 23 problems that Hilbert had identified in 1900 (at the International Congress of Mathematicians in Paris) as crucial mathematical problems for the 20th century. It was not solved until 2025 after a century and a half of intense research carried out by several generations of great mathematicians. The demonstration that it is possible to transition from Newton's reversible mechanical equations for an extraordinarily large number of particles (small hard spheres) interacting through elastic collisions to the irreversible equations of Boltzmann's thermodynamics, and even to those of Euler and Navier-Stokes in hydrodynamics is a miraculous achievement of computational synthesis. We have outlined some elements of this success.

We then presented the basics of statistical physics of *magnetic media* and their *phase transitions*. These are of great interest to us because the elementary

micro-units (atoms) possess an *internal state* (spin), which was not the case for particles in kinetic gas theory. The micro-units can therefore be influenced by the *average field* of the internal states of other units. The concepts of *mean field* and *order parameter* can be applied to a wide variety of situations where micro-units possess internal states.

Critical phenomena of magnetic phase transitions, with their *universal scaling laws*, their critical exponents, and their explanation via the *renormalization group*, represent another highlight of the CMI in physics. They offer a key example of many phenomena encountered in numerous other disciplines. One of the best-known models is the Ising model.

A particularly telling case is that of *spin glasses*, the study of which is of great interest since their equations are very similar to those of *neural networks*. This formal analogy has made it possible to transfer a whole series of formalisms from statistical physics to the modeling of cognitive faculties – a remarkable achievement crowned by the Nobel Prize awarded to John Hopfield and Geoffrey Hinton in 2024. When one considers that this work (typically falling within the scope of the CMI) underlies self-supervised deep learning and generative artificial intelligence, it becomes clear that the CMI is now at the heart of cognitive neuroscience and is completely overturning the traditional dualistic distinction between the natural sciences and the human sciences by unifying them within a naturalistic monism.

After a digression exploring the emergence of purely *geometric* macro-level properties from *physical* micro-level interactions in complex systems such as sand dunes, we continued to explore the gradual transition to the biological, ethological, and social sciences.

To make the transition from neuroscience, we took a moment to consider Friedrich von Hayek's *Sensory Order* dating back to the 1920s. Hayek is a master of CMI in the economic, political, and social sciences. But we must not forget that his theory of *self-organized* complex systems and *spontaneous orders* dates back to the pioneering text *Sensory Order* concerning neuroscience. Hayek was one of the first to understand that there were general laws of self-organization causing macro-structures – in this case psychological – to emerge from a large number of micro-interactions – in this case neurophysiological.

We then moved on to models of "social physics" (econophysics, sociophysics) directly inspired by statistical physics. We have provided a few examples. These models feature scaling laws and critical exponents, demonstrating that there is a certain structural *universality* among CMI models; the epistemology of the CMI is *intrinsically* transdisciplinary.

Ethological disciplines are essential to our discussion because, in these particular complex systems, individual units possess an "interiority," namely that of a living organism endowed with sensorimotor capacities and cognitive faculties (however elementary they may be).

We began with the particularly well-known and spectacular example of flocks of starlings. The individual units (the starlings) fly while remaining grouped together. But they move in a *coordinated* way by applying very simple rules of positioning relative to their neighbors. This is sufficient to give rise, from micro-

local motions, to a macro-collective dynamic possessing a remarkable *global morphological geometry* that varies rapidly. When we examine the computational structure of models like Vicsek's, we find a strong analogy with that of the other models already presented.

The study of *social insects* is equally fascinating. The elementary units follow extremely simple behavioral rules, yet their collective coordination and integration lead to the emergence of highly *architectured*, stable macrostructures such as those found in beehives, termite mounds, or ant colonies. These "architectures without an architect" are the unintended consequences of a myriad of micro-actions. And it is *as such* that they shape the group's social life.

Their formation is led by "stigmergic" processes in the sense of Pierre-Paul Grassé and leads to a sociality endowed with *collective intelligence* named *eusociality* by Suzanne Batra. What is essential is that such macro-structures, which are unintended consequences of micro-actions, absolutely *transcend* the cognitive capacities of individual units.

The analogy between human societies and social insects has often been used to highlight a type of social complexity and collective intelligence that does not seem reducible to a mere evolution of the ethology of the primates that we are biologically. A famous example is that of Bernard Mandeville with his *Fable of the Bees*, well commented on by Hayek and whose argument we have recalled.

In fact, the question concerns the *evolutionary eusociality* of *Homo Sapiens* (in the sense of cultural evolution). As the sociobiologist Edward Wilson stated, "humans are eusocial apes." The question is a difficult one, and, echoing some reflections of Daniel Andler, we have provided some insights into the evolutionary theories of the "social brain," that is, the emergence of a "hypersociality" (coupled with our neurocognitive evolution) in human civilizations.

An interesting example is the description of the emergence of Athenian democracy by Josh Ober, the leading specialist on Thucydides. This is in fact a description of eusociality, with Ober himself referring to the ethological work on ants of his colleagues at Stanford.

We concluded this lengthy discussion with some insights into cooperation models in evolutionary game theory (Robert Axelrod, Karl Sigmund, Martin Novak, Robert May). If we spatialize these theories by introducing neighborhood relations among agents – relations that influence their decision-making – and if we analyze the dynamics of cooperation and defection, we again obtain typical phenomena of complex systems, such as critical phase transitions, but this time with models that closely approximate human social behaviors and a host of political situations.

Thus we concluded this circumnavigation from the physical to the political, which highlighted a wide range of computational syntheses that give full meaning to the CMI.

Bibliography

[1] Amit, D. (1989), *Modeling Brain Function:The World of Attractor Neural Networks*, Cambridge University Press, Cambridge, 1989.

[2] Andler, D. (2016), *La silhouette de l'humain. Quelle place pour le naturalisme dans le monde d'aujourd'hui ?*, Essai NRF, Gallimard, Paris, 2016.

[3] Axelrod R., Cohen M., Riolo, R. (1998), "The Emergence of Social Organization in the Prisoner's Dilemma: How Context Preservation and other Factors Promote Cooperation", *Santa Fe Institute Working Paper*, 99-01-002 (1998).

[4] Bak, P., Tang, C., Wiesenfeld, K. (1987), "Self-organized criticality: An explanation of the $1/f$ noise, *Physical Review Letters*, 59 (1987) 381-384.

[5] Barabási, A.L., Albert, R. (1999), "Emergence of Scaling in Random Networks", *Science*, 286/5439 (1999) 509–512. *arXiv:cond-mat/9910332*.

[6] Barbut, M. (1984), "Note sur quelques indicateurs globaux de l'inégalité : C. Gini, V. Pareto, P. Lévy", *Revue française de sociologie*, XXV, 1984, 602-622.

[7] Bardos, C., Golse, F., Levermore, D. (1991), "Fluid Dynamic Limits of Kinetic Equations I: Formal Derivations", *Journal of Statistical Physics*, 63 (1991) 323–344.

[8] Bernouilli, D. (1738), *Hydrodynamica sive de viribus et motibus fluidorum commentarii*, Johann Reinhold Dulseker, Argentorati, 1738. *Engl transl.* by T. Carmody and H. Kobus, Dover, 1968.

[9] Binmore, K. (1994), *Playing Fair*, Cambridge, MIT Press, 1994.

[10] Birner, J. (2016), "Metaphysical Models of Man in Economics", *Cosmos and Taxis*, 3/2-3 (2016) 78-93.

[11] Bodineau, T., Gallagher, I., Saint-Raymond, L. (2014), "Limite de diffusion linéaire pour un système déterministe de sphères dures", *C. R. Académie des Sciences. Paris*, Ser. I, 352 (2014) 411–419.

[12] Bodineau, T., Gallagher, I., Saint-Raymond, L., Simonella, S. (2023), "Long-time correlations for a hard sphere gas at equilibrium", *Communications on Pure and Applied Mathematics*, 76 (2023), 12, 3852–3911.

[13] Boltzmann, L. (1872), "Weitere Studien über das Wärmegleichgenicht unter Gasmolekülen", *Sitzungsberichte der Akademie der Wissenschaften*, 66 (1872), 275-370. Translation : "Further studies on the thermal equilibrium of gas molecules", *Kinetic Theory 2*, 88-174, (S.G. Brush, ed.), Pergamon Press, Oxford,1966.

[14] Bonabeau, E., Theraulaz, G., Deneubourg, J.-L., et al. (1998), "A model for the emergence of pillars, walls and royal chambers in termite nests", *Philosophical Transactions of the Royal Society of London*, *B*, 353 (1998) 1561-1576.

[15] Bouchaud, J.-P. (2023), "From statistical physics to social sciences: the pitfalls of multi-disciplinarity", *Journal of Physics: Complexity*, 4/4 (2023) 041001.

[16] Bouchaud, J.-P.and Mézard, M. (2000), "Wealth condensation in a simple model of economy". *Physica A: Statistical Mechanics and its Applications*, 282 (2000) 536–545.

[17] Boudon R. (1986), *Theories of Social Change: A Critical Appraisal*, Polity Press, Cambridge,1986.

[18] Bronner, G. (2023), "Why Does Comprehensive Sociology Need to Be Enlarged?", *The Palgrave Handbook of Methodological Individualism I*, Bulle, N., Di Iorio, F. (eds), Palgrave Macmillan, 2023, 271-290.

[19] Bulle, N. (2023), "Methodological Individualism as Holism of the Parts: From Epistemology to Ontology", *The Palgrave Handbook of Methodological Individualism I*, Bulle, N., Di Iorio, F. (eds), Palgrave Macmillan, 2023, 293-319.

[20] Bulle, N., Phan, D. (2017), "Can analytical sociology do without methodological individualism", *Philosophy of the Social Sciences*, 47/6 (2017) 379-409.

[21] Bulle, N., Di Iorio, F. (eds) (2023), *The Palgrave Handbook of Methodological Individualism I, II*, Palgrave Macmillan, 2023.

[22] Cao, Z., Hwa, R. (1999)., "Phase transition in evolutionary games", *International Journal of Modern Physics A*, 14/10 (1999) 1551-1559.

[23] Chaté, H., Ginelli, F., Grégoire, G., Raynaud, F. (2008), "Collective motion of self-propelled particles interacting without cohesion", *Physical Review*, *E* 77, 2008, 046113.

[24] Chavalarias, D.(2020), "From inert matter to the global society - Life as multi-level networks of processes", *Philosophical Transactions of the Royal Society B*, 375/1796, 2020.

[25] Chavalarias,D. (2022), *Toxic Data. Comment les réseaux manipulent nos opinions*, Flammarion, Paris, 2022.

[26] Clapeyron, E. (1834), "Mémoire sur la puissance motrice de la chaleur", *Journal de l'École Polytechnique*, 23 (1834) 153-191.

[27] Coleman J. S. (1987), "Microfoundations and macrosocial behavior", *The Micro-Macro Link* (J. Alexander, B. Giesen, R. Münch, N. Smelser, eds), University of California Press, Berkeley/Los Angeles, 1987, 153-177.

[28] Coleman, J. S. (1990). *Foundations of Social Theory*, Cambridge, MA and London: Belknap Press of Harvard University Press.

[29] Comte, A. (1851-1854), *Système de politique positive*, *I*, *II*, Hermann, Paris 2022.

[30] Deng, Y., Hani, Z., Ma, X. (2025a), "Hilbert's sixth problem: Derivation of fluid equations via Boltzmann's kinetic theory". *arXiv:2503.01800*, 2025.

[31] Deng, Y., Hani, Z., Ma, X. (2025b), "Long time derivation of the Boltzmann equation from hard sphere dynamics". *arXiv:2408.07818*, 2025.

[32] Di Iorio, F. (2015), *Cognitive Autonomy and Methodological Individualism: The Interpretative Foundations of Social Life*, Springer, 2015.

[33] Di Iorio, F. (2016), "Introduction: Methodological Individualism, Structural Constraints, and Social Complexity", *Studies in Emergent Order and Organization*, *Cosmos and Taxis*, 3/2-3 (2016) 1-8.

[34] Di Iorio, F., Chen S.-H. (2019), "On the connection between agent-based simulation and methodological individualism", *Social Science Information*, 58/2 (2019,) 354-376.

[35] Di Nuoscio, E. (2023), "Methodological Individualism, Scientific Explanation, and Hermeneutics", *The Palgrave Handbook of Methodological Individualism I*, Bulle, N., Di Iorio, F. (eds), Palgrave Macmillan, 2023, 181-202.

[36] Douady, S., Hersen, P. (2011), "Dunes, the Collective Behavior of Wind and Sand", *Morphogenesis*, (P. Bourgine, A. Lesne(eds.), Springer, Berlin-Heidelberg, 2011, 107-118.

[37] Duhem, P. (1908), *Sauver les apparences. (*Sôzein ta phainomena). *Essai sur la notion de théorie physique de Platon à Galilée*, Vrin, Paris, 2019

[38] Dupuy, C., Torre, A. (1999), "The morphogenesis of spatialized cooperation relations", *European Journal of Economic and Social Systems*, 13, 1, (1999), 59-70.

[39] Dupuy, J-P. (2000), *The Mechanization of Mind. On the Origins of Cognitive Science*, Princeton University Press, 2000.

[40] Dupuy, J-P. (2004), "Vers l'unité des sciences sociales autour de l'individualisme méthodologique complexe", *Revue du MAUSS*, 24/2 (2004) 310-328.

[41] Dupuy, J.-P., Dumouchel, P., (dir) (1981), *L'auto-organisation : de la physique au politique*, Colloque du Centre Culturel International de Cerisy-la-Salle, 1981. Éditions du Seuil, Paris, 1983. Réédition, Hermann, Paris, 2022.

[42] Elster, J. (2023), "What's the Alternative?", *The Palgrave Handbook of Methodological Individualism I*, Bulle, N., Di Iorio, F. (eds), Palgrave Macmillan, 2023, 117-138.

[43] Epstein, J. (2006), *Generative Social Science. Studies in Agent-Based Computational Modeling*, Princeton University Press, 2006.

[44] Euler, L. (1744), "Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes, sive solutio problematis isoperimetrici lattissimo sensu accepti", *Opera Omnia* : Series 1 (Opera mathematica), Volume 24, Swiss Academy of Sciences. Trad. All. A. Aycock et A. Diener. Reference E65 on the Euler Archive site.

[45] Faccarello, G. (2006), "La liberté du commerce et la naissance de l'idée de marché comme lien social", *Histoire du libéralisme en Europe*, (P. Nemo, J. Petitot eds), Presses Universitaires de France, Paris, 205-253, 2006.

[46] Ferguson, A. (1996), *An Essay on the History of Civil Society*, 1767. Cambridge University Press, 1996.

[47] Galam, S. (2008), "Sociophysics: A review of Galam models", *arXiv:0803.1800*, 2008.

[48] Galam, S., Gefen, Y., Shapir, Y. (1982), "Sociophysics: a New Approach of Sociological Collective Behaviour. I. Mean-Behaviour Description of a Strike", *Journal of Mathematical Sociology*, 9 (1982) 1-13.

[49] Galam, S., Moscovici, M. (1991), "Towards a theory of collective phenomena: Consensus and attitude changes in groups", *European Journal of Social Psychology*, 21 (1991) 49-74.

[50] Gallagher, I., Saint-Raymond, L., Texier, B. (2013), "From Newton to Boltzmann: Hard Spheres and Short-Range Potentials", *arXiv:1208.5753*, 2013.

[51] Gallagher, I. (with T. Bodineau, L. Saint-Raymond and S. Simonella) (2026), "Derivation of the Boltzmann equation from hard-sphere dynamics [after Y. Deng, Z. Hani, and X. Ma]", *Séminaire Bourbaki 1247*, 2026.

[52] Gibbs, J.W. (1902), *Elementary principles in statistical mechanics*, Charles Scribner's sons, New York and Edward Arnold, London, 1902.

[53] Ginelli, F. (2016), "The Physics of the Vicsek Model", *arXiv/:1511.01451*, 2016.

[54] Gordon, D.M.(2010), *Ant Encounters. Interaction Networks and Colony Behavior*, Princeton University Press, 2010.

[55] Grad, H. (1958), "Principles of the kinetic theory of gases", *Handbuch der Physik*, Springer-Verlag, Berlin-Heidelberg, 205-294, 1958.

[56] Grassé, P.P., *Termitologia* (1982-1984), 3 volumes, Masson, Paris, 1982, 1983, 1984.

[57] Greve, J. (2023), "Collective Intentionality and Methodological Individualism", *The Palgrave Handbook of Methodological Individualism II*, Bulle, N., Di Iorio, F. (eds), Palgrave Macmillan, 2023, 3-27.

[58] Hamilton, W. (1964), "The genetical evolution of social behaviour. I&II", *Journal of Theoretical Biology*, 7/1 (1964) 1–16 & 17–52.

[59] Hayek, F. (1952), *The Sensory Order*, London, Routledge & Kegan Paul, 1952.

[60] Hayek, F. (1978), *New Studies in Philosophy, Politics, Economics and the History of Ideas*, University of Chicago Press, 1978.

[61] Hayek, F. (1982), "The Sensory Order After 25 Years", *Cognition and the symbolic processes*, (W.B. Weimer, D.S.. Palermo eds), Vol. 2, 1982, 287–293, Hillsdale, NJ, Erlbaum.

[62] Heider, F., Simmel, M. (1944), "An experimental study of apparent behavior", *American Journal of Psychology*, 57 (1944) 243-259.

[63] Hilbert, D. (1901), "Mathematical Problems". *Bulletin of the American Mathematical Society*, 8 (1901), 437–479.

[64] Hodgson, G. (1986),"Behind methodological individualism", *Cambridge Journal of Economics*, 10 (1986) 211-224.

[65] Hofbauer, J., Sigmund, K. (1998), *Evolutionary games and population dynamics*, Cambridge University Press, 1998.

[66] Hölldobler, B., Wilson, E. (1990), *The Ants*, Harvard University Press,1990.

[67] Hopfield,J. 1982 "Neural networks and physical systems with emergent collective computational abilities", *Proceedings of the National Academy of Sciences*, 79/8 (1982) 2554-2558.

[68] Ising, E. (1925), "Beitrag zur Theorie des Ferromagnetismus", *Zeitschrift für Physik*, 31(1), 253–258, 1925.

[69] Jaynes, E.T. (1957), "Information Theory and Statistical Mechanics", *Physical Review*, 106/4 (1957) 620-630.

[70] Kant, I. (1781), *Kritik der reinen Vernunft*. Engl. trans and ed. *Critique of Pure Reason*, Paul Guyer and Allen Wood, Cambridge University Press, 1997.

[71] Kant, I., (1784), *Idee zu einer allgemeinen Geschichte in weltbürgerlicher Absicht*. Engl. Transl. *Idea for a Universal History with a Cosmopolitan Purpose*, Cambridge University Press, 1991.

[72] Kant, I., (1786), *Metaphysische Anfangsgründe der Naturwissenschaft*, Kants gesammelte Schriften, Band IV, Preussische Akademie der Wissenschaften, Georg Reimer, Berlin, 1911. Engl. transl. Metaphysical Foundations of Natural Science, (M. Friedman, ed.), Cambridge University Press, 2004.

[73] Landau, L., Lifchitz, E., (1958). *Statistical Physics*, Pergamon Press, London-Paris, 1958.

[74] Lanford III, O.E. (1975), *Time evolution of large classical systems*, Lectures Notes in Physics 38, (J. Moser, ed.), 1–111, Springer Verlag, 1975.

[75] Lanford III, O.E. (1976), "On a derivation of the Boltzmann equation", *Astérisque*, 40 (1976) 117-137.

[76] Laurent, A. (1994), *L'individualisme méthodologique*, Collection Que sais-je ? vol. 2906, Presses Universitaires de France, Paris, 1994.

[77] Le Cun, Y., Yoshua Bengio, Y., Hinton, G. (2015), "Deep learning", *Nature*, 521/7553, (2015) 436–444.

[78] Mandeville, B. (1705/1714), *The Fable of the Bees: Or, Private Vices, Publick Benefits* (1705/1714), F. B. Kaye ed., Oxford, Clarendon, 1957.

[79] Manzo, G. (2023), "Agent-Based Models and Methodological Individualism: Are They Fundamentally Linked?", *The Palgrave Handbook of Methodological Individualism I*, Bulle, N., Di Iorio, F. (eds), Palgrave Macmillan, 2023, 681-710.

[80] Nemo, P. (2002), *Histoire des idées politiques aux temps modernes et contemporains*, Presses Universitaires de France, 2002.

[81] Nemo, P., Petitot, J., (eds.) (2006), *Histoire du libéralisme en Europe*, Presses Universitaires de France, Paris, 2006. Italian translation by F. Di Iorio, Rubbettino, Soveria Mannelli, 2012.

[82] Nicole, P. (1671), *Essais de Morale, contenus en divers traités sur plusieurs devoirs importans*,1671. Paris, Presses Universitaires de France, 1999. Paris, Les Belles Lettres, 2016.

[83] Nowak, M., May, R. M. (1993), "The Spatial Dilemmas of Evolution", *International Journal of Bifurcations and Chaos*, 3/1 (1993) 35-78.

[84] Nowak, M., Tarnita, C., Wilson, E. (2010) , "The evolution of eusociality", *Nature*, 466 (2010) 1057–1062.

[85] Ober, J. (2015), *The Rise and Fall of Classical Greece*, Princeton University Press, 2015.

[86] Onsager, L. (1944), "Crystal Statistics. I. A Two-Dimensional Model with an Order-Disorder Transition", *Physical Review*, 65, 117, 1944.

[87] Opp, K. D. (2023), "Methodological Individualism and Micro–Macro Modeling", *The Palgrave Handbook of Methodological Individualism I*, Bulle, N., Di Iorio, F. (eds), Palgrave Macmillan, 2023, 377-405.

[88] Pareto, V., (1885), "La legge della domanda", *Giornale degli Economisti*, 1885, 59-68.

[89] Pentland, A., (2014), *Social Physics*, Penguin Press, 2014.

[90] Petitot, J. (1982), *Introduction aux phénomènes critiques*, 1982. http://jeanpetitot.com/ArticlesPDF/Petitot_CritPh.pdf

[91] Petitot, J. (1992), "Actuality of Transcendental Aesthetics for Modern Physics", *1830-1930 : A Century of Geometry*, (L. Boi, D. Flament, J.-M. Salanskis eds), Berlin, New-York, Springer.

[92] Petitot, J. (1993). "The unity of mathematics as a discovery method: Wiles' example", *XIXth International Congress of History of Science*, Zaragoza, August 1993 & *7th French Philosophy of Mathematics Workshop*, Paris, November 2015.

[93] Petitot, J. (1999), "Le troisième labyrinthe: dynamique des formes et architectonique disjonctive", *L'actualité de Leibniz : les deux labyrinthes* (D. Berlioz, F. Nef eds), Studia Leibnitiana Supplementa, 34, 617-632, Stuttgart, Franz Steiner, 1999.

[94] Petitot, J. (2006), "Modèles formels de la 'main invisible' : de Hayek à la théorie des jeux évolutionniste", *Histoire du libéralisme en Europe*, (Ph. Nemo, J. Petitot eds), Presses Universitaires de France, Paris, 1095-1114, 2006.

[95] Petitot, J. (2008), *Neurogéométrie de la vision. Modèles mathématiques et physiques des architectures fonctionnelles*, Les Éditions de l'Ecole Polytechnique, Paris, 2008. http://jeanpetitot.com/ArticlesPDF/Petitot_NGV_2008.pdf.

[96] Petitot, J. (2010), "Reduction and Emergence in Complex Systems", *Questioning Nineteenth-Century Assumptions about Knowledge, II Reductionism*, (Richard E. Lee, ed.), SUNY Series, NY, 2010.

[97] Petitot, J. (2016), "Complex Methodological Individualism", *Cosmos and Taxis*, 3/2-3 (2016) 27-37.

[98] Petitot J. (2017), *Elements of Neurogeometry. Functional Architectures of Vision*, Springer, 2017.

[99] Petitot, J. (2022). "Axiomatics as a functional strategy for complex proofs. The case of Riemann Hypothesis", *Axiomatic Thinking II*, (F. Ferreira, R. Kahle, G. Sommaruga, eds.), Springer International Publishing, Chapter 8, 165-194, 2022.

[100] Petitot, J. (2023), "Complex Methodological Individualism", *The Palgrave Handbook of Methodological Individualism I*, Bulle, N., Di Iorio, F. (eds), Palgrave Macmillan, 2023, 407-431.

[101] Petitot, J. (2023b), *Schèmes et modèles : la modélisation comme synthèse computationnelle*, https://jeanpetitot.com/ArticlesPDF/Petitot_Modelisation_2023.pdf

[102] Petitot, J. (2024) *Éléments de Neurogéometrie II. Géométrie sous-riemannienne et analyse harmonique non commutative du cortex visuel primaire*, https://hal.science/hal-04722501.

[103] Poincaré, H. (1908), *Thermodynamique*, Gauthiers-Villars, Paris, 1908.

[104] Poundstone, W. (1993), *Prisoner's Dilemma*, Oxford University Press, 1993.

[105] Ramström, G. (2023), "The Analytical Micro–Macro Relationship in Social Science and Its Implications for the Individualism-Holism Debate", *The Palgrave Handbook of Methodological Individualism I*, Bulle, N., Di Iorio, F. (eds), Palgrave Macmillan, 2023, 485-510.

[106] Raub, W. and van de Rijt, A. (2023), "Methodological Individualism and Formal Models", *The Palgrave Handbook of Methodological Individualism II*, Bulle, N., Di Iorio, F. (eds), Palgrave Macmillan, 2023, 373-399.

[107] Robinet, A. (1986), *Architectonique disjonctive, Automates systémiques et Idéalité transcendantale dans l'œuvre de G. W. Leibniz*, Paris, Vrin, 1986.

[108] Rosenstiehl, P., Petitot, J. (1974), "Automate asocial et systèmes acentrés", *Communications*, 22 (1974) 45-62.

[109] Rumelhart D.E., Mc Clelland, J.L. (1986), *Parallel Distributed Processing: Exploration in the MicroStructure of Cognition*, Cambridge, MIT Press, 1986.

[110] Saint-Raymond, L. (2009), *Hydrodynamic limits of the Boltzmann equation*, Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2009.

[111] Saint-Simon, C.-H. (2013), *Œuvres complètes*, Presses Universitaires de France, Paris, 2013.

[112] Samuelson, L. (1997), *Evolutionnary Games and Equilibrium Selection*, Cambridge, MIT Press, 1997.

[113] Sawyer, R. K. (2002), "Nonreductive individualism., I, II", *Philosophy of the Social Sciences*, 32/4 (2002) 537–559 and 33/2 (2003) 203-224.

[114] Searle, J.R. (1995), *The Construction of Social Reality*, Penguin, Londres, 1995.

[115] Shannon, C.S. (1948), *A Mathematical Theory of Communication*, Bell System Technical Journal, 27 (1948) 379-423 and 623-656.

[116] Thom, R. (1972), "Phase transitions as catastrophes", *Statistical Mechanics: new concepts, new problems, new applications* (S.A. Rice, K.F. Freed and C.C. Light, eds), Univ. of Chicago Press, 1972, 93–107. *Œuvres Mathématiques*, vol. III, Société Mathématique de France, Paris, 2021, 237–270.

[117] Thom, R. (1972b), *Stabilité structurelle et Morphogenèse*, Benjamins, Amsterdam, Ediscience, Paris, 1972. Engl transl. *Structural Stability And Morphogenesis*, (trans by D. Fowler) Benjamins, Reading, MA, 1975.

[118] Thucydides (431 BC), *The History of the Peloponnesian War*, translated by Richard Crawley, Loeb Classical Library, 2018. *Internet Classics Archive*, classics.mit.edu//Thucydides/pelopwar.html

[119] Toulouse, G., Pfeuty, P. (1975), *Introduction au groupe de renormalisation et à ses applications*, Presses Universitaires de Grenoble, Grenoble, 1975.

[120] Udehn, L. (2001), *Methodological Individualism: Background, History and Meaning*, Routledge, London, 2001.

[121] Uffink, J., Valente, G. (2010), *Time's Arrow and Lanford's Theorem*, Séminaire Poincaré (Institut Henri Poincaré, Paris), XV, 2010, 141-173.

[122] Van der Waals, J.D. (1873), *Over de continuïteit van den gas en vloeistoftoestand*, Sijthoff, Leiden, 1873. Engl. transl. *On the Continuity of the Gaseous and Liquid States*, North-Holland, 1988.

[123] Vicsek, T., Czirók, A., 1, Ben-Jacob, E., Cohen, I., Shochet, O. (1995), "Novel Type of Phase Transition in a System of Self-Driven Particles", *Physical Review Letters*, 75/6 (1995) 1226-1229.

[124] Villani, C. (2002), "A review of mathematical topics in collisional kinetic theory", *Handbook of Mathematical Fluid Dynamics I*, (S. Friedlander, D. Serre,eds), Elsevier Science, 2002.

[125] Watkins, J.W.N. (1952), "The Principle of Methodological Individualism", *The British Journal for the Philosophy of Science*, 3/10 (1952) 186–189.

[126] Weibull, J. (1996), *Evolutionary Game Theory*, Cambridge, MIT Press, 1996.

[127] Widom, B. (1965), "Surface Tension and Molecular Correlations near the Critical Point", *Journal of Chemical Physics*, 43 (1965) 3892-3897.

[128] Wilson, E. (2012), *The Social Conquest of Earth: humans are eusocial apes*, 2012.

[129] Wilson, K., Kogut, I. (1974), *The Renormalization Group and the ε expansion*, Physics Reports, 12/2 (1974) 75-199.

[130] Zakine, R., Garnier-Brun, J., Becharat, A.-C., Benzaquen, M. (2023), "Socioeconomic agents as active matter in nonequilibrium Sakoda-Schelling models", *arXiv:2307.14270*, 2023.

[131] Zeeman, C. (1977), *Catastrophe Theory.* Addison-Wesley, Reading, MA, 1977.