Springer Proceedings in Mathematics & Statistics

Sergio Albeverio
Elisa Mastrogiacomo
Emanuela Rosazza Gianin
Stefania Ugolini *Editors*

Complexity and Emergence

Lake Como School of Advanced Studies, Italy, July 22–27, 2018

Springer

Springer Proceedings in Mathematics & Statistics

Volume 383

This book series features volumes composed of selected contributions from workshops and conferences in all areas of current research in mathematics and statistics, including data science, operations research and optimization. In addition to an overall evaluation of the interest, scientific quality, and timeliness of each proposal at the hands of the publisher, individual contributions are all refereed to the high quality standards of leading journals in the field. Thus, this series provides the research community with well-edited, authoritative reports on developments in the most exciting areas of mathematical and statistical research today.

More information about this series at https://link.springer.com/bookseries/10533

Sergio Albeverio · Elisa Mastrogiacomo ·
Emanuela Rosazza Gianin · Stefania Ugolini
Editors

Complexity and Emergence

Lake Como School of Advanced Studies,
Italy, July 22–27, 2018

Springer

Editors
Sergio Albeverio
University of Bonn and Hausdorff Center for Mathematics
Bonn, Germany

Elisa Mastrogiacomo
Departments of Economics
Insubria University
Varese, Italy

Emanuela Rosazza Gianin
Department of Statistics and Quantitative Methods
University of Milano-Bicocca
Milano, Italy

Stefania Ugolini
Department of Mathematics
University of Milan
Milano, Italy

ISSN 2194-1009 ISSN 2194-1017 (electronic)
Springer Proceedings in Mathematics & Statistics
ISBN 978-3-030-95702-5 ISBN 978-3-030-95703-2 (eBook)
https://doi.org/10.1007/978-3-030-95703-2

Mathematics Subject Classification: 37Fxx, 97M10, 91G80, 91Bxx, 00A30, 97M50, 92C42, 94A15, 97A17, 00A67

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface to the Volume Advanced School on Complexity and Emergence: Ideas, Methods, with a Special Attention to Mathematics and Its Application to Economics and Finance

Complex systems involve a large number of components and interactions and cannot practically be reduced to a small number of simple components with well-defined relations. The theory of complex systems has evolved from various areas of research, in particular, nonlinear dynamical systems, general system theory, chaotic system theory and the theory of stochastic processes. It is also strongly influenced by network theory, statistical mechanics, computer science (algorithmic complexity), statistical analysis (extreme events, data analysis) and control theory.

Emergent phenomena, presenting essential qualitative differences from those of the components, deserve special interest, and their study (explanation of emergence, prediction) constitutes a major goal in several areas of knowledge.

In the study of complex systems and associated simulations, models inspired by theoretical and mathematical physics play an important role, since they often provide well-defined candidates for the dynamics of the components (microlevel) from which emergent behaviour of the system as a whole (macrolevel) should be explained. In the last decade in areas like meteorology or even climate analysis large-scale models have been constructed on the basis of complex systems theory and are the basis of present-day predictions, respectively, interpretation of past developments, since they can be tested and improved by steadily increasing data sets. The construction of similar models has also been attempted with some success in certain biological contexts (*e.g.* in ecological settings or in the study of some brain functions).

In other areas, and, in particular, in socio-economical sciences, the interplay between microlevel and macrolevel, and the phenomena of emergence arising from this interplay are much less understood. In recent years, however, agent-based models have been developed in the area of econophysics; these have the potential to relate to models also discussed by economists, especially in the study of financial markets. In fact, a financial market generally consists of the traders, furnished by varying amounts of capital and the interaction rules of the trading platform. In contrast to such a simple microscopic structure, the macroscopic behaviour of financial markets appears to be surprisingly rich: the seemingly uncorrelated swings of financial indices and the extreme event of a crash constitute typical phenomena present in systems. This behaviour of the market resulting from the interactions of the investors is an example

of emergence of macrophenomena from an underlying microlevel of interactions. Interdisciplinary efforts to better understand why the evolution of macroeconomical systems can run into "singularities" (*e.g.* financial crisis) are increasingly required, and methods of the area of complex systems (phase transitions, asymptotics, extreme events, stochastic control) are expected to play an important role.

The need to prepare a generation of researchers able to do critical, interdisciplinary work and possibly develop new synergies between different communities, from mathematics to information sciences and theoretical physics, on one hand, and to socio-economical sciences, on the other hand, is much felt. Because of this an "Advanced School on Complexity and Emergence: ideas, methods, with special attention to mathematics and its application to economics and finance" was organized at the international Institute Lake Como School of Advanced Studies (of the Universities of Insubria, Milano and Milano-Bicocca and Pavia), July 22–27, 2018. The present volume gathers contributions that resulted from lectures present at the School. All lectures were given by leading specialists, active in different areas of mathematics, physics, biology as well as philosophy of science, econometrics, jurisprudence and architecture, building bridges among the various communities involved and working jointly on developing new paths in the interdisciplinary subject of complexity and emergence. We believe that the volume will be of interest to those working in mathematics, information sciences, theoretical physics and economics, as well as research workers in those areas, who want to enlarge their spectrum of knowledge towards the world of complex systems and emerging effects.

The contribution by Nihat Ay, Nils Bertschinger, Jürgen Jost, Eckehard Olbrich, Johannes Rauh entitled "Information and Complexity, or: Where is the Information?" presents and discusses basic principles of information theory and complexity, in particular, regarding their quantification, how they are related and how the complexity notions can be quantified in qualified information theoretical terms. They stress the need to discuss the question of "what is information" next to the one "information about what?". They provide a very good survey on what is presently known about these questions. The presentation is greatly enriched by numerous simple examples. The structure of the lectures is as follows. Section 2 presents basic principles of information theory, discussing Shannon information (or entropy) associated with random variables (for simplicity they assume that they have a finite number of values); then they go over to discuss measures for mutual and conditional information, followed by a discussion of maximum entropy. The last topic of this section handles geometric information concepts, in particular, the Kullback-Leibler divergence. Section 3 discusses very clearly different notions of complexity (like algorithmic and, respectively, computational complexity), as well as entropy and its rate; minimal sufficient statistics, stochastic complexity, excess entropy, forecasting complexity and statistical complexity; moreover, Aristotle's definition as "the whole being more than the union of its parts" is analysed. The difference between external and internal complexity is stressed, based on results of the authors themselves, and optimization principles, relating to autopoiesis and sociological approaches by N. Luhmann as well as to concepts of psychology, cognitive sciences, molecular biology,

statistics and statistical learning theory. Hierarchical, complementarity and interaction are also analysed and illustrated using a simple dynamical system. The last section discusses the concepts related to information decomposition.

The contribution by Luciano Boi, entitled "A topological and dynamical approach to the study of complex living systems" focuses on developments in the study of biological phenomena that have emerged from new ideas and methods in mathematics, notably differential geometry, topology and the theory of dynamical systems. The modelling of many processes that occur within living cells and organisms has indeed received a great momentum by such developments. In particular, ideas and methods of mathematics have been applied to discuss: (i) relationships between plasticity of biological structures and emergence of their functionality; (ii) the relation between form and function (in Aristotelian sense) and (iii) the connection between robustness and fragility in complex living organisms. On the way in his rich exposition the author discusses (in Section 1) topics like complex structures in physics (including topological stability, phase transitions, hierarchical approaches, winding mechanism). He stresses especially the presence of geometrical and topological structures in nature (like knotted filamentary structures). For the study of them mathematical methods and ideas naturally combined with insights from physics. So, *e.g.* in the study of stability in condensed matter and its relations with phase transitions in statistical mechanics, still in the introduction and also elsewhere in the article, the author puts in evidence epistemological aspects, in particular, the limits of a hierarchical paradigm of science, going through a rich collection of examples to illustrate his point of view (including, *e.g.* superconductors and DNA structures). In particular, in relation to DNA molecules, he discusses in detail topological structures, symmetry-breaking and linking numbers. Emergence in physical patterns as related to topological changes in the system (that are illustrated using algebraic and differential geometric methods, clearly presented). Complexity is illustrated also in relation with symmetry-breaking phenomena, *e.g.* in morphogenesis and molecular biology, and the special character acquired by the concepts of complexity, emergence and causation in biological contexts. He also discusses self-organization and other systemic properties in biology and measures of chaos, fractability and complexity. The genome is viewed as a complex system that cannot be understood by looking at DNA sequences alone. More generally, the need for an approach in terms of systems biology is stressed. The last sections hint at present development involving complex topology (relations between supercoiling, topological plasticity and biological functions). The author strongly emphasizes the multi-levelness of complex systems and urges a systemic approach, taking fully into account also recent developments in the areas of topology and geometry. The strong point of view that "biological information of living organisms come to be portrayed in the DNA sequence above" (as formulated in Section 15) is illustrated in a very original way, with particular attention to functional aspects of plasticity and supercoiling. Interesting open problems are pointed out in the conclusive remarks.

The contribution by Ilaria Capelli takes us in the world of jurisprudence looked upon as handling of a complex system. In fact, a very innovative point of view is presented through a description of the different levels of complexity in jurisprudence,

linking financial systems and their regulations and "models related to fraud". The contribution starts with a survey of characteristics of complex systems (including nonlinearity, information and entropy, interconnections between agents and environment). Then it goes over to describe complexity aspects and matters of jurisprudence (distinguishing between complex systems and complicated ones). Aspects of emergence and adaptation are stressed, as well as economic and social interrelations, and their interwovenness with various other aspects of jurisprudence. In Section 3, the relation between complexity theory and legal thinking is presented, in particular, connections with N. Luhmann's theory of asymptotic systems is pointed out. In Section 4, the relation between complexity theory and aspect of economics, in particular finance, is discussed, also in their relation to ecological and social problems. In the last three sections, the author deals with complexity jurisprudence in relation with financial systems regulation, including systemic risk and the role of regulation in preventing crisis (in face of systemic risk and fraud). It ends with a most welcome plea for a multidisciplinary approach concerning regulation of financial markets.

The contribution by Francesco Guerra is entitled "The emergence of the order parameter in the interpolating replica trick for disordered statistical mechanical systems" and presents in a condensed but clear and very competent way the results of many decades of investigations of important models of cooperative behaviour emerging from disorder complex systems. Such models originated from the classical Ising-Lenz model of ferromagnetism, a paradigmatic model (originally handling a discrete valued coupled "spin variables" introduced in the 20s to describe ferromagnetic behaviour), but whose applications extend far beyond physics, into areas like economics, sociology, geography, urbanism, biology and medicine. The models of spin glasses and the random emergent model constitute the main topic of investigation in F. Guerra's essay. They can be looked upon as an extension of the abovementioned Ising-Lenz-type models to the case where the coupling between the spin variables is taken to be random. The spin glass model was introduced by Sherrington and Kirkpatrick in the 70s (as a meanfield-type random model), the random energy model was introduced by B. Derrida in the early 80s. They are both suitable to model large systems having random interactions among their variables, and as a consequence exhibiting a complex behaviour. Whereas for the Ising-Lenz-type model the order parameter to describe the macroscopic behaviour is easily identifiable and correspondingly also the parameter describing phase transitions, for the case of spin glasses and the random energy model the identification of suitable order parameter constitutes a much more complex issue. G. Parisi in famous physical work invented a "replica trick" for finding such an order parameter and for producing an explicit formula (called Parisi's formula) for a quantity of great interest associated with such models. The mathematical justification of this formula from first principles has constituted an open problem since many decades, finally solved in hard work by Guerra and his coworkers and Talagrand, see also Panchuck in the references of Guerra's article. In this exposition, Guerra presents the essence of an interpolation method developed by himself and further developed by him with coworkers. By this method the annealed free energy of the replicated system (first replicated an integer number

s of times) is first expressed through variational principles, containing definite variational trial functions depending on special *s*-depending order parameters. Then the interpolation method permits to consider the case of a real number *s* of replicas. The limit where *s* is going to zero is investigated. It is shown that for generic values of *s* there is a unique variational principle, with a trial function and order parameters not depending on *s*. In this way from a suitable bound on the *s*-dependent order parameter entering the variational procedure at least in the random energy model one gets to the limit $s \to 0$. In the case of the spin glass model there is a functional order parameter, but the breaking of the replica symmetry can be interpreted as a phase transition in the numbers of replicas.

The structure of this important paper is as follows. Section 1 is a nice introduction. In Section 2, study of the annealed averages of the replicated free energy is provided together with the concept of an integer number of replicas and a related variational principle. The emergence of replica symmetry in the infinite volume limit is discussed. In Section 3, the extension to a real number *s* of replicas is provided. Section 4 discusses Derrida's random energy model. Section 5 explains very clearly and in detail the main results concerning the Sherrington-Kirkpatrick mean field spin glass model obtained by the very original methods of the author. Section 6 indicates some possible future developments concerning multispecies models, Hopfield network and neural networks.

Let us stress that this fundamental work also has implications for models which are at the best of many attempts to understand complex phenomena in, *e.g.* biology, neurobiology, information science, sociology, economics, geography and urbanism[1].

The contribution by Frédéric Patras and Victor Planas-Bielsa is entitled "Complex systems: from the presocratics to pension funds". The main point of the contribution (presented in its Introduction) is to describe and illustrate how complexity studies inspire new modes of thinking, beyond the traditional mechanistic and reductionistic methods and involving a rethinking of the very foundations of science. In particular, the authors advocate a direct approach to the question of the interplay between complexity and simplicity. Section 1 presents a contemporary view of the philosophical problems raised by presocratic philosophers (in particular, Heraclitus, Parmenides and Democritus), illustrating the actuality of their reflections by a simple example (going back to Poincaré). The authors also beautifully illustrate the relations with contemporary philosophical and epistemological discussions. Section 2 concentrates on the ideas of model and causality, also presented as having their roots in classical Greek philosophy, in particular, Aristotle. Applications of these ideas in two areas of contemporary research (quantum field theory and mathematical finance) are presented. Section 3 discusses various versions of atomism especially in relation with mathematics and its applications, illustrated by an example from economics. The authors present an original point of view of looking at the whole mathematics as a dynamical system, finding support in historical epistemology and

[1] See, *e.g.* references in Stein, D. L., & Newman, C. M. (2013). *Spin glasses and complexity*. Princeton University Press and Chakrabarti, B. K., & Sinha, A. (2021). Development of Econophysics: A Biased Account and Perspective from Kolkata. *Entropy*, 23(2), 254.

philosophy of science, in particular, the tradition from Braunschweig, Bachelard, Cavaillès, Lautman, up to contemporary philosophers. Section 4 takes up one problem of contemporary economics (pension plans) and its relations with complex systems and epistemological issues, arguing for a good use of mathematics (including its atomism) taking critical care of the limits of models.

The contribution of Ferdinando Semboloni with the title "From complex dynamics to the architecture of the city" presents a new theory for the design of a city. Its main point is to take in due account the self-organization components in cities and towns. This requires a reconsideration of the role of urban planning and design from a point of view inherited from enlightenment to a contrasting one, where the city appears as a complex system, to be understood and transformed taking into account its elements of spontaneous dynamics. This presentation of a new design of a new city proceeds as follows. The exposition starts out in Section 3 with a brief review of some basic theories of urban systems. Then Zipf's law of population among cities is described and interpreted as a sign of the presence of complex dynamics in cities. The exposition proceeds by Sections 3 and 4 putting in evidence through beautiful historical considerations how design and construction of a city are intimately related to historical contexts.

The author observes in Section 3 how for a view "from above" cities often appear as spontaneous artefacts and looking from a close distance they often appear as designed. He then analyses the designed cities in Section 4, with historical roots in ethical principles (as ideal cities).

Section 5 discusses the concept of self-organized cities and its relationship with the theory of complex systems.

In Section 6, the idea of artificial self-organized cities is critically analysed. The author argues for an approach where understanding the existing complexity can be used to improve the architects' ability to intervene in an urban asset.

In Section 7, scientific and architectonic ways of knowledge are compared and contrasted, with examples from the history of architectural projects.

Section 8 points out the relations between patterns in architectural contexts (following Christopher Alexander).

Section 9 digresses on imitation in arts and the role of time, with a mentioning of statistical mechanical versus individual behaviour.

Section 10 briefly describes cities as solutions of some problems of communication that are deepened in Section 11 where the role of a centre (of a town) is underlined.

Examples from architecture are mentioned, as well as centres in biological and economic contexts. The genesis of a centre and the centre-area pattern is analysed in Section 12, both in architectural and biological contexts. In Section 13, two basic patterns of a circle and an ellipse formed in cities are analysed.

In Section 14, a self-similarity of systems of centre-area patterns in a city is described and a semi-hierarchical organization of such patterns is pointed out.

In Section 15, a program developed with Netlogo is used to provide an example of city design. The relationships between cells are based on Zipf's law with a distribution in space inspired by Christaller's theory of central places.

In Section 16, the tracing of a road network is illustrated. It combines simulations and manual design.

In Section 17, the theory of space syntax is used with topological tools to analyse a road network and the production of a suitable urban space.

In Section 18, some conclusions are drawn. The author argues for facilitating communication between individuals. For this an analysis of spontaneous dynamics is advocated.

The contribution of Raffaello Seri, Davide Secchi and Mario Martinoli is entitled "Randomness, emergence and causation: a historical perspective of simulation in the social sciences". It gives a comprehensive review of computational models and techniques, in connection with simulations, particularly in their relations with social sciences and the aspects of randomness, emergence and causation. It also contains an historical account of the problems and methods and a discussion of the relation between parts and the whole, typical of complex systems. Section 2 gives an interesting account of the early developments of computers and computational tools, including the Monte Carlo method. Section 3 explains system dynamics (since about the 50s) as an approach to the understanding of the dynamics of composed objects in interaction (in terms of differential equations). Examples are provided (Lotka-Volterra and limits of growth models in biological and population dynamical contexts). Also, the problematic aspects of deterministic versus stochastic models are discussed, with emergence traced back to its economical roots. Section 4 discusses two examples of simulations, discrete-event simulation and query systems. Microsimulation is presented in Section 5, in connection with economics (in particular, contemporary approaches). Simulation by cellular automata is reviewed in Section 6, accompanied by discussion of examples (*e.g.* voting behaviour, Schelling's model). Contemporary agent-based models (ABM) are discussed in Section 7 and the difference between them as computational tools with simulation methods and analysed in detail and illustrated by many examples (taken from social and practical sciences, economics, biology, management and organization research). The advantage of approaches based on ABM is put in evidence, as the most advanced present-day computational approach for computational simulation.

Bonn, Germany — Sergio Albeverio
Varese, Italy — Elisa Mastrogiacomo
Milano, Italy — Emanuela Rosazza Gianin
Milano, Italy — Stefania Ugolini

Acknowledgments

The editors would like to express their gratitude to all contributors and lecturers for the work and the time they invested in preparing the beautiful lectures and the contributions appearing here. Our thanks go also to Ivar Ekeland and Florent McIsaac, as well as Frank Riedel, who presented beautiful lectures on "The emergence problem in economics", respectively, "Knightian Uncertainty in Finance and Economics", but unfortunately were not able to contribute to these Proceedings of the School. The Conference was accompanied by an evening public lecture held by Nadia Robotti, whom we most heartedly thank for her excellent presentation of a very interesting historical personality of last century (namely, the great physicist Bruno Ponte Corvo). She reported on investigations by herself with Francesco Guerra; their work will appear elsewhere. We also thank Fabio Minazzi for an early participation in the organization of the School, unfortunately for family reasons he was not able to attend.

We would also like to thank our friend and colleague Danilo Merlini and all students who participated in the School for lively and stimulating discussions. Last but not least, let us express our gratitude to Giulio Casati, Director of the Lake Como School of Advanced Studies, for inspiring discussions and for giving us the opportunity to carry through the project. We also thank all coworkers of the School for the logistic support. The financial support of the School and of the Alessandro Fondazione Volta in Como. They should also be thanked for their important enhancement of advanced training and research in the Lombardy University System.

A particular hearty thanks to Marina Reizakis of Springer Verlag for her great support in making this volume possible. We also thank Banu Dhayalan and Monica Janet of Springer-Verlag for their technical support in the preparation of the printing process.

Contents

A Topological and Dynamical Approach to the Study of Complex Living Systems

Luciano Boi

Abstract Traditionally one has addressed biological phenomena by analytical and/or statistical tools with the goal to give average laws of the competition between biological species and communities and of their evolution. This approach has certainly permitted to obtain important results. However, many fundamental aspects of biological systems at every scale remain to be explained. In the last four decades new mathematical ideas and methods have emerged, relying notably on differential geometry, topology and dynamical systems, which appear to be very meaningful for the modelling of many processes which occur within the cell and organism contexts and activity. In particular, these ideas and methods promises to be helpful for a better understanding of the following three correlated issues (i) the relationship between the plasticity of biological structures and the emergence of their functionality, (ii) the interaction between form and function: how the first influences the second and, conversely, how, the second may favors adaptive changes of the first, (iii) the connection between the robustness and fragility in complex living organism; in particular the difficulty is to characterize the critical thresholds which may occur a transition from robustness to fragility.

Keywords Analytical methods · Geometry · Topology · Knotting · Twisting · Chirurgy · Dynamical systems · Plasticity · Complexity · Form · Function · Robustness · Fragility

L. Boi (✉)
École des Hautes Études en Sciences Sociales, Centre de Mathématiques (CAMS), 54, bd Raspail, 75006 Paris, France
e-mail: lboi@ehess.fr

S. Albeverio et al. (eds.), *Complexity and Emergence*, Springer Proceedings in Mathematics & Statistics 383, https://doi.org/10.1007/978-3-030-95703-2_1

Introduction

1 Introduction: Complex Structures in Physics

One of the most exciting developments of mathematical physics in the last four decades has been the discovery of numerous intimate relationships between the topology and geometry of knot theory and the dynamics of many domains of "classical" and "non-classical" macroscopic physics, especially soft matter physics and biophysics. Indeed, complex systems of knotted and entangled filamentary structures are ubiquitous in Nature and arise in such disparate contexts as electrodynamics, magnetohydrodynamics, fluid dynamics (vortex structures), superfluidity, dynamical systems, plasma physics, cosmic string theory, chaos of magnetic flows and non-linear phenomena, turbulence, polymer physics and molecular biology (see Boi [1–3], Boyland [4], Hornig [5], Kauffman [6] for a detailed account of the subject). In the last decades, mathematical tools have been developed to identify and analyse the geometrical and topological complex structures and behaviours of such systems and relate this information to energy levels and stable states of matter.[1]

2 Topological Stability in Condensed Matter and Phase Transition

Recent works established the role of topology in understanding exotic forms of condensed matter (for more details we refer to AAVV [8], Kosterlitz and Thouless [9]). Holes and their number define the shape of many kinds of objects, in the following sense they can be squished and stretched without changing essentially their shape, in other words, they posses the property of (a certain) stability. If we add or subtract one or more holes (closed loops) to these objects, they will change their shape and yet conserve some topological stability, and this fact has important physical effects. For example, when we add some holes to a complex Riemann

[1] Recall that four states of matter are observed in everyday life: solid, liquid, gas and plasma. Many other states are known such as Bose–Einstein condensates and neutron-degenerate matter but these only occur in extreme situations such as ultra-cold or ultra-dense matter [7]. Other states, such as quark-gluon plasmas, are believed to be possible but remain theoretical for now. The distinction between these states was first made based on qualitative differences in properties. Matter in the solid state maintains a fixed volume and a stable shape, with component particles (atoms, molecules or ions) close together and fixed into place. Matter in the liquid state maintains a fixed volume, but have a variable shape that adapts to fit its container. Its particles are still close together but move freely. Matter in the gaseous state has both variable volume and shape, adapting both to fit its container. Its particles are neither close together nor fixed in place. Matter in the plasma state has variable volume and shape, but as well as neutral atoms, it contains a significant number of ions and electrons, both of which can move around freely. Plasma is the most common form of visible matter in the universe.

surface—such as a sphere or a torus—by continuous or discontinuous deformations, we change its shape without altering its essential topological properties. Similarly, topological phases have a robustness that allows their properties to remain stable in spite of the impurities or other material details. To explain this robustness, Thouless's group considered all of the wave functions that can describe electrons in a two-dimensional (2D) material and represented this set of possibilities with a curved surface. The shape of this surface can be classified by a number called *the genus of the surface*, which is a topological invariant and which has some stability, like the vortices in the Kosterlitz-Thouless transition (see below). A common example of a topological invariant is the number of holes, or the *genus*, in an object, like a donut or a pretzel. One can distort a pretzel quite a bit, but the hole count doesn't change. With some effort, one can puncture or cut the pretzel to make more or fewer holes, but the hole count always jumps by an integer amount **Z** (there are no half holes). Thouless and his colleagues showed that their derived topological invariant was related to the integers that define the Hall conductance steps. This result account for the quantization and also explained why the Quantum Hall Effect (QHE) is so robust: small changes to the material can affect the set of electrons wave functions, but the topological invariant that describe the set is much harder to change.

The general and fundamental idea behind these results is that topology can be used to characterize the phases of matter. In other words, the study of some topological forms allows to predict and characterize new forms of matter with stable states.

3 Epistemological Remarks on the Limits of the Hierarchical Paradigm of Science

According to the belief of almost all scientists, matter is organized in a hierarchical structure. At the bottom we have elementary particles: quarks, gluons, electrons, etc. At present we don't know what these particles are made off. But we believe we know that quarks, gluons, etc. make up protons, neutrons, etc. These then goes together to make atoms, that are building blocks of molecules. Atoms and molecules give gas, liquids and solids from which we get stars and planets which are grouped together in galaxies, that then form clusters and eventually we arrive at the entire universe. Or from atoms and molecules we get macromolecules like proteins and DNA, that are building blocks of organelles, which together form the cells. From cells we get organs, that put together form organisms: animals and plants of a great variety of species. The totality of individuals and species constitutes the entire ecology.

One branch of science, e.g. particle physics, is concerned with the breaking of systems into smaller and smaller parts. The behaviour and properties are studied at each respective level. However, the theory of dynamical systems, thermodynamics and statistical mechanics are concerned with the opposite quest. Namely, from the interactions between the components, say atoms, at one given level the aim is to understand the collective coherent behaviour which emerges as many atoms or molecules

are but or assembly together and the next level is formed. Often the microscopic details of the properties of the *individual* building blocks are not so crucial. Rather it happens that the collective behaviour is controlled by general properties of the *interaction* between the building atoms or molecules (see Anderson [10], Jost [11], and Kitano [12]).

There is a particular case where it is possible to follow in detail how components at one level assemble and (in some phenomena) organize together and form certain collective coherent structures: *topological defects or topological charges.* These charges can be Coulomb charges in two dimensions, dislocations in two dimensional crystals, vortices in two dimensional superconductors and more. The interaction between the topological charges depends in all cases logarithmically on the spatial separation and this leads to some very general collective behaviour, most spectacular it causes a certain type of phase transition: the Kosterlitz-Thouless transition.

These authors identified a completely new type of phase transition in 2D systems—such as thin fluid films or single-layer materials—where topological defects play a crucial role. Their theory applied to certain kinds of magnets and to superconducting and superfluid films. In the early 1980s David J. Thouless and F. Duncan M. Haldane developed theoretical methods to describe phases of matter that cannot be identified by their pattern of symmetry breaking.

Crystalline solids are a very important class of materials in which the atoms are arranged in periodic patterns. These patterns can be classified by their symmetries. When a liquid solidifies into a crystal, it changes from a phase which is, on macroscopic scale, invariant under both translations and rotations, to a phase where these continuous symmetries are broken down to a finite symmetry group characteristic of the crystal. Another example of such a *phase transition* occurs when a ferromagnet is cooled below the *Curie temperature* T_C or *Curie point* (it is the temperature above which certain materials lose their permanent magnetic properties; in other words, the magnetism is lost at a *critical* temperature), and the atomic magnetic moments, or the spins, line up and give rise to a net magnetization.

The study of magnetism has been very important for our understanding of symmetries in physics. Using new experimental techniques, hidden patterns of symmetries were discovered. For example, there are magnetic materials where the *moments* or the *spins* form a chequerboard pattern where the neighbouring moments are antiparallels. In spite of non-having any net symmetrisation, such *antiferromagnets* are nevertheless ordered states, and the pattern of microscopic spins can be revealed by neutron scattering. The antiferromagnetic order can again be understood in terms of the associated symmetry breaking.

There is a striking similarity between 2D magnet and certain superconducting or superfluid films. The magnetisation is a vector that normally can point in any direction but in certain magnets, the spins are constrained to lie in a plane, say the *xy*-plane, where they are free to rotate. In such an "easy-plane" magnet the direction of the magnetization is determined by a single angle, θ, denoting the rotation around the *z*-axis. It is important now to stress that there are configurations of such planar, or *xy*, spins that are *topologically distinct*. One of such configurations is a

vortex configuration: in fact, we can have a single vortex configuration and a vortex-antivortex pair, which can be smoothly transformed to the ground state. The vortex is a *topological defect* that cannot be transformed into the ground state where all the spins are aligned, by a continuous rotation of the spins. An important result is that the configurations can be classified by their vorticity, which is a *topological invariant* (see Ricca and Moffat [13], Boi [14], Oberti and Ricca [15]).

The classical example of a phase transition is a system going from a disordered phase to an ordered phase as the temperature is lowered below a critical value. More recently, the phase transition concept has been extended to quantum systems at zero temperature. A quantum system can undergo a radical change of its ground-state as a parameter in its Hamiltonian, such as pressure, magnetic field or impurity concentration, is tuned trough a critical value, and such a *quantum phase transition* signals the change from one state of matter to another. This insight has provided an important link between statistical mechanics, quantum many-body physics and high energy physics.

Kosterlitz and Thouless, in the paper already quoted (1973), showed that there is indeed a finite temperature phase transition, but of a new and unexpected nature where the vortex configuration plays an essential role. One year before the work of Kosterlitz and Thouless, Vadim Berenziskij (1972) also recognized the importance of vortex excitations in the *xy*-model. He understood that they could drive a phase transition, but did not correctly describe the nature of this finite temperature transition, which is therefore referred to as the "KT-transition". The XY-model, which allows to describe different systems like the planar magnet and a superfluid,—both have complex order parameters described by a single angle (specifically for a superconductor or superfluid the complex order parameter can be expressed as $\varphi = \sqrt{\rho_s} e^{i\theta}$, where ρ_s is the superfluid density and θ the phase)—is defined by the Hamiltonian

$$H_{XY} = -J \sum_{\langle ij \rangle} \cos(\theta_i - \theta_j) \tag{1}$$

where $\langle ij \rangle$ denotes nearest neighbours and the angular variable, $0 < \theta_i \ll \pi$, can denote either the direction of an XY-spin or the phase of a superfluid.

An important point to be considered is the existence of the thermal fluctuations, which prevent ordering of XY-spins in two dimensions. A more precise statement is based on the large distance behaviour the spin–spin correlation function. To see what happens in two dimensions, we take the continuum limit of the Hamiltonian (Eq. (1)) to get

$$H_{XY} = J/2 \int d^2 r (\nabla \theta(r))^2. \tag{2}$$

A simplification is to extend the range of the angular variable to $-\infty < \theta < \infty$ to get a free field Hamiltonian and thus Gauss fluctuations, and a direct calculation using a short distance cut-off a give

$$\left\langle e^{i(\theta(r)-\theta(0))}\right\rangle \sim (a/r)k_B T/2\pi J. \tag{3}$$

This is a power law even at high temperatures, where an exponential fall-off would be expected. Kosterlitz and Thouless resolved the apparent paradox by showing that there is indeed a finite temperature phase transition but of a new kind and essentially topological in nature.

The glitch in the argument leading to Eq. (3) is that the periodic, or $U(1)$, nature of θ cannot be ignored, since that amounts to neglecting vortex configurations. A vortex is characterized by a non-zero value of the vorticity

$$v = 1/2\pi \int_C d^2r \cdot \nabla\theta(r) \tag{4}$$

where C is any curve enclosing the centre position of the vortex. The integral measures the total rotation of the spin vector along the curve, so after dividing with 2π, v is simply the number of full turns it makes when circling the vortex. From this, we also understand there can also be antivortices, where the spins rotate in the opposite direction. For a rotationally symmetrical vortex with $v = \pm 1$ it follows from last equation that $|\nabla\theta(r)| = 1/r$, so the energy cost for a single vortex becomes

$$E_v = J/2 \int d^2r(1/r)^2 = J\,\pi \ln L/a \tag{5}$$

where L is the size of the system, and a a short distance cut-off that can be thought of as the size of the vortex core.

Even in the case in which the energy cost for a single vortex is very large, and cannot be excited by thermal fluctuations, the effect of the vortex cannot be neglected. In other words, the vortex is a topological object whose shape and properties may induce a change of structure and behaviour in the physical system, precisely a phase transition, even in absence of a critical parameter breaking the symmetry of the system.

We can understand the essence of this new type of topological phase transition by a quite simple thermodynamic argument. Although the energy of a single vortex diverges as *ln L*, this is not true for vortex-antivortex pairs since they have zero total vorticity. The energy required to create such a pair is $J2\pi$ *ln r/a*, where r is the separation between the vortices. Such pairs can thus be thermally excited, and the law temperature phase will host a gas of such pairs. The insight by Kosterlitz and Thouless was that at a certain temperature T_{KT} the pairs will break up into individual vortices. It is this vortex pair unbinding transition that will take the system to a high temperature phase with exponentially decaying correlations.

4 The Spin-Like Representation of the KT Topological Phase Transition

The general idea behind the KT-transition, which is an effect that shows up in thin films of magnetic material, can be simply explained in terms of a geometric-spins representation. Imagine a thin film of stuff where each atom's spin likes to point in the same direction as its neighbours. Also suppose that each spin must point in the plane of the material; in other words, all the spins of the magnet material are lined up. Physically, this means something precise: it means that the energy is low. When the stuff is very cold, its energy will be as low as possible, so the spins will line up. When one heats up the thin film, it gets a bit more energy so the spins can do more interesting things. An interesting possibility is when they form a vortex: the spins swirl around like the flow of water in a whirl-pool; each spin is fairly close to being lined up to its neighbours, except near the middle where they are moving in a strange way. The total energy of a vortex is enormous. The reason is that fairly close is not good enough. The spins fall to perfectly line up with their neighbours even far away from the middle. (In fact, the energy would be i*nfinite* if our thin film or material went on forever.) So, even if you heat up your substance, there won't be enough energy to make many vortices. One might then think that vortices were irrelevant in this story. But there is another possibility, called an "anti-vortex": A single vortex has a huge energy, just like a vortex. So again, it might seem antivortices irrelevant if one is wondering what the stuff will do when it has just a little energy.

At this point Kosterlitz and Thouless noticed that the combination of a vortex and an antivortex has much less energy than either one alone! So, when the thin film of stuff is hot enough, the spins will form vortex-antivortex pairs. We can imagine (the moving pictures of) how this might happen. A vortex-antivortex pair can appear out of nothing… and then (somehow magically) disappear again. Thank to this process, at low temperatures the thin film will contain a dilute "gas" of vortex-antivortex pairs. Each vortex will stick to an antivortex, since it takes a lot of energy to separate them. These vortex-antivortex pairs act a bit like particles: they move around, bump into each other, and so on. But unlike most ordinary particles, they can appear out of nothing, or disappear, in this process.

As one heats up the thin film, one gets more and more vortex-antivortex pairs, since there is more energy available to create them. But here a real surprising thing occurs. Kosterlitz and Thouless showed that as one turns up the heat, there is a certain temperature at which the vortex-antivortex pairs suddenly "unbind" and break apart! Why? Because at this point suddenly something happens at some particular temperature. It is bit like how ice suddenly turns into liquid water when it warms above its melting point. A sudden change in behaviour like this is called a phase transition. So, the Kosterlitz-Thouless transition is the sudden unbinding of the vortex-antivortex pairs as one heats up a thin film of stuff where the spins are confined to a plane and they tend to line up. Superconducting materials too can exhibit a Kosterlitz-Thouless transition. Indeed, their work was the key that unlocked a treasure room full up of strange new states of matter, called "topological phases"; these states show clearly

that the "usual" concepts of symmetry and symmetry-breaking are enriched and must be completed by the concept of topological deformations whose resulting forms may generate new states of matter and a more subtle notion of stability of these states, in spite of the thermal fluctuations and local spatial dislocations that apply to them. Such stable new states of condensed matter, but almost the same can be said of soft and living matter, emerge from these topological deformations and from some topological invariants, such as the genus, which are related (and allow for characterising) some classes of "objects" (surfaces and manifolds).

Vortex are one of the most significant (and surprising) topological objects showing interesting physical properties and behaviours. Let's start with the question: How can a vortex and an antivortex be defined? Elementarily in this way. If one walk around either one and look at the little arrows, the arrows turn around—one full turn. It is a vortex if when one walk around it clockwise the little arrows make a full turn clockwise, like in the first picture above. And it is an antivortex if when one walk around it clockwise the little arrows make a full turn counter-clockwise, like in the second picture above. Vortex can be characterized topologically by its *winding number* whose value is 1 (1 positive for a counter-clockwise turning), while the antivortex has winding number -1 (1 negative for clockwise turning). Roughly speaking, the *winding number* of a closed curve in the plane around a given point is an integer ($\mathbf{Z}$) representing the total number of times that curve travels counter-clockwise around the point. The winding number depends on the orientation of the curve, and is negative if the curve travel around the point clockwise (for more details see White [16], Gray [17], and Roe [18]).

Definition. *The winding number is one the most basic invariant in topology. It measures the number of times a moving point P goes around a fixed point Q, provided that P travels on a continuous path that never goes through Q and the final position of P is the same as its starting position.*

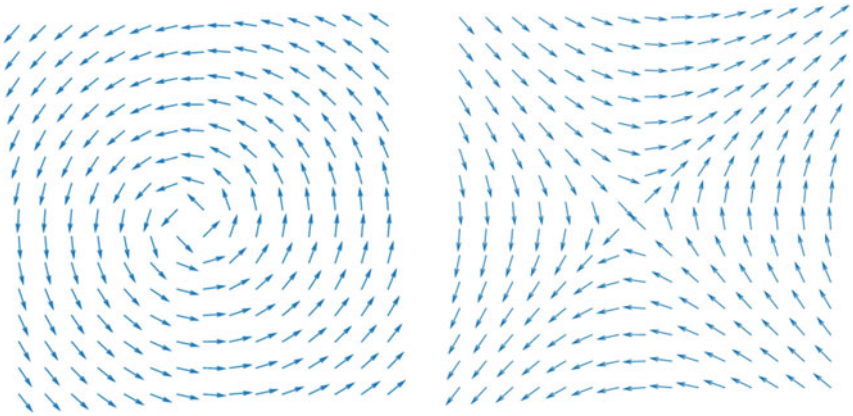


Fig. 1 A vortex: the arrows are oriented toward the same direction. An antivortex: the arrows are oriented toward different directions

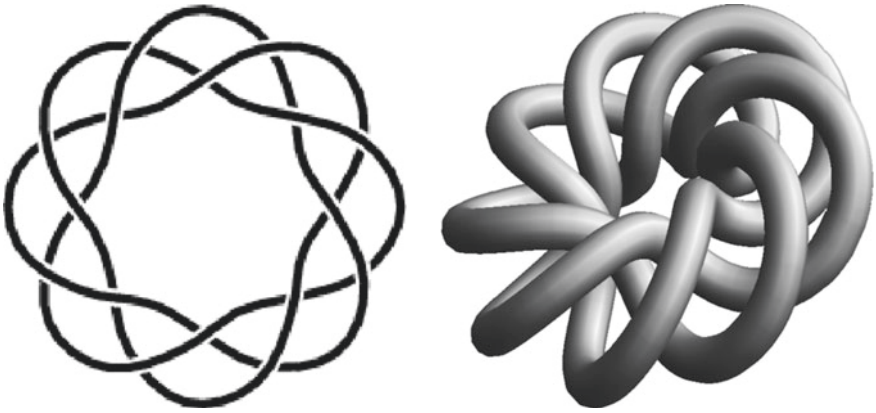


Fig. 2 A simple representation of a torus knot. Precisely, it is a torus knot K(3, -8) represented as a knot diagram (left) and in three-dimensional space with volume (right)

This simple idea has far-reaching applications, firstly in almost all fields of mathematics and in theoretical physics (string theory and topological quantum field theories), but also in condensed matter physics and in molecular biology.

Recall that the winding number is one of the most important concepts in the topological theory of knots. The class of torus knots can be characterised as follows (see Oberti and Ricca [19]).

"Torus knots and unknots are rotationally symmetric closed curves standardly embedded on a mathematical torus T in P^3. Each torus knot is de ned by a pair of co-prime integers $p > 1$ and $q > 1$, denoting the number of full turns around T done by the curve along the longitudinal (or toroidal) direction and meridian (or poloidal) direction, respectively (see Figs. 2 and 3). The *unknot* is given by either $p = 1$ or $q = 1$. When $p = q = 1$ the unknot is a twisted circle on T, and when p and q are rational, but not relatively prime, we have links (with number of components given by the greatest common divisor between p and q). The ratio $w = q/p$ ($w > 0$) defines the *winding number* and is a measure of the knot topology. When $w = q/p$ is irrational the curve forms a dense set covering T almost everywhere. Two limits are of interest: (i) a *poloidal* hollow ring covered by infinitely many poloidal coils, when p is fixed and finite, and $q \to \infty$; (ii) a *toroidal* hollow ring, when q is fixed and finite, and $p \to \infty$" (2019, 17284).

Oberti and Ricca provided a mathematical proof of the effects of winding number—a topological invariant torus knots—are of primary importance on the motion of vortex knots in the context of classical and, ideal (i.e. inviscid) fluid mechanics. This has been done by considering the full Biot-Savart law.

(Recall that this law enables us to calculate the magnitude and direction of the magnetic field produced by a current in a wire. The Biot-Savart law states that at any point P, the magnetic field dB due to an element dl of a current-carrying wire is given by

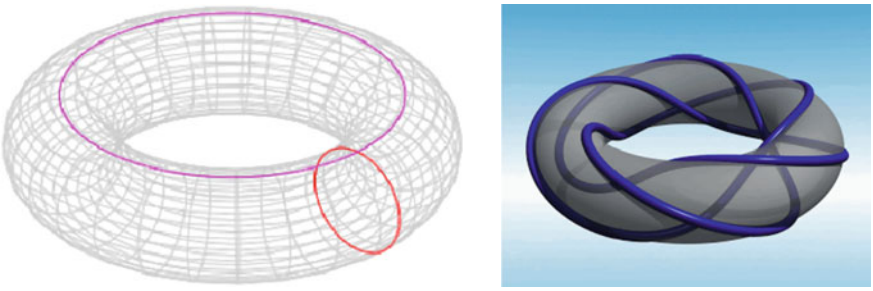


Fig. 3 The outside shell of a donut or bagel is a good concrete representation of what we call a mathematical torus. A torus knot is a knot (i.e. a tangled string in our 3-dimensional space) laying on the surface of the torus shape, without intersecting itself in any place. The torus is made up of two types of circles (**a**), the meridian curve, which wraps one the short way around the torus, and the longitudinal curve, which wraps once the long way around the torus (**b**). Topologically, the torus is the product of two smooth circles $T \cong C \times C'$. Torus knots can be named as $T(p, q)$, where the knot $K(p, q)$ crosses over the meridian curve (a circle) p times, and the longitudinal curve (the other circle) q times. An example would be the (3, 5) torus knot $K(3, 5)$ in (**b**). It crosses around the meridian curve three times. (From *Wolfram Mathworld*.)

$$dB = \mu_0/4\pi\, Idl \times r/r^2. \tag{6}$$

The constant μ_0 is known as the permeability of free space and is exactly

$$\mu_0 = 4\pi \times 10^{-7}\, T \cdot m/A \tag{7}$$

In the SI system. The infinitesimal wire segment dl is the direction as the current I (assumed positive), r is the distance from dl to P and r is a unit vector that points from dl to P.)

The influence of winding number is found to be comparable to curvature effects (which, as we know, are essential in almost all domains of fundamental physics and beyond) and for thin filaments it is more important than the vorticity distribution over the vortex cross-section. The authors also determined, in full generality, precise relations between winding number, knot complexity and relative velocity contributions. They computed the speed for most common knot types and showed that the propagation velocity increases with the number of toroidal coils; moreover, they proved that for increasing ratio and number of poloidal coils vortex can be even reversed.

For a formal definition of the linking number consider a curve in the xy-plane defined by the parameter equations:

$$x = x(t) \text{ and } y = y(t) \text{ for } 0 \leq t \leq 1. \tag{8}$$

If we think of the parameter t as time, then these equations specify the motion of an object in the plane between $t = 0$ and $t = 1$. The path of this motion is a curve as long as the functions $x(t)$ and $y(t)$ are continuous. The curve is closed as long as the position of the object is the same at $t = 0$ and $t = 1$. We can define the winding

number of such a curve using the polar coordinate system. Assuming the curve does not pass through the origin, we can rewrite the parametric equations in polar form

$$r = r(t) \text{ and } \theta = \theta(t) \text{ for } 0 \leq t \leq 1. \tag{9}$$

The function $r(t)$ and $\theta(t)$ are required to be continuous, with $r > 0$. Because the initial and final positions are the same, $\theta(0)$ and $\theta(1)$ must differ by an integer multiple of 2π. This integer is the winding number:

$$w = \theta(1) - \theta(0)/2\pi \tag{10}$$

This defines the winding number of a curve around the origin in the *xy*-plane. By translating the coordinate system, we can extend this definition to include winding numbers around any point.

In differential geometry, the winding number of a differentiable curve can be expressed as a line integral:

$$w = 1/2\pi \int_C (x/r^2 dy - y/r^2 dx). \tag{11}$$

In topology, the winding number is an alternate term for the degree of a continuous mapping [20]. The above example for a curve winding around a point has a simple topological interpretation. The complement of a point in the plane is homotopy equivalent to the circle, such that *maps* from the circle to itself are really all that is needed to be considered. It can be shown that each such map can be continuously deformed to (is isomorphic to) one of the standard maps $S^1 \to S^1$: $s \to s^n$, where multiplication in the circle is defined by identifying it with the complex unit circle. The set of homotopy classes of maps from a circle to a topological space forms a group, which is called the *first homotopy group* or *fundamental group of that space.* The fundamental group of the circle is the group of the integers **Z**; and the winding number of a complex curve is just its homotopy class. Maps from the 3-sphere to itself are also classified by an integer which is also called the *winding number* or sometimes *Pontryagin index*.

The simplest example of the concept of linking coefficient or linking number is that of two non-intersecting closed rectifiable curves L_1, L_2 in R^3, given by the so-called Gauss integral:

$$I = 1/4\pi \int_{L1} \int_{L2} (x_1 - x_2) dx_1 dx_2 / |x_1 - x_2|^3 \tag{12}$$

(Here x_1 and x_2 are the radius vectors of L_1 and L_2).

The concept of linking number generalizes to the case of a pair of closed oriented manifolds M^{k-1} and M^{n-k} in R^n: the linking number is equal to the degree of the mapping χ of the oriented direct product $M^{k-1} \times M^{n-k}$ into the sphere $S^{n-1} \subset R^n$, where $\chi(x, y)$, $x \in M^{n-k}$, is the point at which S^{n-1} is cut by a ray through the origin

parallel to the vector (x, y). The linking number is equal to the intersection index (in homology) of any k-chain C^k such that $\partial C^k = \alpha z^{k-1}$ with the cycle z^{n-k}, divided by α. This member is independent of the choice of C^k. (For more details, we refer to Dubrovin et al. [21], and Ranicki [22])

5 The Effectiveness of the Concept of Winding

The notion of *winding* is effective and has a precise geometrical and dynamical meaning in physics and in biology, especially in fluid mechanics and in molecular biology (see Ophl and Roberts [23], White [24], and Boi [25]). Consider, for example, the structure and function of DNA. As it well-known, the DNA must be compacted more than ten thousand folds in the cell, and this might explain why almost any protein that binds to DNA will bend it. Moreover, since the curvature K[2] of an entire DNA double-helix segment depends on the torsional stress which applies to DNA strands, these strands must form a twisted curve, i.e. a curve of double curvature in the three-dimensional space of the cell nucleus (see Olsen and Bohr [26]). DNA double-helix must coil many times in a very ordered way to form thus a chromatin structure, otherwise, if the chromosomes of the human cell were in the form of a random coil, they will not fit within the nucleus (see Almouzni [27], Ridgway and Almouzni [28], Cozzarelli and Holmes [29], and Boi [30]). The DNA double-helix coils first by overwinding or underwinding of the duplex. The supercoiled form of a circular DNA molecule minimizes to the highest the space volume it occupies in the nucleus.

It is worth to note that the supercoiled DNA molecule follow the most important property of minimal surfaces (as helicoid and catenoid): *they minimize the area.* Supercoil condense DNA and promote the disentanglement of topological domains. Recall that a surface $\sum \subset R^3$ is minimal if and only if every point $p \in \sum$ has a

[2] The *Gaussian curvature* (or *total curvature*) K of a surface S is an intrinsic property of space independent of the coordinate system used to describe it; in other words, it depends only on distance that are measured on the surface, not on the way it is isometrically embedded in Euclidean space. The Gaussian curvature of a regular surface in R^3 at a point p is formally defined as $\mathrm{K}(p) = \det(S(p))$, where S is the shape operator and det denotes the determinant. If x: $U \to R^3$ is a regular patch, then the Gaussian curvature is given by $\mathrm{K} = eg - f^2/\mathrm{EG} - \mathrm{F}^2$, where E, F, and G are coefficients of the second fundamental form. The Gaussian curvature can be given entirely in terms of the first fundamental form (as demonstrated by C.F Gauss in its *Theorema Egregium*, 1827) $ds^2 = \mathrm{E}du^2 + 2\mathrm{F}dudv + \mathrm{G}dv^2$, and the metric discriminant $g \equiv \mathrm{EG} - \mathrm{F}^2$, by

$$K = 1/\sqrt{g}\Big[\partial/\partial v(\sqrt{g}/E\,\Gamma_{11}^2) - \partial/\partial u(\sqrt{g}/E\,\Gamma_{12}^2)\Big],$$

where Γ_{ij}^k are Christoffel symbols of the first kind. The Gaussian curvature is positive at an elliptic point, negative at a hyperbolic point, and it is zero at a parabolic point or a flat point. Since the Gaussian curvature depends on the intrinsic metric only, i.e. on the coefficients of the first fundamental form, it is invariant under isometric deformations of the surface (see Gray [17, pp. 373–380 and 481–500)].

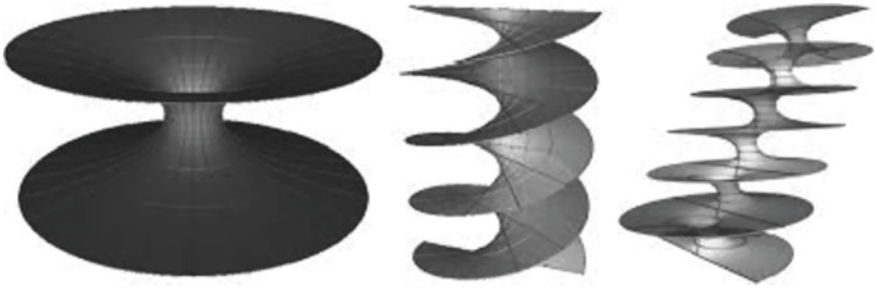


Fig. 4 Properly embedded minimal planar domains $M \subset R^3$ with infinite topology (equivalently, with an infinite number of ends) are Riemann minimal surfaces, in the precise sense that such an M must be congruent to a homothetic scaling of one of the classical examples found by Riemann in 1860. Recent work (see Meeks and Pérez [32]) demonstrate that the plane, the catenoid and the helicoid are the only properly embedded minimal surfaces of genus zero (without holes) with finite topology (equivalently, with a finite number of holes). This is so since the surfaces $\Re_s$ converges to a catenoid as $s \to 0$ and to a helicoid as s $\to \infty$. Then the moduli space M of all properly embedded non-planar, minimal planar domains in R^3 is homeomorphic to the closed unit interval [0, 1]. One of the fundamental results obtained by Meeks and Pérez (stated as a *Theorem*) asserts that *Up to scaling and rigid motion, any connected, properly embedded, minimal planar domain in* R^3 *is a plane, a helicoid, a catenoid or one of the minimal examples. In particular, for every such surface there exists a foliation of* R^3 *by parallel planes, where each plane intersects the surface transversely in a circle or a straight line* (2009, 283) (Figure from Meeks and Pérez [32])

neighbourhood with least-area relative to its boundary. Equivalently, a surface $\sum \subset R^3$ is minimal if and only if its mean curvature[3] vanishes identically. The shape of DNA and of any protein are minimal surfaces, or at least close approximations to them (see Hildebrandt and Tromba [31], Meeks and Pérez [32]). Why are minimal surfaces formed and how does this shape affect the local and global environment of the system? For DNA, the following explications can be evoked. DNA-proteins condensation in chromosomes calls for a packing mechanism that is fast, reliable and that causes as little change as possible in the interaction between the DNA molecule itself and the intracellular fluid in which it is immersed (Figs. 4, 5 and 6).

It is now understood that DNA is packed along hierarchical levels of rising complexity (see Boi [33]). The lowest level is of course the DNA double-helix itself. This then transforms in the nucleosomes, whose assembly form the chromatin, level

[3] The *mean curvature* of a surface $\sum$ in 3-dimensional space R^3 if the half of the sum of the principal curvatures k_1 and k_2, calculated at a point p of this surface: $H(p) = k_1 + k_2/2$. The mean curvature of a surface in R^3 can be expressed by means of the coefficients of the first and second fundamental forms, calculated at a point $p \in \sum$. For a m-dimensional submanifold M of a n-dimensional Euclidean space of codimension $n - m > 1$, the mean curvature generalizes to the notion of the mean curvature normal

$$V_p = 1/m \sum\nolimits_j 1^{n-m}[\mathrm{Tr}\, A(e_j)]e_j,$$

where $e_1, \ldots, e_{n-m}$ is the orthonormal form of the normal space of M at p and $A(e_j)$: $T_pM \to T_pM$ (T_pM denotes the tangent space to M at p) is the shape operator of M at p in the direction e_j, which is related to the second fundamental form V of M at p by $\langle A(e_j)(X), Y\rangle = \langle V(X, Y), e_j\rangle$.

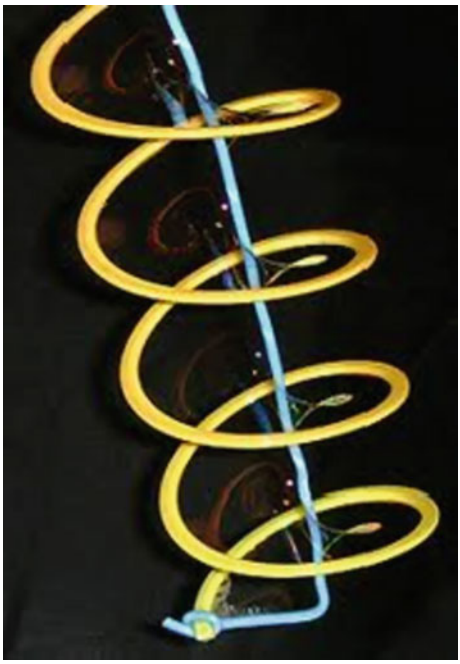


Fig. 5 A minimal surface. It looks like a one-strand right-handed model of DNA helix

where the double helix is wound around a highly specific protein cluster creating the bead-on-a-string form of chromatin. At the next level, the bead-on-a-string is condensed to a chromatin fibre, again aided by a highly specific protein, which is further compacted to the final *metaphase chromosome* (for more details see Gasser [34], Cremer et al. [35]).

The question is now how to account for a more than 10,000-fold compactification that takes place rapidly (in less than 0.5 ms) and yet is so gentle that the fragile DNA molecule, that will break when pipetted, is preserved intact throughout the transformation, and will survive through a large number of repetitions of this folding. The main reason for this remarkable property of a DNA molecule is its general shape. The double-helix sits on a helicoid,[4] and therefore it shares the property of that

[4] The helicoid, H, is a single periodic minimal surface swept out by horizontal lines moving at a constant speed up the x_3-axis while rotating at constant speed. It is invariant under any vertical screw motion around the x_3-axis, in particular vertical translations by 2π. Otherwise stated, the (circular) helicoid is the minimal surface having a (circular) helix at its boundary. It is the only ruled surface (i.e., a surface formed by a motion of a straight line) other than the plane. For many years, the helicoid remained the only known example of a complete embedded minimal surface of finite topology with infinite curvature. However, in 1992 a second example, known as Hoffman's minimal surface and consisting of a helicoid with a hole, was discovered. The helicoid is the only non-rotary surface which can glide along itself. It can be described as a straight line that rotates at a constant angular rate around a fixed axis, intersects the axis at a constant angle α, and at the same time becomes gradually displaced at a constant rate k along this axis. The equations of a helicoid in cylindrical coordinates is $z = c\theta$; in cartesian coordinates, it is $y/x = \tan(z/c)$; and it can be given

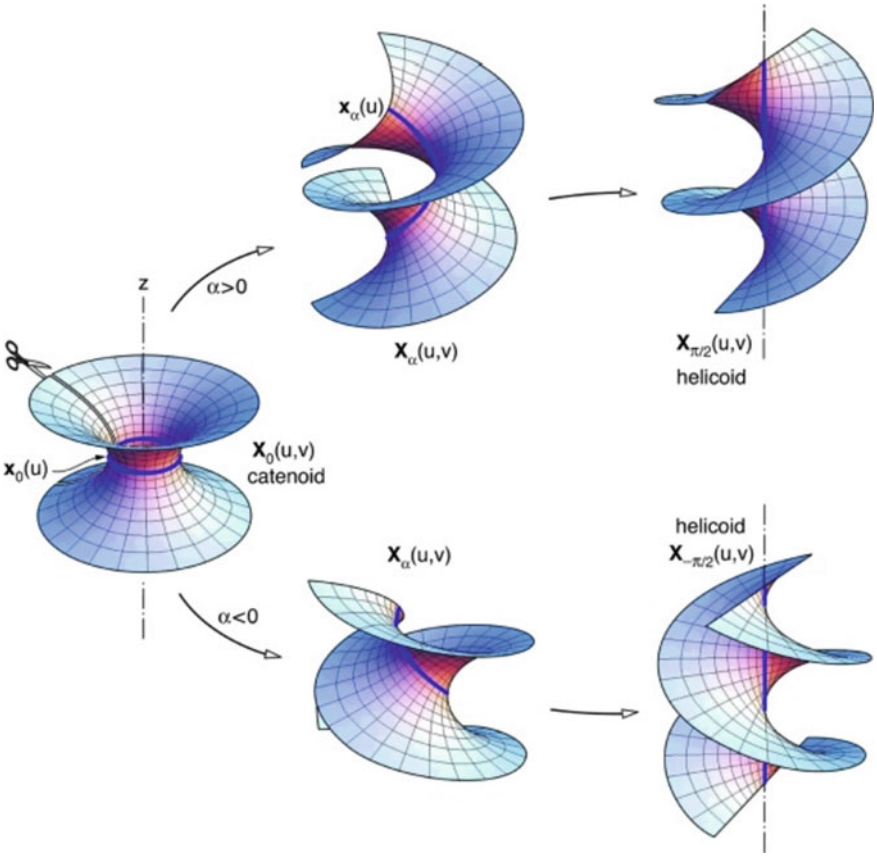


Fig. 6 The steps of the transformation of a catenoid (by cutting along $x_0(u)$) into a right-handed ($x_{\pi/2}(u, v)$ helicoid, and a left-handed helicoid ($x_{-\pi/2}(u, v)$. (From *Wolfram Mathworld*.)

in parametric form by $x = u \cos v$, $y = u \sin v$, $z = cv$. The Gaussian curvature of a helicoid is given by $K = -c/(c^2 + u^2)^2$, and the mean curvature is $H = 0$, making the helicoid a minimal surface. The helicoid can be continuously deformed into a catenoid by the transformation: $x\ (u, v) = \cos \alpha \sin h\ v \sin u + \sin \alpha \cos h\ v \cos u$; $y\ (u, v) = -\cos \alpha \sin h\ v \cos u + \sin \alpha \cos h\ v \sin u$; $z\ (u, v) = u \cos \alpha = v \sin \alpha$, where $\alpha = 0$ corresponds to a helicoid and $\alpha = \pi/2$ to a catenoid.

Locally the helicoid is isometric to the catenoid. In fact, such local isometry can be achieved as endpoint of a continuous one-parameter family of isometric deformations which are all minimal surfaces. If a twisted curve C (i.e., one with torsion $\tau \neq 0$) rotates about a fixed axis A and, at the same time, is displaced parallel to A such that the speed of displacement is always proportional to the angular velocity of rotation, then C generates a generalized helicoid. (Recall that the *torsion* of a space curve is the rate of change of the curve's osculating plane; the torsion τ is positive for a right-handed curve, and negative for left-handed curve; a curve with curvature $K \neq 0$ is planar iff $\tau = 0$; the torsion can be defined by $= N \cdot B'$, where N is the unit normal vector and B is the unit binormal vector.).

A *helix*, sometimes also called a *coil*, is a curve for which the tangent makes a constant angle with a fixed line. The shortest path between two points on a cylinder (one not directly above the other) is

surface. The most important of these is the way the helicoid can be deformed applying the Bonnet transformation,[5] which preserve the mean curvature. This means that the surface remains a minimal surface, and that is of great importance to a highly solvated molecule like DNA, since the metric of the parallel surfaces remain the same. That is,

---

a fractional turn of a helix, as can be seen by cutting the cylinder along one of its sides, flattening it out, and noting that a straight line connecting the points becomes helical upon re-wrapping. Helices come in enantiomorphic left- (coils counterclockwise as it "goes away") and right-handed form (coils clockwise). Helices in a double-stranded molecule of DNA are right-handed. (Contrary to DNA molecule, many large helical structures in plants and animals are left-handed). The helix is a space curve with parametric functions $x = r\,\cos t$, $y = r\,\sin t$, $z = ct$, for $t \in [0, 2\pi]$, where r is the radius of the helix and 2π c is a constant giving the vertical separation of the helix's loops. The curvature of the helix is given by $K = r/r^2 + c^2$, and the locus of the centers of curvature of a helix is another helix. The arc length is given by $s = \sqrt{r^2 + c^2}t$. The torsion of a helix is given by $\tau = c/r^2 + c^2$, so $K/\tau = r/c$, which is constant. In fact, a necessary and sufficient condition for a curve to be a helix is that the ratio of curvature to torsion be constant. The minimal surface of a helix is a helicoid.

[5] The *Bonnet transformation* (after its inventor Pierre-Ossian Bonnet, 1819–1892) is a very special transformation which has many interesting applications in mathematics (minima surfaces, Riemann surfaces, differential geometry, etc.) and in other sciences, especially chemistry and biology. To start let us denote by k the curvature of a space curve, and by τ its torsion. We learned that at any (regular) point a space-curve can be approximated by a circle and the curvature at that point is simply the curvature of the circle that best approximates the curve at that point. The rate of recession of the circles of curvature from the original plane is the torsion of the curve. If the torsion is zero, the rate of recession from the plane is zero and the curve is planar. Let now k_1 and k_2 be the principal curvatures of a surface patch $\sigma(u, v)$. The Gaussian curvature of σ is the product of k_1 and k_2, $K = k \cdot k_2$, and its mean curvature if the half of k_1 and k_2: $H = ½(k_1 + k_2)$. K is a measure of the metric of the surface. When a surface is subjected to a transformation that leaves the metric invariant (bending without stretching), K will not change. An important characteristic of H is that it changes when the surface changes orientation. When this is forbidden from symmetry considerations, H must be identically equal to zero, and the surface is a minimal surface. In the special case of minimal surfaces, fixing H and K restricts us not to a single surface, but to a family of surfaces. The transformation from one member of the family to another is known as the Bonnet transformation. This transformation is highly restrictive. It is continuous and it preserves both the mean and the Gaussian curvature of every point on the surface. The perhaps most well-known example of a Bonnet transformation is that allowing to transform a catenoid a catenoid in a helicoid. Mathematically the Bonnet transformation can be described as the weighted sum of two minimal surfaces: $S = \cos\theta\, S'' + \sin\theta\, S'$, where θ is the Bonnet angle (parameter). The Bonnet transformation is well-defined in the sense that it can follow one path only. As a minimal surface undergoes the Bonnet transformation the individual points move along ellipses. Each point on the ellipse corresponds to a unique surface in the family. This puts the Bonnet transformation (and therefore the minimal surface) in a very special position. The sphere cannot be bent at all without changing the curvature K. This is known as the (metrical) rigidity of the sphere. The plan on the other hand has an infinite number of possible modes of isometric K-preserving deformations, something that might be called the floppiness of the plane. For a minimal surface there exists one such mode only if the minimal surface properties are to be preserved. For a minimal surface undergoing the Bonnet transformation, H and K are invariant. The Bonnet transformation preserves all local axes of rotation and roto-inversion normal to the surface. This determines what symmetries are possible for periodic minimal surfaces generated by the Bonnet transformation. Periodic minimal surfaces have a specific obstacle to overcome when undergoing the Bonnet transformation. The transformation continuous only for surfaces of genus zero, that is for disc-like surfaces. When a surface of higher genus is transformed, it must be cut open. In the case of periodic minimal surfaces, these cuts close up at the Bonnet angles where the surface is one more periodic. Some important characteristics of the Bonnet transformation are

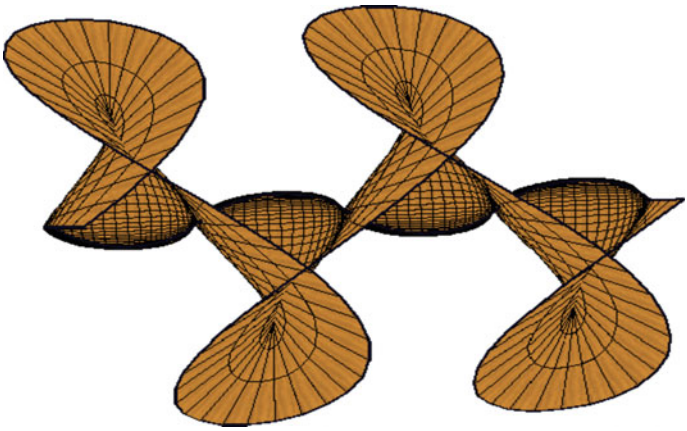


Fig. 7 A (1-periodic) Bonnet transformation

the solvation[6] shells will remain unperturbed during folding if this can be described as a Bonnet transformation. Nevertheless, globally, the difference is huge: 10,000-fold compactification is accomplished, but only by moving bulk solvent, not by changing solvation shells. Still another important property of Bonnet transformation is that it imposes simultaneity on the system. In a mathematical Bonnet transformation, all points on the surface move in unison. In the DNA molecule, which is only a good approximation of the helicoid, the partial folding of the molecule at one location will lead to the imposition of a similar structure at nearby sites (Fig. 7).

The trigger to start folding must come from outside the molecule if the Bonnet transformation mechanism is to be used. Since the local environment is virtually unchanged, the impetus must be a global effect. This is supplied by the protein's histones. By binding to sites that are distant from each other along the helicoid surface and bringing them close together in 3-space, they pull the DNA strand together. The

thus: (i) It is isometric, i.e. it preserves distances along the surface. (ii) It is the only isometric transformation of a minimal surfaces through minimal surfaces. (iii) The parallel surfaces undergo Bonnet-like transformations as well. (iv) The Bonnet transformation preserves all axes of rotation and roto-inversion normal to the surface.

To biologists, the Bonnet transformation is an attractive way to describe complex reorganizations, which influence the physiological functions of macromolecules like DNA and proteins. Many macromolecules, indeed, attain minimal surface shape, and for the kind of them that undergo dynamic structural changes, the Bonnet transformation result to be very (physiologically and evolutionary) advantageous. This notably because: there will be a well-defined, low energy path from one state to another; the transformation is isometric, and therefore no bonds are stretched; it will proceed along a well-defined path since the isometry is unique; and, perhaps most important in the case off biomolecules, the parallel surfaces undergo a Bonnet transformation as well, leaving any hydration shell virtually unperturbed. Bonnet transformation is a very interesting example of a geometric structure which optimize the biological processing and work of living matter.

[6] Solvation can be defined as a process in which a solute is transferred from a fixed position in the ideal gas to a fixed position in the liquid (solution) phase.

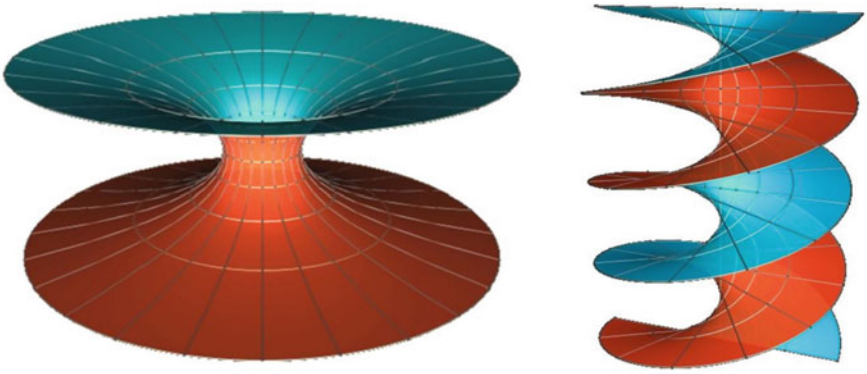


Fig. 8 A catenoid and a helicoid (From *Wolfram Mathworld*.)

mechanism is similar the coiling and super-coiling and an elastic (flexible) cord (Fig. 8).

The general scheme of the folding would then be this: the loosely curled DNA strand is swimming in a soup of intracellular fluid, containing the histones. At a critical pH, the conditions become just right for the histones to bind to the DNA strand. Once the first histone is in place, the folding will be self-catalysed, since the binding induces a Bonnet transformation upon the nearby parts of the DNA strands, creating ideal binding sites for free histones. This auto-catalysis yields an ever-accelerating process that propagates through the entire DNA strand like a sonic wave, dramatic on the global scale, but gentle on the local, thereby ensuring structural integrity of the genetic material. Considering the fact that the ensuring condensation of the chromatin fibre into the metaphase chromosome is achieved by further winding of the molecule, it is fair to assume that this follows a similar mechanism, creating a self-similar sequence, a cascade of Bonnet transformations.

The winding of the two strands of the DNA double helix about each other, represented by filled and open strings, can be measured by the linking number between the strands. It is equal to the one-half the number of the signed crossing of the two strands in any projection of the molecule. The linking number is a very important knot or link invariant [36] it is, more precisely, an invariant of oriented links. Suppose that D is an oriented regular diagram of a 2-component link $L = \{k_1, k_2\}$. Further, suppose that the crossing points of D at which the projection of k_1 and k_2 intersect are $p_1, p_2, \ldots, p_n$. (We consider only the crossing points of the projections of k_1 and k_2, which are not self-intersections of the knot component.) Then $1/2\{\text{sign}\ (p_1) + \text{sign}\ (p_2) + \ldots + (\text{sign}\ (p_n)$ is called the *linking number* of k_1 and k_2, which is usually denoted $Lk\ (k_1, k_2)$. This number has some striking properties, the most important of which is that the linking number $Lk\ (k_1, k_2)$ is an invariant of L, that is, it is the same for two or more diagrams of L. To see how the linking number works, consider how it applies to a real DNA, say the polyoma-virus DNA (a small, non-enveloped virus, which is widespread in nature, and which can cause different diseases, from cancers to disorders of the central nervous system). Remember that this DNA can be

resolved through sedimentation into three components: I and II, which are circular, and III, which is linear. It has been determined experimentally that the average linking number for a population of relaxed circular molecules of this DNA is about +500. On the other hand, for a population of closed circular molecules (the supercoiled molecules that makes up component I) the average linking number is about +475. Thus, the closed circular molecules of the polyoma-virus DNA are underwound, having a deficit in the winding number of about 25. This finding suggests a way to define supercoiling [37]. It is equal to ΔLk, that is the difference between the linking number of a molecule in the natural closed circular state and the linking number of the same molecule in the relaxed closed circular state (when the energy of deformation is at a minimum and the writhing number is zero). For example, for the DNA's both the polyoma-virus and the monkey virus SV40, ΔLk is approximately –25. It can now be understood why a deficit in the linking number of a molecule of DNA causes the molecule to supercoil. A linear molecule of DNA in solution normally assumes a form known as the *B* configuration,[7] in which the nucleotide bases are approximately perpendicular to the helical axis with 3.4 angström units between them and in which there are about 10 bases pairs for each turn of the double helix. This is a configuration of minimizing energy, and if the molecule is bent or twisted, its energy is increased. If a long molecule is simple circular, however, the diameter of the circle is large compared with the thickness of the double helix. Hence the curvature of the molecule is small and its energy is increased only slightly. As a result, nicked circular molecules such as components II of polyoma-virus DNA hardly depart from the *B* configuration. The situation is quite different for a closed circular molecule with a deficit in linking number. To satisfy the condition that the value of Lk be less than that of a relaxed molecule (say 475 rather than 500) the double helix would have to be untwisted, a transformation that would substantially increase the deformation energy of the molecule. By supercoiling, however, the closed circular molecule minimizes the amount by which it departs from *B* configuration. More precisely, as the analysis of the ribbon model revealed, one way that underwound DNA can reduce its deformation energy is by writhing. Since writhing and twisting are interconvertible, it is apparent that by changing the extent of writhing it is possible to minimize the twist of a molecule, thereby minimizing the twisting component of its deformation energy. On the other hand, writhing always introduces some curvature, and so it increases the bending contribution to the energy of the molecule. Therefore, the supercoiled configuration that the underwound DNA molecule assumes is one that minimizing twist while introducing the smallest possible amount of bending (see White [24], Summers [38], and Wolffe [39]) (Fig. 9).

The *writhe Wr* can simply be thought of in terms of the number of times the rubber rod crosses over itself. It is a measure of the shape of the DNA as a three-dimensional curve through space. One can count the of crossovers of the DNA in a single view in order to estimate *Wr*. All we need to do to get *Wr* accurately is to count the number

[7] *B*-DNA is the classical configuration of DNA with a right-handed double helix and 10 nucleotides per turn of the helix.

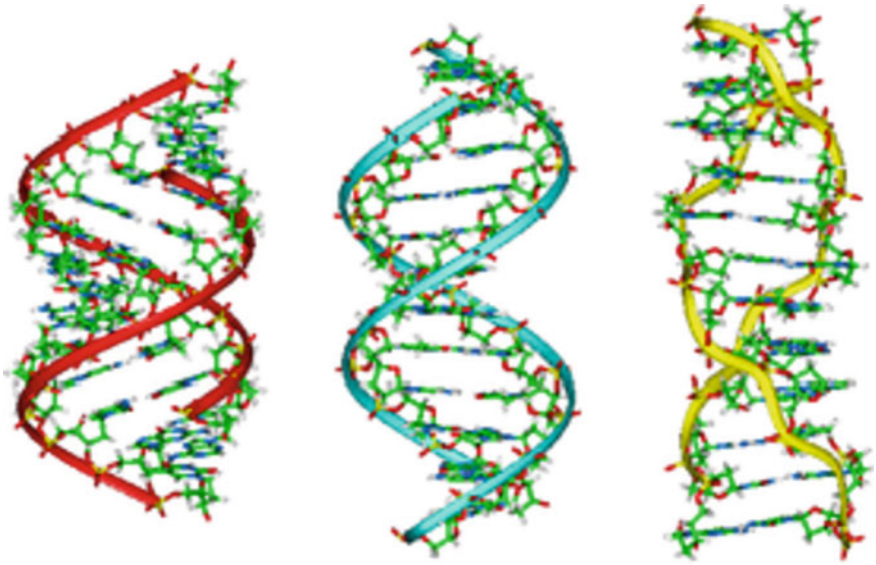


Fig. 9 The B-DNA is in the middle, on the left A-DNA, and Z-DNA on the right. Other configurations of DNA exist and have been observed

of crossovers that can been sought in many different randomly chosen views of the structure, and then take the average of all of these to get the actual value of *Wr*.

The *twisting Tw* is a measure of the turning of a vector v_{ac} (a unit vector along the line joining a to c) around the axis A. a is a unique point of the axis A and c a unique point of the backbone curve C (perpendicular to A). Since the curve C (representing a strand of DNA) winds helically about A, the vector v_{ac} turns around A, and as the point a move along A, the vector v_{ac} changes. The infinitesimal change in v_{ac}, denoted dv_{ac}, will have a component tangent to the axis and a component perpendicular to the axis. *Tw* is the measure of the total perpendicular component of the change of v_{ac} as the point a traverse the entire length of the DNA. This is given by the integral

$$Tw = 1/2\pi \int_A dv_{ac} \cdot T \times v_{ac}, \tag{13}$$

where T is the tangent vector along the curve A. When A is a straight line or planar, dv_{ac} is always perpendicular to A, so that in this case *Tw* is simply the number of time that v_{ac} winds about the axis. It can be easily demonstrated that *Tw* is positive if the winding is right-handed and negative if left-handed. Furthermore, if the DNA is closed the initial and final positions of v_{ac} are the same. Thus, if the DNA is closed (and in the circular or ribbon model) and its axis planar, *Tw* must necessarily be an integer.

6 Topological Changes and Emergence of Physical Patterns

In fluid mechanics a very interesting phenomenon is that of topological changes, which do occur when dissipative effects become predominant over the coherency of structures. When this happens there is a dramatic change of fluid patterns, often on small time-scales compared to evolution. The change occurs through the formation and disappearance of physical reconnections in the fluid pattern. In real fluids, for example, vortex and magnetic tubes do interact and reconnect freely. From a dynamical system viewpoint Smale [40], Thom [41, 42]), reconnections take place when the vector field lines (streamlines, vortex lines or magnetic lines) cross each other. If two field lines meet, the point of crossing is a true nodal point, like a bifurcation in a path. Dissipative effects allow the reconnection to proceed through such points (for more details we refer to Hornig [5]).

Analytical and numerical studies of flow patterns show that bifurcations of the field lines occur when configurations are degenerate, as with interfacial flows in the vicinity of a solid boundary. When these events dominate the physics, we can still use an alliance of topological, probabilistic and combinatorial techniques to predict average properties and long-term evolution (see Ricca and Moffat [13], Boi [14, 43]).

As local processes, reconnections are difficult to describe and are still a puzzle for theorists. One seemingly helpful mathematical approach, which must be mediated by detailed knowledge of the particular physical process, involves techniques of "oriented surgery", performed on the bundle of constitutive vector field lines.

Recall that, mathematically, *surgery*[8] is a procedure for changing one manifold into another (of the same dimension n) by excising a copy of $S^r \times D^{n-r}$ for some r, and replacing it by $D^{r+1} \times S^{n-r-1}$, which has the same boundary, $S^r \times S^{n-r-1}$. In general terms, the surgery problem can be stated as follows. Let X be a topological space. When is X homotopy equivalent to a closed manifold? We first need to consider a map of degree one $f: M \to X$ from a closed manifold M to a finite Poincaré complex X covered by some bundle data. The surgery problem is to change it to a homotopy equivalence without losing the structure of a closed manifold. We also need to state the following theorem.

Theorem. Let X be a connected finite n-dimensional Poincaré complex. Let f: $TM \oplus \mathbf{R}^a \to \xi$ be a normal map of degree one covering f: $M \to X$. Then we can

[8] *Surgery theory* is one of the most interesting method in differential topology and algebraic geometry, introduced in the 1960s by Browder, Novikov and Wall. It investigates the homotopy type of manifolds using a combination of topology and algebra. It has been used, initially by Milnor, to generate a finite-dimensional manifold from another in a well-ordered way. Originally developed for differentiable (or smooth) manifolds, surgery techniques also apply to piecewise linear (PL-) and topological manifolds. Surgery refers to cutting out parts of the manifold and replacing it with a part of another manifold, matching up along the cut or boundary. Surgery theory is deeply connected with other important mathematical methods and techniques, notably with s-cobordism theorem, h-cobordism theorems, Whitehead torsion, Dehn surgery, manifold handlebody decomposition knot theory and the Perelman program on geometrization conjecture (deletion of singularities developing in simply connected oriented (compact) 3-manifolds). (See A. Ranicki, *High-Dimensional Knot Theory—Algebraic Surgery in Codimension 2*, Springer-Verlag, Berlin Heidelberg, 1998.

carry out a finite sequence of surgery steps to obtain a normal map of degree one g: $TN \oplus \mathbf{R}^{a+b} \to \xi \oplus \mathbf{R}^b$ covering g: $N \to X$ such that (f, f) and (g, g) are normally bordant and g is k-connected, where $n = 2k$ or $n = 2k + 1$.

Then the final version of the surgery problem can be stated as follow.

Suppose we have some normal map (f, f) from a closed manifold M to a finite Poincaré complex X. Can we change M and f leaving X fixed by finitely many surgery steps to get a normal map (g, g) from a closed manifold N to X such that g is a homotopy equivalence?

Another important result which can be useful for what follows concern knots. By definition, a *knot* is a smoothly embedded circle in $\mathbb{R}^3$ or S^3. The union of finitely many disjoint knots is a *link*. By removing a thickened link in S^3 (a union of solid tori) and gluing it back with a twist, one obtains a new 3-manifold (space) M. An important result of W. B. R. Likorish (1962) showed that all closed, orientable, connected 3-manifolds can be obtained by performing Dehn surgery on links in S^3 (see Rolfsen [44] and Lescop [45]).

Let's briefly explain the notion of Dehn surgery [46], which is of great importance in the theory of 3-manifolds. Given a 3-manifold M and a knot $K \subset M$, one removes a solid torus neighborhood of K from M and sews this solid torus back into the resulting complement using self-diffeomorphism of the boundary 2-torus. These self-diffeomorphisms up to isotopy are identified with $SL(2, \mathbf{Z})$ via the action of the diffeomorphism on the first homology H^1 of the 2-torus T^2. Dehn then constructed manifolds by this method, starting a $(2, q)$-torus knot in S^3—these are knots lying on the surface of the standard torus and wrapping that torus linearly twice in one direction and q times in the other (q must be odd). Dehn identified which of the non-identity surgeries on $(2, q)$-torus knots generate manifolds with the homology of the 3-sphere and showed that they all have non-trivial fundamental groups (usually infinite). When the knot is the $(2, 3)$-torus knot there is surgery that generate a manifold with the same group as Poincare's example (later proved to be diffeomorphic to Poincaré example). The example of Dehn surgery on $(2, q)$-torus knot was much better understood after the work of Seifert [47] and Seifert and Threlfall [48]. They considered three-manifolds that admit locally free circle actions, now called *Seifert fibrations*. They showed that all the examples coming from Dehn surgery on $(2, q)$-torus knots were such manifolds and they showed how to compute the fundamental group of these manifolds. In particular, Dehn's extraordinary example was shown to have the same fundamental group as Poincaré's original example—it is the pre-image in $SU(2)$ of the symmetries of the regular icosahedron.

The remarkable mathematical fact is that Dehn surgery gives a method for constructing one manifold from another, especially one 3-manifold from another, by a kind of cut-and-paste procedure. This is also philosophically much relevant because it allows to thinking space as a generated process, as a dynamical entity which is not fixed or given a priori, rather as a result of certain operations and transformations. The question then is: How can we from some space find those deformations that applying to it give us the possibility to generate new spaces? This displaces significantly the standard philosophical question, which was to ascertain whether the space

is already comprised in our mental schemes or not, and which consists even now in responding to the question whether space (and time) exists or not.

A standard application of Dehn is surgery along a link L in a 3-sphere S^3. This works in two steps:

1. Form the complement in the 3-sphere of a tubular neighborhood of the embedded link $L \to S^3$, of the form $L \times \text{int}(D^2)$. This is called *Dehn drilling*. The result is a 3-manifold with boundary M; whose boundary $\partial M \cong L \times S^1$ can be viewed as the boundary of a disjoint union of solid torr $L \times D^2$.
2. To each of the connected components $C_1, \ldots, C_n$ of ∂M, apply an (orientation-preserving) homeomorphism, say $\phi_1, \ldots, \phi_n$. The union $\phi_1 \cup \ldots \cup \phi_n$ is a homeomorphism ϕ: $\partial M \to \partial M$. Then perform a Dehn filling by constructing the pushout of an evident span:

$$\begin{array}{ccc} \partial M & \to & \partial M, \\ \downarrow & & \downarrow \\ M & & L \times D^2, \end{array}$$

thus, refilling the drilled portion, but in a new way (along ϕ). This gives a new 3-manifold N.

An important result state that

Theorem (Likorish-Wallace) *Every connected oriented closed 3-manifold N arises by performing an integral Dehn surgery along a link $L \to S^3$ (i.e., surgery along a framed link).*

Now return to the phenomenon of topological reconnection. Vector field lines in the fluid flow are oriented curves, whose arrows give the direction of the field they represent. A physical knot can be seen as a knotted tubular bundle made of oriented curves. When two strands of the bundle come into contact, vector lines of one strand may recombine with vector lines of the other by a "cut and connect" process, which preserves orientation.

The process of surgery shows clearly how a local (geometrical) event, as a bifurcation, a breaking symmetry or other kinds of singularities, may have a global (topological) effect generating an overall reorganisation of the system. When this happens, we have a complex change of topology of the system. Recent studies (see Hornig [5], Boyland [4]) show that the efficiency of reconnection seems to be strongly influenced by local geometric properties given, for example, by the relative inclination of the tube strands. Orientation-preserving surgery and efficiency of the process are therefore two important features for topological diagnosis of fluid flows.

7 Complexity of Time and of the Universe

The reversibility of time in physical elementary process—both in classical and in quantum mechanics, as well as in the relativistic theories—is commonly accepted and very well established; that means that the fundamental laws of physics are invariant under time reversal. However, it is an obvious fact that most phenomena in Nature distinguish a direction of time; time is irreversible in complex systems (see Albeverio and Blanchard [49]). Electromagnetic waves are observed in their retarded form only, where the fields causally follow from their sources. The increase of entropy, as expressed in the second law of thermodynamics, also defines a time direction. This is directly connected with the psychological arrow of time—we remember the past but not the future. In quantum mechanics it is the irreversible measurement process and in cosmology the expansion of the Universe, as well as the local growing of inhomogeneities, which determine a direction of time.

In complex systems consisting of many particles or other elements, disorder (chaos) inevitably increases as a result of the random nature of numerous interactions (see Devaney [50]. Entropy is that very measure of the degree of chaos. It is very important that when creating a more ordered state in a system, by influencing it from within a larger system, we inevitably insert additional disorder into this larger system. The laws of thermodynamics state that the "chaos" added to the larger system is inevitably greater than the 'order' introduced into the smaller system. Hence the "chaos", and "entropy", in the whole world must grow, even though order may be established in some parts of the world. One realizes then that the second law of thermodynamics is of great importance for the evolution of the Universe. Indeed, exchange of energy between the world and "other systems" being impossible, the Universe must be treated as an isolated system. Therefore, all types of energy in the Universe must ultimately convert to heat spread uniformly through matter, after which all macroscopic motion peters out. Even though the law of conservation of energy is not violated, the energy does not disappear and remains in the form of heat, it 'loses all forces', any possibility of transformation, any possibility of doing the work of motion. This bleak state became known as the 'thermal death' of the Universe. The irreversible process in the Universe is thus the growth of entropy. The question however remains open: can this process entirely dictate the direction of flow of time? I guess that we shall search for some other key feature of time and of space–time if we want be able to give a satisfactory answer to that question.

8 The Complexity and the Breaking-Symmetry Phenomenon

It must be pointed out that one of the most important upheaval in the scientific vision of nature in the last century has been the discovering that spontaneous breaking

symmetries, bifurcations and singularities are three mechanisms which play a fundamental role for the organisation of physical and living matter and the unfolding of natural phenomena (see Boi [51, 52]). These mechanisms are very deeply related, because each time that a physical or living system bifurcate, the immediate consequence is that the symmetry of the system breaks down and instead of that a new broader symmetry will appear. Besides, the fact that a system may bifurcate at some moment of his evolution means that its unfolding stops to be (mathematically speaking) continuous or linear and become discontinuous and non-linear. In many situations, this non-linearity (of partial differential equations) lead to the emergence of new order–disorder transition phenomena which exhibits non-equilibrium states which mathematically can be represented by time dependent equations, and these states are a source of instability, bifurcation and symmetry breaking phenomena (see Prigogine and Nicolis [53]). Many of these macroscopic and local dynamical laws and phenomena manifest time-asymmetry or irreversibility, which is a feature of key significance.

Let me first mention some examples and fields in which spontaneous symmetry-breaking manifests itself as a primary feature of the problem (we refer to Jost [11] for more details).

Morphogenesis and molecular biology. A striking example of symmetry breaking in a biological system is the breakdown of rotational symmetry in the *Fucus* seaweed egg. At a critical stage in the development of the egg a transition is made from a spherically symmetric membrane potential distribution to a polarized state with an axial symmetry, and a net trans-cellular current leaving one pole and entering the opposite.

Wave propagation in neural networks. Bifurcation phenomena in simple mathematical models of excitatory inhibitory neural networks have been discussed recently by many peoples. Neural networks are aggregates of nerve cells which interact with other neurons in the network in either an excitatory or inhibitory way, and so it is plausible to expect these networks to exhibit such non-linear collective phenomena as bifurcation, threshold effects, and hysteresis. One can model these networks by a system of equations

$$\mu Y = -Y + S(KY + P)$$

where Y is a two-component vector, S is a non-linear vector-valued function, K is a linear convolution operator, and P is the external stimulus.

Phase transitions in statistical mechanics. The notion of symmetry breaking is fundamental to phase transitions, yet much harder to treat mathematically. Until the renormalization theories developed in the last four decades, the primary approach to phase transition was, in one way or another, a mean-field approximation coupled with a bifurcation analysis of the mean-field equations.

9 The Specificity of Biological Complexity, Emergence and Causation

The important fact that the specificity of complex biological activity does not arise from the specificity of the individual molecules that are involved, as these components frequently function in many different processes. For instance, genes that affect memory formation in the fruit fly encode proteins in the cyclic AMP (camp) signaling pathway that are not specific to memory. It is the particular cellular compartment and environment in which a second messenger, such as camp, is released that allow a gene product to have a unique effect. Biological specificity results from the way in which these components assemble and function together. Interactions between the parts, as well as influences from the environment, give rise to new features, such as network behaviour which are absent in the isolated components (see Cornish-Bowden and Cárdenas [54], Liljenström [55]).

More precisely, it can be said that complex biological levels of functionality result from self-organized processes. For self-organization to act on macroscopic cellular structures, three requirements must be fulfilled: (i) a cellular structure must be dynamic; (ii) material must be continuously exchanged; and (iii) an overall stable configuration must be generated from dynamic components. Interactions between the parts, as well as influences from the environment, give rise to new features, such as network and collective behaviors which are absent in the isolated components (see Misteli [56], Karsenti [57]).

Consequently 'emergence' has appeared as a new concept that complements 'reduction' when reduction fails [58]. Emergent properties resist any attempt at being predicted or deduced by explicitly calculation or any other means. In this regard, *emergent* properties differ from *resultant* properties, which can be defined from low-level configurations and information. An important aspect of emergent properties is that *they have their own causal power*, which is not reducible to the power of their constituents (for a thorough treatment of this topic see Boi [59]).

The key concepts here are those of 'organization' and 'regulation', first of all because organization and regulation become cause in the living matter of morphological, functional and mental novelties [60]. According to the principle of emergence, the natural and living worlds are organized into stages and levels that have evolved over different evolutionary times through continuous and discontinuous processes.

Reductionists advocate the idea of 'upward causation' by which molecular states generally bring about higher-level phenomena, whereas proponents of emergence admit 'downward causation' by which higher-level systems may influence lower-level configurations. The importance of admitting 'downward causation' in the analysis of complex living systems (i.e. presenting and ever-increasing coupled activity of plasticity and complexity) appear clearly in epigenetics phenomena, for example in the fact that chromatin forms and its structural modifications play a crucial role in the increasing complexity of gene regulatory networks [61], in the emergence of cellular functions and in development, as well as in the neurocognitive plasticity.

10 Systems Biology and Complexity

The principal challenge facing systems biology is complexity. Systems biology defines and analyses the interrelationships of all the elements in a functioning system in order to understand how the systems works. At the core of the challenge is the need for a new approach, a shift from reductionism to an integrative perspective. More precisely, what is needed is to provide a conceptual framework for system biology research. The concept of a complex system, i.e. a system of subsystems each belonging to a certain category of living entities such as proteins, tissues, organs, etc., need first to be defined in general mathematical terms.

It is rather clear, however, that for a deeper understanding in systems biology investigations should go beyond building numerical mathematical or computer models—important as they are. Biological phenomena cannot be predicted with the level of numerical precision as in classical physics. Explanations in terms of how the categories of systems are organized to function in ever changing conditions are more revealing. Non-numerical mathematical tools are appropriate for the task. Such a categorical perspective led us to propose that the core of understanding in systems biology depends on the search for organizing principles than solely on construction of predictive descriptions (i.e. models) that exactly outline the evolution of systems in space and time.

Biological systems are difficult to study because they are complex in several ways. One of the most important aspects of biological complexity is multi-levelness: the structural and functional organization of the human body into tissues and organs systems composed of cells. From molecules to organs, levels are inter-related and interdependent, so that the organism is able to conserve and adopt the integrity of its structural and functional organization against a setting of continuous changes within the organism and its environment. This capacity, usually described as ‘robustness’, is a consequence of non-linear spatial–temporal intra- and inter-cellular interactions.

11 Self-Organization and the Causal Role Played by Systemic Properties in Biology

Self-organization, that is the capacity of any complex living organism to intrinsically produce new properties and behaviors of organization and regulation, cannot be addressed by purely reductionist approaches (For a comprehensive treatment of this subject see Kauffman [62], Nicolis and Prigogine [63], and Boi [30, 52]). Living organisms present the following two fundamental features. (1) They are thermodynamically open systems, that is, they are in a state of permanent flux, continuously exchanging energy and matter with their environment. (2) They are characterized by a complex organization, which results from a vast network of molecular and cellular interactions involving a high degree of nonlinearity. Under appropriate conditions,

the combination of these two features, *openness* and *nonlinearity*, enables complex systems to exhibit properties that are *emergent* or *self-organizing*.

In biological systems, such properties may express themselves through the spontaneous formation, from (almost) random molecular interactions, of long-range correlated, macroscopic dynamical patterns in space and time—the process of *self-organization*. The dynamical states that result from self-organizing processes may have features such as excitability, bi-stability, periodicity, chaos or spatial–temporal patterns formation, and all of these can be observed in biological systems.

To show the effective causal role played by systemic properties in biology, let us consider the three following examples

(1) The first concerns the many cases of cooperative feedback inhibition of metabolic pathways, which are now well-known, such as the inhibition of aspartokinase in bacteria by lysine. This type of observation is often explained by supposing that the biosynthetic flux is regulated by this feedback inhibition, and would be subject to uncontrolled variations if there were no feedback loop. However, this explanation is wrong, because fluxes can be controlled perfectly well without feedback inhibition, whether cooperative or not. The need comes not from flux control but concentration control: without feedback inhibition in this pathway the rate at which lysine would be synthetized would still match the rate at which it is used in protein synthesis, but there would be huge and potentially damaging variations in the concentration of lysine and the intermediates in the pathway from aspartate. This sensitivity of metabolite concentrations to perturbations has major implications for the regulatory design of metabolism in living organisms.

(2) The second example concerns the failure of genome sequencing to provide an effective explanation of how living organisms develop and evolve. There are at least two fundamental reasons for this failure. (i) The first is related to the essential fact that the expression of genome, i.e. its state of activity, stands beyond the gene sequences, and depends much more upon the peculiar spatial organization of the genome into the chromatin and the chromosome [64]. Moreover, the functional properties of genomes are strongly determined by their cellular organization. The functional relevance of spatial and temporal genome organization must be stressed at three interdependent levels: the organization of nuclear processes; the organization of chromatin into higher-order domains; and the spatial arrangement of chromosomes and genes within the nuclear space.

    Each of these levels has regulatory potential, and all are interdependent. There is an increasing evidence that the higher-order, topological organization of the genomes exert a fundamental influence on their functional properties, and on many cellular processes, including expression and genome stability [35, 65].

(3) The third example concerns the relationship between genotype and phenotype. We know that for more than half a century the prevalent ‘dogma’ was to think that the genotype completely and unidirectionally determine the phenotype and

hence the fate of any complex living organism. Now, to be more precise, the problem is not much that genome sequences contain no phenotypic information, but that we do not have reliable methods for undertaking all of the steps involved in deducing a phenotype from them. "A list of putative gene products, or even a list of putative enzymes, is not a phenotype, and converting it into a phenotype requires construction of plausible metabolic map, which then need further work to convert it into a possible phenotype. Finally, the possible phenotype can only become a real phenotype when all relevant kinetic and regulatory properties are taken into account, together with information about how all the components are organized into a three-dimensional whole—even a four-dimensional whole, given that the times when different components are made may be just as important as where they are placed" [66].

12 Systems Biology, Reductionism and Emergence

To understand disease-relevant processes, we therefore require methodologies that allow us to study non-linear spatial–temporal systems with multiple interconnected levels of structural and functional organization. Non-linear dynamics plays an important role for the explanation of highly non-linear biological behaviors such as biochemical and cellular rhythms or oscillations (see Liljenström [55]). According to biodynamics, biological systems are seen as open systems of non-linearly interacting elements. Consequently, the field of biodynamics might be defined as the study of the complex web of non-linear dynamical interactions between and among molecules, cells and tissues, which give rise to the emergent functions of a biological system as a whole. The work of non-linear dynamical interactions favors the self-organization of emergent macroscopic patterns, including temporal oscillations and spatial–temporal wave patterns, especially in chemical and biological systems. Numerous examples are now known at all levels of biological organization. The formation of biological rhythms and oscillatory dynamical states of different periodicities plays a fundamental role in living organisms.

Biological complexity and specificity results from the way in which single components like molecules, genes and cells self-organize and function together when constituting a whole (a tissue, an organ, an organism), say a whole system including different subsystems. Not only the interactions between the parts and the influence from the environment (think of epigenetic factors, both chemical and spatial, that mediate the complex relationship between the genomes and the micro- and macro biophysical environments), but also the systemic properties of the whole that exert an action on the components, give rise to new features, such as network behavior and functional properties, which are absent in the isolated components.

Self-organizing processes may give rise to new, unexpected properties and behaviors in living systems, also called *emergent properties*. Emergent properties can be defined as properties that are possessed, either triggered or constrained, by a dynamical system as a whole but not by its constituent parts. Otherwise stated, emergent

phenomena are phenomena that are expressed at higher levels of organization in the system but not at the lower levels. The concept of self-organization implies the existence of a dynamical interdependence between the molecular and cellular interactions at the microscopic level and the emerging global structure at the macroscopic organismic level. (See Karsenti [57], Misteli [56].)

We need to consider 'emergence' as an effective new concept that complements 'reduction' when reduction fails, and allow to consider those specific systemic properties of the whole responsible for biological organization and regulation at higher levels. Emergent properties do not result from properties pertaining to simple (and isolated) components of biological systems. They resist any attempt at being predicated or deduced by explicitly calculation or any other analytical means. In this regard, emergent properties differ from 'resultant' properties, which can be predicted from lower-level components. "For instance, the resultant mass of a multi-component protein assembly is simply equal to the sum of the masses of each individual component. However, the way in which we taste the saltiness of sodium chloride is not reducible to the properties of sodium and chlorine gas. *An important aspect of emergent properties is that they have their own causal power, which is not reducible to the powers of their constituents*. For instance, the experience of pain can alter human behavior, but the lower-level chemical reactions in the neurons that are involved in the perception of pain are not the cause of the altered behavior, as the pain itself has a causal efficacy." (see Van Regenmortel [67]). Advocating the reductionist idea of 'upward causation' means to maintain that molecular components and states suffice to determine higher-level processes occurring in biological systems.

We think that, in fact, there is an active combination of upward and downward processes. The upward process indicates that, under non-equilibrium constraints, molecular interactions tend to spontaneously synchronize their behavior, which initiates the beginning of a collective, macroscopically ordered state. At the same time, the downward process indicates that the newly forming macroscopic state acts upon the microscopic interactions to force further synchronizations. Through the continuing, *energy-driven* interplay between microscopic and macroscopic processes, the emergent, self-organizing structure is then stabilized and actively maintained.

This important issue has been studied in depth by Ilya Prigogine and his collaborators [63, 53]. In particular, they showed the importance of fluctuation in the emerging of instabilities. The general insight is the recognition that a deterministic causal description cannot be adequate for large systems involving a macroscopic number of degrees of freedom. The main raison is that the very existence of many degrees of freedom implies necessarily (and spontaneously) the appearance of *fluctuations*, i.e., of spontaneous deviations from some average macroscopic behavior. In most cases, the appearance of fluctuations may be treated as a random event obeying probabilistic laws. On the other hand, once the fluctuation is produced, the systems respond according to definite macroscopic laws. In general, one expects that fluctuations although measurable, should remain small compared to the macroscopic average. As Nicolis and Prigogine underlined (1973, 91):

> An important result of equilibrium statistical mechanics shows that this cannot be so in the neighborhood of critical points of phase transition (liquid-vapor transition, etc.). In these cases, small thermal fluctuations are amplified, attain a macroscopic level and drive the system to a new phase. The evolution of the new phase occurs as an abrupt change *beyond an instability point* of the reference state. In the critical region around the instability the system exhibits a markedly coherent behavior which is frequently combined to an increase of spatial order. Several examples of similar phenomena in systems maintained far from thermal equilibrium are also known. In all these phenomena the important feature is that there exists a critical point of transition (threshold) in the neighborhood of which the least macroscopic fluctuation is amplified and drives the system to the new state.

13 Chaos, Fractality and Complexity

Chaos and complexity are often deeply related, in the sense that chaotic phenomena can generate some complexity and complex systems may exhibit chaotic behaviours. Another important generator of complexity is fractal geometry [68]. The structures entailed or generated by fractal geometry are wat give order to chaos. This relationship can best be seen in the *Mandelbrot set*, which can be described as a complex and beautiful mathematical object. Its most interesting characteristic is that it has and extraordinarily efficiently organized storehouse of images, and as such is the example for excellence of order in chaos. In essence, chaos theory and fractal geometry radically question our understanding of equilibria, and therefore of harmony and order, in nature as well as in other contexts. They offer a new integral model which can encompass a part of the true complexity of nature for the first time. It is highly likely that the new methods and concepts will allow us, for example, a much more adequate understanding of ecology and climatic changes, and thus they could contribute to our more tackling our enormous global problems. More precisely, the Mandelbrot set visually constitute a picture of a complex number c for which the images of the point c under iteration of the map $z \to z^2 + c$ never heads of to ∞. Very deep geometrical structures and transformations underlies these objects, which are intimately related to the Kleinian modular functions and groups, the Schottky group and the Julia set. All these mathematical objects are concerned with the exploration of a family of unusually symmetrical shapes, which arise when two spiral motions of a very special kind are allowed to interact. These shapes display intricate "fractal" complexity on every scale from very large to very small. Sometimes the interaction of two spiral motions is quite regular and harmonious, sometimes it is total disorder and sometimes (and this is the most intriguing case) it has layer upon layer of structure teetering on the very brink of chaos. In a more precise mathematical language, these spiral motions are maps of the complex plane to itself that can be expressed by simple formulas like $T(z) = z^2$ and $T(z) = \sqrt{z}$.

Let us stress the following essential theoretical points about chaotic (non-linear) systems and fractal geometry.

(i) Simple deterministic systems with only a few elements can generate random behaviour, and that randomness is fundamental; gathering more information

does not make it disappear. This fundamental randomness has come to be called chaos. On the other hand, as R. M. May pointed out [69], mathematically simple nonlinear systems governed by simple sets of differential equations (for example, in population biology) may generate very complicated dynamics and hence a huge growth of their intrinsic complexity [70].

(ii) Deterministic chaos entail, as one of its most significant formulation, that chaos occurs when the error propagation seen as a signal in a time process, grows to the same size or scale as the original signal. To speak of deterministic chaos is thus only apparently a paradox. In principle, the future is completely determined by the past; but, in fact, small uncertainties, much like minute errors of measurement which enter into calculations, are amplified, with the effect that even though the behaviour is predictable in the short term, it is unpredictable over the long term. The discovery of such behaviour is one of the important achievements of chaos theory.

(iii) Deterministic chaos implies that the strong causality principle is incorrect because its presumptions are erroneous and the conclusions are also wrong. Natural laws do not exclude the possibility of chaos; in other words, determinism and predictability are not equivalent. And an even more surprising finding of recent chaos theory has been the discovery that these effects are observable in many very simple feedback systems. Moreover, chaos and order can be observed simultaneously in the same system. There may be a linear progression of errors characterizing a deterministic system which is governed by the causality principle, while (in the same system) there can also be an exponential progression of errors (i.e., the butterfly effect) indicating that the causality principle breaks down.

(iv) A most important point is that a definition of chaos entails the three fundamental properties: mixing and transitivity of mixing, dense periodic points, and sensitive dependence on initial conditions (for a detailed presentation see Devaney [50]). There is a very important interdependence of these three properties of chaos. Consider now the question of the inheritance of chaotic mappings, namely: given that a mapping f is chaotic and that g is related to f, can we conclude that g is chaotic as well? There are, indeed, several examples of mappings which are chaotic and which are related to each other. The notion of being related is topological in essence, and can be properly explained by the notions of *topological conjugacy* and *topological semiconjugacy* (for a detailed description of this notion see Boi [25]).

(v) Fractal geometry and the associated notion of dimension is intimately linked to the question of measurement, and it showed that curves, surfaces, and volumes can be complex and performed at different levels that certain ordinary measurements become meaningless. However, there is a way to measure de degree of complexity by evaluating how fast length, or surface, or volume increases if we measure with respect to smaller and smaller scales. The fundamental idea is to assume that the two quantities are related by a law, which allows us to compute one quantity from the other. The kind of law which seems to be relevant is a power law of the form $y \propto x^d$. Such a law also turns

> out to be very useful for the discussion of *dimension* (in fractal geometry). Dimension is not easy to understand; at the turn of the nineteenth century it was one of the major problems in mathematics to determine what dimension means and which properties it has. Of course, for familiar geometric objects the answer to the question "What is the 'dimension' of a set of points?" is clear: lines and smooth curves are one-dimensional, planes and smooth surfaces are two-dimensional, solid are three-dimensional, and so on.

Roughly, a definition could be that *the dimension is the minimum number of coordinates needed to describe every point in the set.* For instance, a smooth curve is one-dimensional because every point on it is determined by one number, the arc length from some fixed reference point on the curve. Astonishingly, in the last decades of last century it turns out that some of the many different notions of dimension: topological dimension, Hausdorff dimension, fractal dimension, self-similarity dimension, capacity dimension, information dimension, Euclidean dimension, are in fact all related. Some of them, however, make sense in certain situations, but not at all in others, where alternative definitions are more helpful. Sometimes they all make sense and are the same. Sometimes several notions make sense but do not agree.

The self-similarity dimension is actually a special form of Mandelbrot's fractal dimension, which in turn was motivated by the Hausdorff fundamental work [71]. Recall that *the Hausdorff dimension of an object is a numerical measure of its complexity.* It interpolates our usual idea of dimension so as to give an idea of how convoluted and crinkled or fragmented the object is at fine scales.

14 The Difference Between Self-Assembly and Self-Organization

It is unknown what determines the different shapes and sizes of cellular organelles, why specific structures form in particular places, and how cellular architecture is affected by function and vice versa. Two fundamentally different mechanisms exist to generate macromolecular structures: self-assembly and self-organization (see Lehn [72, 73]). Whereas self-assembly involves the physical association of molecules into an equilibrium structure, self-organization involves the physical interaction of molecules in a steady-state structure. The processes of self-organization have been extensively studied by many scientists and notably by Jean-Marie Lehn, who showed the importance of the transition from molecular information to self-organized functions in complex matter. In (2004, 250), he stressed some important points:

> A step beyond preorganization (of supramolecular chemistry) consists in the design of systems undergoing self-organization, i.e. systems capable of spontaneously generating well-defined supramolecular architectures by self-assembly from their components. Self-organization processes, directed by the molecular information stored in the components and read out at the supramolecular level through specific interactions, represent the operation of programmed chemical systems. (…) Self-organization processes also give access to advanced supramolecular materials, such as supramolecular polymers and liquid crystals,

> and provide an original approach to nanoscience and nanotechnology. In particular, the spontaneous but controlled generation of well-defined, functional supramolecular architectures of nanometric size through self-organization represents a means of performing programmed engineering and processing of nanomaterials. (…) Supramolecular chemistry is intrinsically a dynamic chemistry, in view of the lability of the interactions connecting the molecular components of a supramolecular entity and the resulting ability of supramolecular species to exchange their constituents. (…) The merging of the features, information and programmability, dynamics and reversibility, constitution and structural diversity, points towards the emergence of adaptative and evolutionary chemistry. Together with the corresponding fields of physics and biology, it constitutes a science of informed matter, of organized, adaptative complex matter.

In biology self-organization play a crucial role and we have many fundamental examples of self-organized processes. Concerning the difference between self-assembly and self-organization, for example, virus and phage proteins self-assemble to true equilibrium and form stable, static structures. In contrast, most cellular structures (i.e. the cytoskeleton, nuclear sub-compartments, or exocytic and endocytic compartments) are open for exchange of energy and matter and are governed by steady-state dynamics. The concept of self-organization is based on observations of chemical reactions far from equilibrium, and it is well established in chemistry, physics, and ecology. Self-organization in the context of cell biology can be defined as the capacity of a macromolecular complex or organelle to determine its own structure based on the functional interactions of its components [74]. In a self-organizing system, the interactions of its molecular parts (and not the molecular parts themselves) determine its architectural and functional features. The processes that occur within a self-organized structure are not underpinned by a rigid architectural framework; rather, they determine its organization. For self-organization to act on macroscopic cellular structures, three requirements must be fulfilled:

(i) a cellular structure must be dynamic,
(ii) material must be continuously exchanged, and
(iii) an overall stable configuration must be generated from dynamic components. Observations from recent advances in live cell imaging indicate that many cellular structures fulfill the requirements for self-organization. I suggest to think of self-organization as a more general mechanism for the formation, maintenance, and function of cellular organization than self-assembly.

Let us now consider the relationship between self-organization and the property of plasticity in living systems. Macroscopic cellular structures are characterized by two apparently contradictory properties. On one hand, they must be architecturally stable, on the other hand they must be flexible and prepared for change. Self-organization ensures structural stability without loss of plasticity. Fluctuations in the interaction properties of its components do not have deleterious effects on the structure as a whole. This means also that some form of robustness is conserved by the system. However, global and persistent changes rapidly result in morphological and functional changes. The basis of the responsiveness of self-organized structures is the transient nature of the interactions among their components. The dynamic interplay of components generates frequent windows of opportunity during which proteins

can change their interaction partners or be modified. The effective availability of components is controlled by posttranslational modifications via signal transduction pathways. Why does a cell not simply build stable, static structures?

The absence of gradual intermediates in the reorganization of cellular structures is consistent with self-organization, as self-organizing systems are frequently in a state of criticality; that is, a point at which system properties can change suddenly. Self-organization is an elegant, efficient way to organize complex structures. The properties that determine the organization are the intrinsic properties of the structure's components. For example, in protein polymers, the protein–protein interaction properties determine the architecture; in membrane structures, the flow of membrane determines the architecture. Self-organized structures do not require external mechanisms to establish, maintain, and regulate their architecture.

Thus, self-organization is a simple but effective way to optimally organize cellular structures. The study of the dynamic behavior of dynamic cellular structures cannot be described accurately by conventional equilibrium dynamics or by static observations. To understand the behavior of dynamic systems, the kinetic characteristics of their components must be known. In contrast to the study of molecular mechanism, it is not sufficient to understand in detail the behavior of single molecules; rather, the rules that govern the collective behavior of systems must be uncovered. In contrast to the mechanism of self-assembly which involves the physical association of molecules into an equilibrium structure (for example, virus and phage proteins self-assemble to true equilibrium and form stable, static structures), the concept of self-organization is based on observations of chemical reactions far from equilibrium, and the associated processes involve the physical interactions of molecules in a steady-state structure. It is well established in chemistry, physics, and ecology.

15 Complexity of Biological Systems: The Genome as a Complex System

The study of biological systems involves the qualitative-quantitative and simultaneous integration of different and multiple biological components and their relationships with one another. For example, the components may be proteins, while their relationships may be described by signal transduction pathways. The cellular processing is a complex-dynamic system with hundreds of thousands of bio-molecules interacting with one another to perform life's many functions. To fully understand the multi-layers information and organization "program" of life, a comprehensive description of protein–protein, cell–cell, cell-organism and organism-environment interactions is required. Understanding how genes ad their proteins products and cells and their intra- and extra-interactions generate the complexity and diversity that we know as life is perhaps one of the greatest challenge new millennium.

Recent theoretical studies and a huge amount of experimental data point forward the need for a profound change in our way of thinking biological phenomena, and their modelling. Let's summarize some important findings:

In the last two decades it has become more and more clear that the linear sequence map of human genome is an incomplete description of our genetic information and processing. This is because information on genome functions and gene regulation is also encoded in the way DNA sequence folded up with proteins into chromosome within nucleus [75]. This allowed for the conclusion that *the biological information on living organisms cannot be portrayed in the DNA sequence alone*. In a post-genomic, (epigenomic or/and proteomic?) era, the importance of chromatin-chromosome/epigenetic interface has become increasingly apparent.

The genome of eukaryotes is a highly complex system, which is regulated at (at least) four major (hierarchical and network-like) levels: (i) the DNA sequence level, (ii) the DNA-proteins and chromatin level, (iii) the nuclear level, which includes the dynamics and three-dimensional spatial organization of the chromosome inside the nucleus, (iv) the cell regulation in response to internal and external signals and factors, which is able to remodel the genome structure and function. There is an increasing evidence that such a higher order organization of chromatin structure and dynamics contribute in an essential way to the regulation of gene expression and therefore to cell activity [76–79].

We have to take into account epigenetic in order to understand some features of our genome, its topological forms and the ways in which its functions. The two properties are closely related. Epigenetics encompass the many processes that cannot be accounted for by the simple genetic code [80], and the term refers to extra layers of instructions, that is of biological organization and information (notably cellular, organismal and environmental) that influences gene activity without altering the DNA sequence (for more details see Esteller and Almouzni [81], Cavalli and Heard [82]).

The development in recent years of epigenetics entails the emergence of a more integrative and global approach to the study of biological forms and functions. To tackle the whole human epigenome and to deal with the entire organisms, it is needed to elucidate the relationship between the different level of plasticity of protein complexes associated with chromatin remodeling and gene regulation, and the various levels of complexity exhibited by the phenotypic patterns during embryogenesis. The landscape of genetic expression revealed by epigenetics studies appear to be much more complex than that showed by DNA sequence alone, and it clearly results from diverse layers of biological information (DNA folding, histone modifications, the complex regulatory roles of DNA methylation, chromatin remodeling complexes, spatial organization of chromosomes, architecture of nuclear bodies, cell morphology and mobility), which intervenes at different stages of the spatial and temporal development and evolution of a living human organism.

16 The Need for a Systems Biology Approach

Another reflection, which relates to the previous one, is aimed at highlighting the importance of a systems biology approach. System biology is about *interactions* rather than about constituents, although knowing the constituents of the system under study may be a prerequisite for starting description and modeling. Interactions often bring about new properties—sometimes called *emergent properties*—, for instance, a system may start oscillating although the constituent alone would not.

Another important example is that evolutionary biologists have wondered for long jump-like transitions can occur in evolution. From the viewpoint of systems theory, the answer arises from bifurcations. In a non-linear system, at certain points in parameter space, called *critical points*, bifurcations occur, that is, a small change in a parameter leads to a qualitative change in system behaviour—e.g. a switch from steady state to oscillation. It is clear that the number of potential interactions within a system is far greater than the number of constituents. If only pairwise interactions were allowed, the former number would be n^2 if the latter number were denoted by n. The number of interactions is even larger if interactions within triples and larger sets are allowed, as is the case in multi-protein complexes.

In the sense of systems biology, a biological object or being is a *system* if emergent properties result from it. Genomics has certainly been a very important and fruitful undertaking and gave us much new insights into molecular biology. However, much of molecular biology is based on reductionism and simple determinism. It is an extreme exaggeration to say that the human genome has been "deciphered". Besides the fact that not to all ORFs functions have been assigned yet, it should be acknowledged that even if all functions were known, we would be far from understanding the phenomenon of life because knowledge of all the individual gene products does not say much about the interactions between them. According to a system's view of life, the study of the dynamics and interaction networks is essential for understanding the ways in which living organisms regulate their cellular activity and organize their physiological growth.

One of the major goals of systems biology is to find appropriate ways of diagramming and mathematically describing the specific, complex interactions within and between living cells. Because complex systems have emergent properties, their behaviour cannot be understood or predicted simply by analyzing the structure of their components. The constituents of a complex system interact in many ways, including negative feed-back and feed-forward control, which lead to dynamic features that cannot be captured satisfactorily by linear mathematical models that disregard cooperativity and non-additive effects. In view of the complexity of informational pathways and networks, new types of mathematics are required for modeling these systems [83].

17 The Complex Topology of Biological Structures: From Topoisomerases to Supercoiling

Complex living systems may be studied from the phenomenological and from the theoretical point of view as well. Theoretically, one studies the 3-dimensional structure of DNA and DNA–protein complexes through geometrical and topological methods [37, 84]. This study follows two steps: Firstly, one tries to show that certain topological deformations associated to the macromolecular structures during the cell cycle take part in the dynamics of chromatin, chromosome, and therefore in the cell metabolism. Then, to elucidate the way in which these deformations modulate the action of different regulatory systems, ensuring in particular the transition of this action from local-target mechanisms to global functional processes.

These interactions between topological changes and dynamical processes constitute a deep and largely unexplored meeting point for mathematics ad biology. One observes first that certain geometric properties and topological patterns work like dynamical principles, which are involved in the organization and growth of living systems, and next that these properties and patterns display intricate biological plasticity and complexity on every scale, from the very marge (the organism as an integrated system, as a whole) to the very small (the molecule as a component, which is continuously changing its conformation—for instance when it auto-assembly for forming macro-molecules aggregates with news patterns and functions—and its function according to the context in which is involved.

Plasticity and complexity are two key features of DNA structure. First, it has to be stressed that the double-helix structure is both a geometrical entity and a topological form. This topological form is itself a manifestation of two remarkable operations, namely linking and knotting. DNA within cell is a very long molecule with a striking complex topology. Topological properties of DNA are defined as those that can be changed only by breakage and reunion of the backbone; in fact, by surgery and gluing. The very significant point is that this complex topology of DNA is essential for maintaining and reproducing life, as well as for assuring the evolution of organisms.

We want now go a little bit into details by observing that the topology of DNA in vivo is set by a remarkable group of enzymes called topoisomerases [85]. These enzymes essentially promote the passage of DNA segments through each other until a stable state is achieved. This functional stability is thus made possible thanks to a conformational topological flexibility of the double-helix, and the continuous remodeling of nuclear structures is as well required for cell activity to be performed.

DNA has three important topological properties: (i) the linking number between two strands of the double-helix, (ii) the interlocking of separate DNA rings in what we call catenanes, (iii) the knotting and unknotting of the two DNA strands. To these mathematical properties there correspond three important physical and phenomenological features, which are: (a) As the number of crossing in a knot or catenane increases, the number of possible isomers grows exponentially; (b) the linking number of DNA in all organisms is less than the energetically most stable value in unconstrained (relaxed) DNA; this puts the DNA under stress, which causes it

to buckle and coil in a regular way called negative (–)-supercoiling; (c) the name supercoiling arises because it is the coiling of a molecule, which is itself formed by the coiling two strands about each other.

The ability of a region of DNA, thanks to supercoiling, to affect processes that occur hundreds to thousands of basic pairs away is an essential factor of the physiological behavior of nuclear structures and of the chromosome. Dynamic changes of DNA supercoiling act as a driving force behind the alterations of genetic activity and DNA compaction in eukaryotes and prokaryotes [86]. Both these processes involve formation of spatially organized nucleoprotein structures by DNA architectural proteins. Supercoiling is a fundamental property of DNA and chromatin. It is modulated by polymerase and topoisomerases activity and, through controlled constraints, by DNA/chromatin binding proteins. DNA supercoiling plays a key role in gene expression and genome organization. Indeed, recent studies makes it increasingly clear that DNA double helix carries at least two types of encoded information and these include the well-known genetic code and the structural information determining the changes of form and the physical properties of genetic material. It should be stressed that the two layers of information interact in a very complex and dynamic way.

18 The Functional Role of Topological Plasticity

We want again emphasize the fact that the geometrical structure and topological form of nuclear components (DNA double-helix, chromatin and chromosome) play an important role in the genome functions, cell differentiation (during embryo's development) and organism growth. Let's consider un example regarding gene functions, namely the role of enzymes topoisomerase [87]. They remarkably convert DNA double-helix from one topological conformation to another through some specific qualitative manipulations, and therefore they play an important role for assuring the crucial genetic events of replication, transcription, recombination and repair of the genetic material. Moreover, certain topological mechanisms, such as supercoiling and knotting, are involved in the fundamental biological process of the compaction of chromatin into the chromosome during the interphase and metaphase.

Another good example illustrating the complex relationship between form and function in biological systems concern proteins. The standard approach in the study of proteins focuses on one protein at a time. However, recent studies show that their biological function appear to be more a correlate of macro-molecular geometry than of chemical details. Any effective picture of protein structure must provide at the same time a model (an explanation) for the common character of all proteins, as exemplified by their many chemical and physical similarities, and for the highly specific nature of each protein type. These studies stressed the topological determinants of protein folding (see Dokholyan et al. [88], Gromov [89]). For some of them, one can show that topological properties of protein conformation determine their kinetic ability to fold. Roughly one speaks of a *macroscopic measure* of the protein *contact network*

topology, the *average graph connecti*vity, by constructing graphs that are based on the geometry of protein conformations. It has been found that the average connectivity is higher for conformations with a higher folding probability than for those with a highly probability to unfold. As a protein unfold, it encounters dynamic constraints that emerge as a consequence of its being folded into a particular low-resolution structure or topology. Fo example, it often occurs that parts of protein are entangled or wrapped within its interior, and for these "frustrated" parts to unfold requires the rest of protein to reorganize and at least partially unfold first. At this level of resolution, topological constraints can impose a time order on unfolding events and occasionally this order can be reorganized in a protein's actual nucleation process of folding "pathway" despite the extreme complexity of its interactions.

19 The Topological State of Supercoiling and Its Biological Functions

We have thus three interrelated theoretical and experimental facts, which now I would like to stress. First, DNA condensation is a driving force for double-helix unlinking and chromosome portioning, by folding in topological domains. Second, condensation is achieved by supercoiling, which is a topological state of the macromolecules enhanced by three kind of deformations (or, technically speaking, embeddings[9]), namely twisting, writhing and knotting. If DNA is modelled as a ribbon in three-space whose axis is not flat in the plane, we can define the *twist* of the ribbon abstractly as *the integral of the incremental twist of the ribbon about the axis, integrated as we traverse the axis once*; so, it simply measures how much the ribbon twists about the axis from the frame of reference of the axis: it needs to be an integer.[10] The *writhe*

[9] Let's remember the definition of the concepts of embedding and immersion. Let X and Y be topological spaces, an *embedding* of X into Y is a one-to-one bi-continuous map of X into Y—that is a homeomorphism of X onto $f(X)$. If X and Y are smooth manifolds, an *immersion* of X into Y is a smooth function, f, from X into Y that is locally one-to-one but not necessarily globally one-to-one. Equivalently, we can define an *immersion* as a smooth map whose differential, df, is non-singular. An immersed manifold is an image of an immersion. For example, the limaçon, the graph in polar coordinates of $r = 1/2 + \cos(\theta)$, is an immersed circle in the plane. Another important example, the most familiar image of the Möbius band is a "circular band with a half-twist"—call this subset M_1. This is an image of one embedding of this 2-dimensional manifold into ambient space R^3. This manifold is non-orientable.

[10] Recall the important concept of Dehn twist. A *Dehn twist* is a type of self-homeomorphism on a 3-dimensional orientable closed manifold. Let $\sum$ be an orientable 2-manifold of genus g, and consider an embedded circle: $S^1 \to \sum$. The embedded circle has a tubular neighborhood in $\sum$, homeomorphic to an annulus $S^1 \times [0, 1]$. Representing elements of S^1 by complex numbers z of norm 1, a *twist* on the annulus may be defined by sending $(z, t) \to (\exp(2\pi it)\, z, t)$, which equals the identity on the boundary circles where $t = 0$, $t = 1$. This define a self-homeomorphism (mod boundary) on the annulus. The corresponding Dhen twist on $\mathbf{Z}$ is obtained by extending the self-homememomorphism on the annulus to the entire surface, by taking a point p to itself if p is outside the annulus. If were working in the category of smooth manifolds, one may modify the twist on the annulus by taking $(z, t) \to (\exp(2\pi i f(t))\, z, t)$ where f is a smooth bump function such that $f(t) = 0$

measures how much the axis of the ribbon is contorted in space. Because negative (–)-supercoiling in bacteria arises from topological misalignment and not a protein corset, it has the flexibility to do work. Thirdly, supercoiling result from topological strain and the contortion of DNA by proteins, notably the nucleosome histone octet and the structural maintenance of chromosomes (SMC) proteins.

There are three ways, which have been experimentally observed in vivo, in which condensation of chromosome by supercoiling occurs, and to each of them corresponds a topological model for explaining the compaction of chromosomes in the cell nucleus. These ways are: (i) (–)-supercoiling by gyrase compacts the chromosomes such that random passages by topoisomerase IV disentangle them. In particular, topoisomerase IV is responsible for decatenation or unlinking of DNA.[11] (ii) With the second type of condensation via supercoiling, that is by core histones, DNA is compacted in independent successive stages such that the total compaction is the product of compaction of each stage. The first stage of this compaction is via solenoidal[12] wrapping of DNA in the nucleosome. Although the compaction achieved

---

on a small neighborhood of $t = 0$ and $f(t) = 1$ on a small neighborhood of $t = 1$. This modification ensures that we get a self-diffeomorphism on the surface $\sum$ as a smooth manifold.

One important theorem was first proved by Dehn, then later, in a simplified version, by Likorish:

Theorem (Dehn-Likorish). *The Dehn twists generate the mapping class of* $\sum$*, of orientations-preserving homeomorphisms considered modulo isotopy.*

(In fact, Likorish described 3 $g - 1$ explicit embedded circles for a surface $\sum$ of genus g whose corresponding twists give the generators.).

One example that is easy visualized is the mapping class group $MCG(\sum)$ of the torus $\sum = S^1 \to S^1$. Here the canonical map $MCG(\sum) \to$ Aut $(\pi_1(\sum)) \cong$ Aut $(\mathbf{Z} \times \mathbf{Z})$ is an isomorphism (where automorphisms of $\mathbf{Z}$ are required to preserve orientation, i.e., are elements of the modular group SL_2 $(\mathbf{Z})$. Thus, in this case the Dehn twist can be visualized in terms of their action on loops representing elements of $\pi_1(\sum)$.

[11] To be more precise, topoisomerase IV fulfils another function, which consists in relaxing positive supercoils generated during DNA replication.

[12] In rigorous mathematical terms, a solenoid is a compact connected topological space (i.e. a continuum) that may be obtained as the inverse limit of an inverse system of topological groups and continuous homeomorphisms

$$(S_i, f_i),\ f_{i+1} \to S_i, \quad i \geq 0,$$

where each S_i is a circle and f_i is the map that uniformly wraps the circle S_{i+1} n_i-time ($n_i \geq 2$) around the circle S_i. This construction can be carried out geometrically in three-dimensional Euclidean space R^3. A solenoid is a one-dimensional homogeneous indecomposable continuum that has the structure of a compact topological group. In the geometrical theory of dynamical systems, a solenoid can arise as a one-dimensional expanding attractor, or Smale-Williams attractor, and forms an important example in the theory of hyperbolic dynamical systems (see Williams [90]). Geometrically, each solenoid may be constructed as the intersection of a nested system of embedded solid tori in R^3. Fix a sequence of natural numbers $\{n_i\}$, $n_i \geq 2$. Let $T_0 = S^1 \times D$ be a solid torus. For each $i \geq 0$, choose a solid torus T_{i+1} that is wrapped longitudinally n_i times inside the solid torus T_i. Then their intersection $\Lambda = \cap_{i \geq 0} T_i$ is homeomorphic to the solenoid constructed as the inverse limit of the system of circles with the map determined by the sequence $\{n_i\}$. A variant of a geometrical solenoid is the expanding attractor in the theory of smooth dynamical systems constructed by Stephen Smale. Denote the angular coordinate on the circle S^1 by t (t is defined modulo 2π) and consider the complex coordinate z on the two-dimensional unit disk D. Let f be

is modest, the nucleosome provides a fundamental structure for genome organization and function. The structure of a nucleosome reveals a scaffolding that forces DNA ordered solenoidal supercoils [91]. (iii) The third type of compaction cum supercoiling, that by condensing, is needed for the formation of mitotic chromosomes from the open interphase forms.

There are two general classes of supercoil, known as interwound supercoil and toroidal supercoil. The circular DNA (that is, with the ends of the molecule fixed) consists of a series of open spirals that wind around an imaginary ring, or toroid; this kind of supercoiling is known as *toroidal*. But the circular DNA can also wind above and below itself several times, and this kind of supercoiling is called *interwound*. In practice, real DNA supercoils may contain portions of both the toroidal and the interwound geometries. Thus, where certain parts of DNA are highly curved, on amount of either the base sequence or due to wrapping around a protein, one may find toroidal structures, since the DNA in a toroidal supercoil is highly curved throughout! Alternatively, if such curved portions of DNA are not very long, they may locate themselves at the two strongly curved end-loops of an interwound supercoil, as shown in Figs. 10 and 11. Sometimes the interwound and toroidal geometries may occur together, as in the looped-linear DNA which is shown schematically in Figs. 10 and 11. On a small scale, within any loop, the coiling is toroidal on account of the wrapping of DNA around protein spools, but on a large scale, over the full length of any loop, the structure is interwound. One can see this kind of arrangement in many natural structures, especially in trees.

Now, what are the relative stabilities of these two forms of DNA supercoiling? In other words, under which conditions will a DNA molecule be interwound, and when will it be toroidal? The interwound shape is usually very stable, and most underwound overwound DNA molecules will naturally adopt an interwound shape, in the absence of other forces. But the proteins that associate with DNA in living

---

the map of the solid torus $T = S^1 \times D$ into itself given by the explicit formula $f(t, z) = (2t, 1/4z + 1/2e^{it})$. This map is a smooth embedding of T into itself that preserves the foliation by meridional disks (it is essential that ¼ < ½ and $\frac{1}{4} + \frac{1}{2} < 1$). If T is imagined as a rubber tube, the map f stretches it in the longitudinal direction, contract each meridional disk, and wraps the deformed tube twice inside T with twisting, but without self-intersection. The hyperbolic set Λ of the discrete dynamical system (T, f) is the intersection of the sequence of nested solid tori described above, where T_i is the image of T under the ith iteration of the map f. This set is a one-dimensional (in the sense of topological dimension) attractor, and the dynamics of f on Λ has the following interesting properties: (i) meridional disks are the stable manifolds, each of which intersects Λ over a Cantor set, (ii) periodic points of f are dense in Λ, (iii) the map f is topologically transitive in f.

General theory of solenoids and expanding attractors, not necessarily one-dimensional, was developed by R.F. Williams and involves a projective system of infinitely many copies of a compact branched manifold in place of the circle, together with an expanding self-immersion. (Recall that branched manifolds are a generalization of a differentiable manifolds which may have singularities of very restricted type and admits a well-defined tangent space at each point; a branched manifold is covered by n-dimensional "coordinate charts", each of which involves one or several "branches" homeomorphically projecting into the same differentiable n-disk in R^n.) A solid torus wrapped twice around inside another solid torus in R^3 is a very nice example of branched manifold. Another beautiful example of branched manifold is the Smale-Williams solenoid, or Smale-Williams expanding attractors, which are one-dimensional hyperbolic attractors.

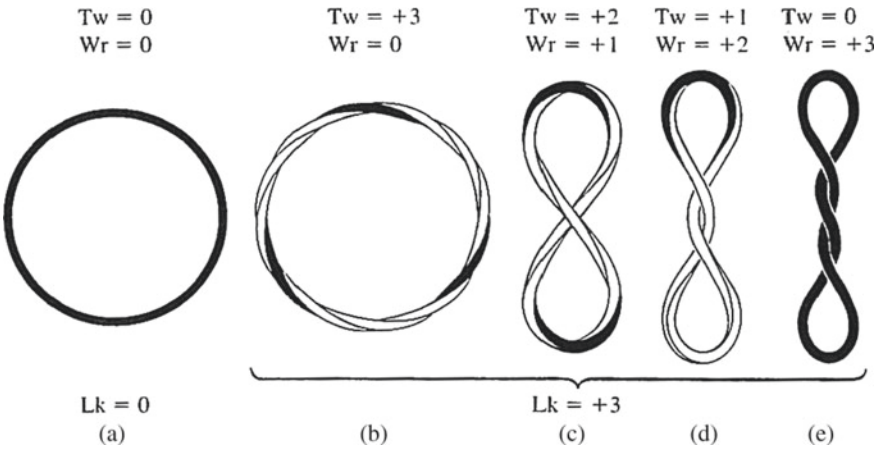


Fig. 10 Five closely related circular DNA molecules: **a** and **b** show open circles, while **c**, **d** and **e** show interwound supercoils. The DNA in its stress-free, relaxed form is drawn as a rubber rod of square cross-section, with one face black (From Boi [25])

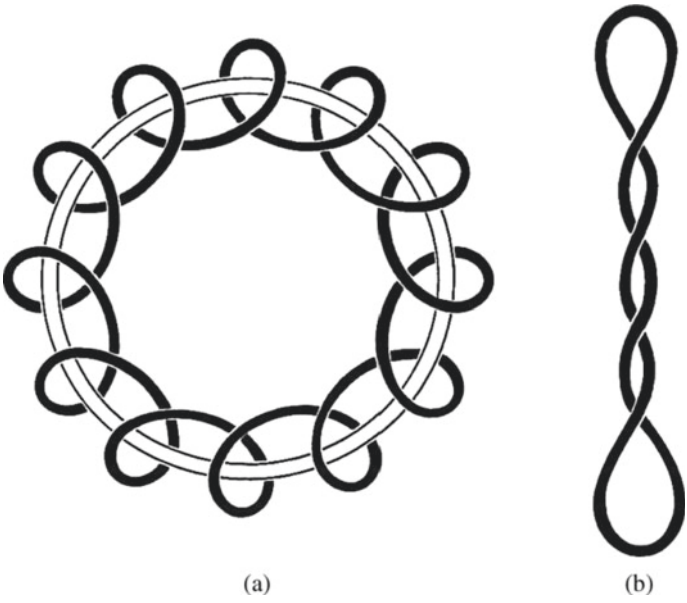


Fig. 11 Two general varieties of DNA supercoil. In **a**, the DNA coils into a series of spirals about an imaginary toroid or ring (shown here by open lines); and so this kind of wrapping is known as 'toroidal'. In **b**, the DNA crosses over and under itself repeatedly; and so this kind of wrapping is known as 'interwound' (From Boi [25])

cells can sometimes change the situation dramatically, and favors the toroidal over the interwound form by wrapping the DNA around themselves.

20 Conclusive Remarks

To conclude, let's emphasize the important fact that the complex topology of DNA is essential for the evolution of all organisms, hence for life. In particular, it is needed for the process known as DNA replication, whereby a replica of the DNA is made and one copy is passed on to each daughter cell. Some of the open and key problems relating to the functional processing of biological systems seem to be deeply linked to the following issues: (i) to the conformational, organizational and biological role of topoisomerase which, because of their extreme structural plasticity and functional complexity, still remain to be elucidated. (ii) To the DNA–protein supercoiling process, because it links the biological activity of DNA to its tertiary structure (that is, the different geometrical structures and general topologies of DNA and DNA-proteins complexes) [92] and not to its sequence (see Misteli [65]). Indeed, almost all cellular processes seem to be deeply related to the way in which supercoiling is realized. (iii) To the three-dimensional organization of chromatin, which is a nucleoprotein complex and the stuff chromosomes are made of. This organization not only compacts DNA but also plays a fundamental role in regulating interactions with DNA during its metabolism.

In conclusion, it can be said that DNA double helix is an elegant and robust structure that allows the genetic information to be stored, protected, replicated and repaired. However, the interwound nature of the double helix also has the potential to impose a number of topological constraints on the genetic material that affect all of its physiological functions (see Boi [51]). This means that DNA double helix admit different elastic deformations [93], in spite of some geometric rigidity characterizing its basic structure, and constrained particularly by bending, coiling and twisting. As long as the ends of DNA strands are fixed in space, as is the case for circular molecules or long linear molecules that are attached to solid supports, DNA can be considered to be a dynamically topological system. Therefore, topological properties of DNA are designed as those that cannot be altered without breaking one or both strands of the double helix. The topological nature of DNA double-helix is also demonstrated by the fact that it adopts various conformations. This intertwining of topological remodelling of biological structures and acquisition of functionality by them in different contexts is very likely one of the most fundamental features of biological systems. Variability of forms and enhancement of functions continuously interact and both are essential for the evolution, generation and preservation of living beings.

Acknowledgements This paper is an enlarged version of lectures given at the Summer School on Complexity held in 2018 in Como (Villa del Grumello). I am grateful to the organizers for kind

hospitality, and to Sergio Albeverio, Jürgen Jost, Elisa Mastrogiacomo and Stefania Ugolini for useful comments and discussion.

References

1. Boi, L.: Epigenetic phenomena, chromatin dynamics, and gene expression. New theoretical approaches in the study of living systems. Biol. Forum **101**(3), 405–442 (2008)
2. Boi, L.: Mathematical knot theory. In: Françoise, J.-P., Naber, G., Sun, T.S. (eds.) Encyclopedia of Mathematical Physics, pp. 399–406. Elsevier, Oxford (2006)
3. Boi, L.: Topological knots models in physics and biology. In Geometries of Nature, Living Systems and Human Cognition: New Interactions of Mathematics with Natural Sciences and Humanities, pp. 203-278. World Scientific, Singapore (2005)
4. Boyland, P.: Fluid mechanics and mathematical structures. In: Ricca, R.L. (ed.) An Introduction to the Geometry and Topology of Fluid Flows, pp. 105–134. Kluwer, Dordrecht (2001)
5. Hornig, G.: The geometry of reconnection. In: Ricca, R.L. (ed.) An Introduction to the Geometry and Topology of Fluid Flows. Kluwer, Dordrecht (2001)
6. Kauffman, L.H.: Knots and physics. Series on Knots and Everything, vol. 1. World Scientific, Singapore (2001)
7. Albeverio, S., et al. (a cura di): Extreme Events in Nature and Society. Springer-Verlag, Heidelberg (2006)
8. AAVV: Topological phase transitions and topological phases of matter. The Nobel Prize in Physics, The Royal Swedish Academy of Sciences, 4 October 2016
9. Kosterlitz, J.M., Thouless, D.J.: Ordering, metastabilities and phase transition in two-dimensional systems. J. Phys. C: Solid State Phys. **6**(7), 1181–1198 (1973)
10. Anderson, P.W.: More is different. Broken symmetry and the nature of the hierarchical structure of science. Science **177**, 393–396 (1972)
11. Jost, J.: On the notion of complexity. Theory Biosci. **117**, 161–171 (1995)
12. Kitano, H.: Systems biology: a brief overview. Science **295**, 1662–1664 (2002)
13. Ricca, R.L., Moffat, H.K.: The helicity of a knotted vortex filament. In: Moffat, H.K. (ed.) Topological Aspects of Dynamics of Fluid and Plasmas, pp. 225–236. Kluwer, Dordrecht (1992)
14. Boi, L.: Topological knot theory and macroscopic physics. In: Françoise, J.-P., Naber, G., Sun, T.S. (eds.) Encyclopedia of Mathematical Physics, pp. 271–277. Elsevier, Oxford (2006)
15. Oberti, C., Ricca, R.L.: Influence of winding number on vortex knots dynamics. Sci. Rep. **9**, 17284–17297 (2009)
16. White, J.H.: Self-linking and the Gauss integral in higher dimensions. Amer. J. Math. **91**, 693–728 (1969)
17. Gray, A.: Modern Differential Geometry of Curves and Surfaces, with Mathematica. CRS Press, Boca Raton, FL (1997)
18. Roe, J.: Winding Around: The Winding Number in Topology, Geometry and Analysis. American Mathematical Society, Student Mathematical Library, vol. 76. Providence (2015)
19. Oberti, C., Ricca R.L.: Influence of winding number on vortex knots dynamics. Sci. Rep. **9**(1) (2019)
20. Massey, W.S.: Algebraic Topology: An Introduction. Hardcourt (1967)
21. Dubrovin, B.A., Fomenko, A.T., Novikov, S.P.: Modern Geometry and Applications. Part II: The Geometry and Topology of Manifolds. Graduate Texts in Mathematics, vol. 104. Springer, New York (1985)
22. Ranicki, A.: High-Dimensional Knot Theory—Algebraic Surgery in Codimension 2. Springer-Verlag, Berlin & Heidelberg (1998)
23. Ophl, W.F., Roberts, G.W.: Topological considerations in the theory of replication of DNA. J. Math. Biol. **6**, 383–402 (1978)

24. White, J.H.: An Introduction to the Geometry and Topology of DNA Structures. CRC Press, Boca Raton (1989)
25. Boi, L.: When topology meets biology 'for life'. Remarks on the way in which topological form modulates biological function. In: Bartocci, C., Boi, L., Sinigaglia, C. (eds.) New Trends in Geometry and its Role in the Natural and Life Sciences, pp. 241–303. Imperial College Press, London (2011)
26. Olsen, K., Bohr, J.: The generic geometry of helices and their close-packed structures. Theoret. Chem. Acc. **125**, 207–215 (2010)
27. Almouzni, G.: Assembly of spaced chromatin: involvement of ATP and DNA topoisomerases. EMBO J. **7**, 4355–4365 (1988)
28. Ridgway, P., Almouzni, G.: Chromatin assembly and organization. J. Cell Sci. **114**, 2711–2722 (2001)
29. Cozzarelli, N.R., Holmes, V.F.: Closing the ring: Links between SMC proteins and chromosome portioning, condensation, and supercoiling. Proc. Natl. Acad. Sci. USA **97**(4), 1322–1324 (2000)
30. Boi, L.: Geometrical modeling of DNA and the structural complexity of the chromosome. J. Biophys. (2021) (forthcoming)
31. Hildebrandt, S., Tromba, A.: Mathematics and Optimal Form. Scientific American Library (1985)
32. Meeks III, W.H., Pérez, J.: Properly embedded minimal plana domains with infinite topology are riemann minimal examples. In: Jerison, D., et al. (eds.) Current Developments in Mathematics (2008), pp. 281–346. International Press, Boston, MA (2009)
33. Boi, L.: Geometry of dynamical systems and topological stability: from bifurcations, chaos and fractals to dynamics in the natural and life sciences. Int. J. Bifurc. Chaos **21**(3), 815–867 (2011)
34. Gasser, S.: Visualizing chromatin dynamics in interphase nuclei. Science **296**, 1412–1416 (2002)
35. Cremer, T., et al.: Higher order chromatin architecture in the cell nucleus: on the way from structure to function. Biol. Cell **96**, 555–567 (2004)
36. Atiyah, M.: The geometry and physics of knots. Lezioni Fermiane, Accademia Nazionale dei Lincei, Edizioni della Scuola Normale Superiore, Pisa (1989)
37. Boi, L.: Geometrical and topological modeling of supercoiling in supramolecular structures. Biophys. Rev. Lett. **2**(3), 1–13 (2007)
38. Summers, D.W.L.: Knot theory and DNA. In: New Scientific Applications of Geometry and Topology. PSAM, vol. 45. American Mathematical Society (1992)
39. Wolffe, A.P.: Chromatin. Structure and Function. Academic Press, London (1998)
40. Smale, S.: Differentiable dynamical systems. Bull. Amer. Math. Soc. **13**, 747–817 (1967)
41. Thom, R.: Structural Stability and Morphogenesis. Benjamin, New York. (1972)
42. Thom, R.: Paraboles et catastrophes. Flammarion, Paris (1983)
43. Boi, L.: Topological ideas and structures in fluid dynamics. JP J. Geom. Topol. **8**(2), 151–184 (2008)
44. Rolfsen, D.: Knots and Links. Mathematical Lectures Series 7. Publish or Perish, Huston (1990)
45. Lescop, C.: An introduction to finite type invariants of knots and 3-manifolds. In: ICPAM-ICTP Research School, Symplectic Geometry and Geometric Topology. Meknès, Marocco (2012)
46. Dehn, M.: Über die Topologie des dreidimensionalen Raumes. Math. Ann. **69**(1) 137–168 (1910)
47. Seifert, H.: Topologie dreidimensionaler gefaserter Räume. Acta Math. **60**, 147–238 (1933)
48. Seifert, H., Threlfall, W.: Lehrbuch der Topologie. Teubner, Leipzig (1934)
49. Albeverio, S., Blanchard, P. (eds.): Direction of Time. Spinger, Heidelberg (2014)
50. Devaney, R.L.: Introduction to Chaotic Dynamical Systems. Addison-Wesley, Redwood City, CA (1989)
51. Boi, L.: Plasticity and complexity in biology: topological organization, regulatory protein networks and mechanism of gene expression. In: Terzis, G., Arp, R. (eds.) Information and Living Systems. Philosophical and Scientific Perspectives, pp. 205-250. The MIT Press, Cambridge, Mass (2011)

52. Boi, L.: Remarks on the geometry of complex systems and self-organization. In: Isonomia-Epistemologia, special issue on Complessità e riduzionismo, pp. 21–36. Urbino (2012).
53. Prigogine, I., Nicolis, G.: Fluctuations in the mechanism of instabilities. In: Proceedings of the 3rd International Conference from Theoretical Physics to Biology (1971), pp. 89–109. S. Karger, Basel (1973)
54. Cornish-Bowden, A., Cárdenas, M.L.: Systems biology may work when we learn to understand the parts in terms of the whole. Biochem. Soc. Trans. **33**, 516–519 (2005)
55. Liljenström, H.: Multi-scale causation in brain dynamics. In: Kozma, R., Freeman, W. (eds.) Cognitive Phase Transitions in the Cerebral Cortex: Enhancing the Neuron Doctrine by Modelling Neural Fields, pp. 177–186. Springer, New York (2016)
56. Misteli, T.: The concept of self-organization if cellular architecture. J. Cell Biol. **155**(2), 181–185 (2001)
57. Karsenti, E.: Self-organization processes in living matter. Interdisc. Sci. Rev. **32**(6), 21–38 (2007)
58. Morange, M.: Post-genomic, between reduction and emergence. Synthese **151**, 355–360 (2006)
59. Boi, L.: The interlacing of upward and downward causation in complex living systems: on interactions, emergence and wholeness. In: Paolini Paoletti, M., Orilia, F. (eds.) Philosophical and Scientific Perspectives on Downward Causation, pp. 180–203. Routledge, London (2017)
60. Stewart, I.: Life's Other Secrets: The New Mathematics of the Living World. Allen Lane, London (1998)
61. Li, E.: Chromatin modification and epigenetic reprogramming in mammalian development. Nat. Rev. Genet. **3**, 662–673 (2002)
62. Kauffmann, S.: The Origins of Order. Self-Organization and Selection in Evolution. Oxford University Press, New York (1993)
63. Nicolis, G., Prigogine, I.: Exploring Complexity: An Introduction (1989)
64. Noble, D.: The Music of Life. Biology Beyond the Genome. Oxford University Press, Oxford (2006)
65. Misteli, T.: Beyond the sequence: cellular organization of genome function. Cell **128**, 787–800 (2007)
66. Cornish-Bowden A.: Putting the systems back into systems biology. Perspect. Biol. Med. **49**(4) 1–9 (2006)
67. Van Regenmortel, M.H.V.: Reductionism and the search for structure-function relationships in antibody molecules. J. Mol. Recognition **15**, 240–247 (2004)
68. Mandelbrot, B.: The Fractal Geometry of Nature. W. F. Freeman, New York (1982)
69. May, R.: Simple mathematical models with very complicated dynamics. Nature **261**, 459–464 (1976)
70. Nguyen, D.C.: Topological Dynamics of Random Systems. Clarendon Press, Oxford (1997)
71. Hausdorff, F.: Mengenlehre. de Gruyter, Berlin (1935)
72. Lehn, J.-M.: Toward self-organization and complex matter. Science **295**, 2400–2403 (2002)
73. Lehn, J.-M.: Supramolecular chemistry: from molecular information towards self-organization and complex matter. Rep. Prog. Phys. **67**(3), 249–284 (2004)
74. Nedelec, F., Surrey, T., Karsenti, E.: Self-organization and forces in the microtubule cytoskeleton. Curr. Opin. Cell. Biol. **15**(1), 118–124 (2003)
75. Scherrer, K., Jost, J.: Gene and genon concept: coding versus regulation. Theory Biosci. **126**, 65–113 (2007)
76. Jaenisch, R., Bird, A.: Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat. Genet. **33**, 245–254 (2003)
77. Ehrenhofer-Murray, A.E.: Chromatin dynamics at DNA replication, transcription and repair. Eur. J. Biochem. **271**, 2335–2349 (2004)
78. Felsenfeld, G.: Chromatin: an essential part of transcriptional apparatus. Nature (London) **421**(355), 219–223 (1992)
79. Görisch, S.M., et al.: Nuclear body movement is determined by chromatin accessibility and dynamics. Proc. Natl. Acad. Sci. USA **101**, 13221–13226 (2004)
80. Waddington, C.H.: The Strategy of Genes. Allen & Unwin, London (1957)

81. Esteller, M., Almouzni, G.: How epigenetics integrates nuclear functions. In: Workshop on Epigenetics and Chromatin: Transcriptional Regulation and Beyond, EMBO Rep. 6, pp. 624-628 (2005)
82. Cavalli, G., Heard, E.: Advances in epigenetics link genetics to the environment and disease. Nature **571**, 489–499 (2019)
83. Jost, J.: Biologie und Mathematik. Springer-Verlag, Berlin/Heidelberg (2019)
84. Cozzarelli, N.R.: Evolution of DNA topology: implications for its biological role. In: New Scientific Applications of Geometry and Topology, PSAM, vol. 45. American Mathematical Society (1992)
85. Wang, J.C.: DNA topoisomerases. Ann. Rev. Biochem. **65**, 635–692 (1996)
86. Muskhelishvili, G., Travers, A.: The regulatory role of DNA supercoiling in nucleoprotein complex assembly and genetic activity. Biophys. Rev. **8**, 5–22 (2016)
87. Roca, J.: The mechanisms of DNA topoisomerases. Trends Biochem. Sci. **20**, 156–160 (1995)
88. Dokholyan, N.V., et al.: Topological determinants of protein folding. Proc. Natl. Acad. Sci. **99**(13), 8637–8641 (2002)
89. Gromov, M.: Crystals, proteins, stability and isoperimetry. Preprint IHES, pp. 1–30, August 3, 2010
90. Williams, R.F.: Expanding attractors. Publications mathématiques de l'IHÉS **43**, 169–203 (1974)
91. Kepes, F., Vaillant, C.: Transcription-based solenoidal model of chromosomes. Complexus **1**, 171–180 (2003)
92. Harteis, S., Schneider, S.: Making the bend: DNA tertiary structure and protein-DNA interactions. Int. J. Mol. Sci. **15**(7), 12335–12363 (2014)
93. Durickovic, B., Goriely, A., Maddocks, J.H.: Twist and stretch of helices via the Kirchhoff-Love rod model of elastic filaments. Phys. Rev. Lett. **111**, 108103–108105 (2013)

The Complexity Theory and Financial Systems Regulation

Ilaria Capelli

Abstract Complexity theory provides a revolutionary point of view when dealing with classical legal problems and in understanding the intimate interconnections between law and other social sciences. Complexity theory understands the legal system as a complex adaptive system like the economy: these two systems complexly interact with each other, as well as with the other complex social and physical systems they are interconnected with. Furthermore, complexity theory shows us how there is no perfect form of regulation available. A complex version of the law would be flexible and adaptable without, however, being able to provide a "perfect and stable" solution to regulation. Such a system would only deliver a lower number of mistakes and more adapt reactions. Complexity theory is considered an important tool for better understanding, and therefore better regulating, global financial systems. Especially after the great crisis, there is a deep interest in using complex system models, with the aim of forecasting the long-lasting effects of the regulation system and to predict a potential collapse of the financial system. Complexity theory introduces in the debate ideas and concepts that are new for law scholars and also urges the institutions and the scholars themselves to analyze the reality in overall terms, involving at once legal skills, system modeling, economic skills and the knowledge of the technological tools.

Keywords Complex systems · Legal thinking · Financial system · Regulation effects · New regulatory tools

Summary

(1) The characteristics of complex systems. (2) Complexity theory and the law: complex vs. complicated. (3) Complexity theory and legal thinking. (4) Complexity theory and the financial system. (5) Complexity jurisprudence and financial systems regulation. (6) The financial system regulation and systemic risk. (7) Complex systems models and the role of fraud.

I. Capelli (✉)
Department of Economics, University of Insubria, Varese, Italy
e-mail: ilaria.capelli@uninsubria.it

S. Albeverio et al. (eds.), *Complexity and Emergence*, Springer Proceedings in Mathematics & Statistics 383, https://doi.org/10.1007/978-3-030-95703-2_2

1 The Characteristics of Complex Systems

The aim of this short presentation is to define and explain how law scholars are using complexity theory to make sense of law. In relation to law, complexity theory provides a revolutionary point of view when dealing with classical legal problems and in understanding the intimate interconnections between law and other social sciences.

Even if the origins of complexity theory can be traced back to early work on cybernetics and information theory [34], it's generally believed that the first notion of a distinctive theory of complexity was theorized by the Belgian physical chemist Ilya Prigogine. This scientist, who won the Nobel Prize for Chemistry in 1977, worked on far-from-equilibrium systems and introduced the notion of 'order out of chaos' [28], which can be considered as the first aphorism of a theory of complex systems.

As highlighted by the real beginning, the complexity theory is inherently interdisciplinary. The US Santa Fe Institute (established in 1986) hosted some of the leading academics in the field of complexity theory: Kauffman and Holland, who both worked on biological complexity [16, 17, 19] Arthur, who worked on economic complexity [1, 2]; scholars in the field of mathematics and computer sciences, and scientists interested in the continental philosophy of Deleuze and Guattari [11] and Morin [24].

It's well known that complexity theory first developed in the sphere of natural sciences. According to this theory, order arises without the need for a guiding hand or a central control: in a complex system, the structural order emerges spontaneously as the result of the interactions between the elements which compose the system, as they encounter new information.

Through this theory, natural science scholars explain the structure of insect colonies while neurologists adopt it in an effort of discovering the relationship between the mind and the brain. Physics and chemistry reveal the existence of complex systems like the Great Red Spot vortex of Jupiter; biology uncovers the existence of complex adaptive systems such as ant colonies and the immune system.

Social sciences as well found how complexity theory and this revolutionary approach explain the importance of connectivity and dynamic network organization, unpredictability, systemic instabilities, and rapid change.

When talking about complexity, the main difference between social and natural sciences is that social sciences involve human behavior, consciousness, and different ways of reacting to events, each one determined by personal ethical, moral or religious beliefs. The social world is composed of complex adaptive systems such as those of language and 'symbolic interactions', as well as political systems.

Due to this interdisciplinary kind of approach, there is not a generally agreed definition of the complexity theory. Hence, there is a collection of different theories and approaches related to complexity theory that began to grow in the 1990s, as an alternative approach in response to the behavior of systems not readily understood using traditional approaches.

The absence of a general definition of the complexity theory has not prevented scholars in both the natural and social sciences from looking to the language of complexity with the aim of investigating applying it to almost every field of study. For example, complexity can be used when studying economics, law, cosmology, biology, and artificial intelligence.

In all of these domains, complexity theory allows scientists to identify common principles that guide the dynamics and the evolution of systems across all of these domains. These principles reflect a deeper order that deeply influences the physical and social world in which we live by giving it a structure [25].

Complexity theorists recognize a series of qualities in complex systems, such as ceaseless creativity and transformation, adaptation, evolution, order out of chaos, far-from-equilibrium processes, and spontaneous self-organization.

Complexity theory has revolutionized many areas of the natural sciences, and its core insights have been adopted by social sciences to provide a new and more accurate way of dealing with human social existence by emphasizing the importance of connectivity and dynamic network organization, unpredictability, systemic instabilities, and rapid change.

The interdisciplinarity and the absence of an agreed general definition lead to identifying this approach as a set of tools, but perhaps a more accurate description is the one of 'a conceptual framework, a way of thinking, and a way of seeing the world' [25].

Complexity theorists are conscious of this situation and the absence of a general definition is by now an agreed characteristic of this view of the world. As the philosopher and sociologist Edgar Morin said, it is not possible to uncover general laws of complexity. The key is to make sense of the relationships between the whole and the parts by focusing on the notions of order and disorder.

A possible alternative could be the search for the laws of complexity. This complexity theory, however, would be a "restrictive complexity" influenced by the paradigm of classical science, and because of that, some important elements would be inevitably missed.

The correct way of engaging with complexity is to focus on the interactions between the system and its component elements, the interactions between the component elements, and their interactions with elements outside of the system. Despite the absence of a common definition of a complex system, there is some literary consensus on the characteristics of complex systems. These include:

(1) Complex systems are self-organizing. There is no controlling power or central control in a complex system, which is the result of the actions and interactions of micro-level component elements;
(2) Complex systems change over time with the flow of new matter, energy or information into the system, generating new characteristics;
(3) Complex systems not only interact with agents and elements in the external environment but also with other complex systems (for example, the economic system); in this way complex systems will increase their complexity by interconnecting and interacting with other complex systems;

(4) Whilst complex systems may remain stable for long periods, their non-linearity means that radical change can happen quickly and unexpectedly, with complex systems existing somewhere between entropy (where the system decays over time) and chaos (where too much activity makes stable structures impossible to maintain) [6, 9].

2 Complexity Theory and the Law: Complex Versus Complicated

There are some legal scholars who seek to apply the insights from complexity theory to law. Complexity gives law scholars a new way to describe law systems. However, facing the complexity theory is not simple for lawyers.

Obstacles arise, first of all, from the problem related to the ambiguous meaning of the word "complex".

Generally speaking, there is a widespread belief according to which the legal system is something "complex" and the two expressions, "complex" and "legal system", are systematically used in close proximity. Because of that, for example, it's possible to read that "the English legal system is a vast and complex subject" [12] and legal systems are analyzed as complex systems by scholars [27].

The complexity theory approach is not focused on the complicatedness of the law or its intricacy; regarding the law, complexity theorists analyze how agents interact and the aggregated product of their interactions.

Complicatedness and complexity are not the same.

The legal system is complicated: a lawyer must study law for many years just to have the privilege of taking the bar exam. The distinction between complicatedness and complexity lays in the essence of the complex adaptive systems theory: in a complicated world, the composing elements of a system maintain a degree of independence from one another. Thus, removing one of such elements (reducing the level of complication) does not fundamentally alter the system's behavior (apart from the one directly influenced by the removed piece).

A system is complex when the dependencies between the elements become important: in such a system, removing a single element radically changes the system behavior, to an extent that goes well beyond what is embodied by the removed element. Complex adaptive systems theory analyzes the inter-agent dependencies and the effect they produce in the system.

Systematic and adaption are the main properties of the theory of complexity: in a system, agents interact adaptively. An example could be the dilemma of the two-player prisoner's theory. Bilateral adaptive system models are rather simple; on the other hand, multi-agent adaptive systems can be more complicated.

This effectively applies to physical and biological systems. Concerning social systems, the main difference is that humans are the intentional designers of law, writing the rules for other social systems.

Nonetheless, the legal system is a complex adaptive system since it regulates complex adaptive systems, and the law itself co-evolves following the social system it aims to regulate.

The direct consequence of co-evolution is the impossibility to create a perfect legal system.

Complexity theory affects the way in which we understand the idea of law. More specifically, the way in which we relate to law as a regulatory tool: the law is a self-organizing system in which an interactive network of many parts (actors and institutions) operate with no overall guiding hand.

There are several levels of analysis of law and complexity. First of all, we can recognize that there are complex adaptive system properties in the economy, poverty, war, crime and other domains the lawmakers attempt to manage and regulate through law. At a deeper level, complexity theorists assert that the legal system is a complex adaptive system, due to the fact that economic and other social systems (regulated by the legal systems) exhibit complex adaptive system properties. Then, if the economy and the legal systems are both complex adaptive systems, the two systems interact complexly with each other and also with the other complex social and physical systems connected to them.

Furthermore, if law complexly affects the economy and other systems and the economy and other systems complexly affect law, there is a high probability that law is complexly affecting itself.

Thanks to this idea of law as a complex adaptive system co-evolving with its regulatory targets, complexity theory acquires the role of a possible new approach to legal theory or, in other words, a new jurisprudence.

Jurisprudence includes various approaches such as natural law, positivism, sociological jurisprudence, realism, critical legal studies, feminist jurisprudence, postmodernist jurisprudence and critical race theory [13].

Even in such a wide overview, complexity theory adds a qualitatively new point of view that involves interconnections, systemic properties, unpredictability, porous boundaries, some element of bottom-up organization and rapid innovations in law and regulation.

The regulation can be analyzed in very different fields, such as the regulation of on-street sex work [7] or the legal complexity of the governance of global financial markets.

3 Complexity Theory and Legal Thinking

Complexity theory poses a great influence on the natural and social sciences. However, traditional legal thinking is dominated by the theory of autopoietic systems.

The theory of autopoiesis sees the law's autonomy as the self-reproduction of a communication network and define its relation to society as an interference with other autonomous communication networks.

German sociologist Niklas Luhmann developed a comprehensive theory of what he called "autopoietic" or "self-referential systems": he theorized the existence of a major social system, composed by various social subsystems, such as state, economy, science, religion, education, art, family, and law.

These ideas were originally developed in the field of biological research conducted in the early 1970s by biologists Humberto Maturana and Francisco Varela on visual cognition in frogs and pigeons. From this biological research, they sought to develop a general theory of living systems.

Luhmann developed his idea of autopoiesis as "operative closure", describing societal subsystems as closed with respect to code. The code is a distinction between two opposed values, such as true/false for the scientific system and legal/illegal for the legal system. The code defines the societal subsystem's unity and is unique to that system (at least as a code).

In Luhmann's opinion, society is mainly divided into various subsystems that perform some unique functions: economy, politics, law, science, education, religion, art, mass media, and the family.

The political and legal systems, however, are partly regionally and not simply functionally differentiated due to the fact that these subsystems are influenced by territory. The legal system, notably, is territorially differentiated into different legal orders—even if it's still possible to talk about a global legal system—without centralized legislation or decision-making capacity.

The various subsystems regularly communicate and high number of exchanges take place between the individual subsystems and the environment. The subsystems receive inputs from the surrounding environment, and these inputs are processed and converted into outputs fed back to the environment. Consequently, the outputs' effects on the environment are processed and flow back into the subsystems, completing the feedback loop.

In this framework, the systems are cognitively open but operatively closed or normatively closed.

This theory could be clearer if we consider that, in Luhmann's opinion, the societal systems and the subsystems are not the traditional associations among individuals, but a network of communications. In other words, communication is the key, since the societal systems and subsystems are networks of communications.

The presence of communications differentiates societal systems from the nonsocial environment, which is characterized by the absence of communication.

The societal systems are operatively closed because the information from the surrounding environment does not penetrate the system: information is always internally produced within the system according to internal procedures, standards, and criteria. In other words, the communications within a system only interconnect with other communications within that same system.

A system's operative closure is the basis for its autonomy or autopoiesis. The term "autopoiesis" means literally self-production. Autopoietic systems produce themselves through their own operations, their information and their structures.

In Luhmann's opinion, the idea of relative autonomy or relative autopoiesis does not exist, because there are only two alternatives: either a system is autonomous and autopoietic, or it is not.

The system's operative closure does not deny a system's dependence on its environment but leads to investigate societal systems from within and define them in terms of their own operations or, in other words, in terms of the communication that characterizes the individual societal system.

Consequently, the system's evolution is a transformation and a renewal of a system's relation to its environment.

Luhmann analyzed modern society's evolution with the help of structural couplings among functional subsystems. Through the concept of 'structural coupling', Luhmann describes evolution as coordination or coevolution without an intentional guide or direction between two systems, such as law and politics.

Luhmann's thinking aligns with the mainline of legal positivism: the law is a separate subsystem of society, distinct from politics as well as from morality. Only law—not politics, morality or any other system of communication—determines what law is.

The idea of operative closure excludes extralegal concepts and criteria from the law since they have no legal relevance. In this context it's important to note that, when these extralegal concepts and criteria influence judicial decision (for example, in constitutional decisions) they tend to be "juridified". In other words, when incorporated in the legal system, the extralegal concepts and criteria have different and specific legal meaning.

This can be identified as the biggest difference between the theory of autopoiesis and the complexity theory. As previously described, complexity theory understands the legal system as a complex adaptive system like the economy. These two systems complexly interact with each other, as well as with the other complex social and physical systems they are interconnected with.

Furthermore, complexity theory shows us how there is no perfect form of regulation available. A complex version of the law would be flexible and adaptable without, however, being able to provide a "perfect and stable" solution to regulation. Such a system would only deliver a lower number of mistakes and more adapt reactions.

This is clear if we think about the numerous references to complexity theory across all scientific disciplines, such as physics, chemistry, and biology. The literature has drawn on a wide range of sources in these disciplines to produce socially influential metaphors *inter alia*, such as the "butterfly effects" and the "tipping points".

Being an adaptive system, the law must adapt itself to the constantly changing reality that surrounds it and the continuous changing of regulations, especially in the field of economics and finance. The endless effort of adaptation to the circumstances is what lawmakers all over the world are currently experiencing.

For these reasons, complexity can lead to a better understanding of the quick evolution of economy and finance, and it represents an effective set of tools for the regulation of the financial system.

4 Complexity Theory and the Financial System

It is thought that the global financial crisis indicates the failure of neoclassical economics. This idea favors the application of complexity theory to the financial system [8].

Even before the great crisis, complexity theory has already been applied to financial systems [32]. The main idea was to use the complexity theory's approach to develop prediction models, more specifically, to try and predict stock market crashes.

In particular, the approach utilizes the features of Santa Fe complexity economics with the aim of developing a model of financial markets considering it a complex and adaptive system in constant evolution [32].

A few years later, complexity theory was, for the first time, applied to financial regulation: complexity theory is considered an important tool for better understanding, and therefore better regulating, global financial systems [18].

It's essential to point out that these approaches to financial systems were determinedly trans-disciplinary, bringing together ecologists and engineers with financial professionals and regulators. This means that the complexity theory framework validated ecology and the complexity theory of ecosystems as an instrument to rethink banking and financial systems and their regulation.

After the great crisis, Andrew Haldane was the first central banker to recognize the potential importance of complex systems theory in understanding financial systems and financial regulation. He put specific emphasis on bifurcation points and tipping points in complex financial systems and crisis prediction [15]. The current stage of development of the complexity theory paradigm for understanding financial systems and the global financial system is well delineated by Battiston et al. [4]. The complexity tools can improve the ability to avert crises or, at least, to lessen their impact and better manage them when they occur.

The application of complexity theory to the financial systems raises a daunting set of problems requiring new and innovative approaches. It's necessary to take into account a large number of interacting individual elements, such as people, companies, and countries.

Furthermore, these new approaches challenge conventional thinking: complex systems are dynamic rather than static, and they appear very difficult to predict and control. The individual elements of a system are directly influenced by the behavior of the system as a whole while, at the same time, their interactions lead to emergent behavior at the aggregate level of the system.

In particular, complex systems are permeated by non-linear or network interactions amongst the component agents. This means, *inter alia*, that the 'common sense' connection between the size of an event and its consequences are not interconnected due to the fact that small changes have the capacity to trigger large scale events.

In current times, the authors are pointing in the direction of a new complexity agenda specifically for financial regulation. They emphasize the need for lawyers to develop new useful approaches to better manage complexity [20].

5 Complexity Jurisprudence and Financial Systems Regulation

Complexity jurisprudence sees law, regulation, and management of financial systems as part of a fluid network characterized by continuous change and transformation [26].

From this point of view, the financial system configures itself as a complex adaptive system. This means that it has an adaptive capacity which is, in other words, the capacity to explore, to experiment, and to innovate.

Adaptive capacity relates to other systems and to its own self-organization, emergence, and transformations. This predisposition leads to ecological resilience as the ability to explore regimes, to systemically mutate and to create new regimes of self-organization and emergence in new dissipative structures [26].

In this approach, compliant with the complexity jurisprudence, the issues concerning regulation and management of the financial system are compared with the resilience theory of adaptive management. According to complexity jurisprudence, the resilience theory of adaptive management is the practical application of the complexity theory and deals with the complex problems of managing and governing complex adaptive systems.

The main idea is that the system's ecological resilience implies the presence of creative processes of self-organization, emergence, innovation, and adaptation in the managed system. Therefore, adaptive management considered as an outcome of the complexity jurisprudence aims to manage systems so that they operate in ecological resilience and are able to generate self-organization, emergence, innovations, and adaptations, increasing the managed system's own adaptive capacity to manage itself. The main consequence is a highly flexible regulation.

The lawmaker acts by trial and error, and the possibility to change the rules is surely kept in mind. There are many examples of this new approach of lawmaking: we can consider, for example, the Italian statutes which explicitly provide for the option of rule changes.

The adaptive management approach is very far from the creation of fixed rules made to regulate a determined phenomenon forever, as seen in the property rules inside the 1942 Italian Civil Code. Instead, adaptive management proceeds with monitoring, measuring and mapping the managed complex adaptive system, expecting regulatory interventions and modifications into the managed system.

This is the reason why adaptive management is a constant process: it needs to incessantly learn, experiment, adapt and innovate.

Furthermore, the regulatory goals can change, together with the evolving processes of regulation and management.

The real difference is not between old fixed rules and new variable rules, but it's the awareness of the continual changes and evolution of the system. Consequently, lawmaker's interventions are trial and error. Indeed, in the traditional approach, the regulation is defined as strongly front-ended while, on the contrary, adaptive management is back-ended where experimentation and changes of approach are enabled and discretion to change decisions is retained.

Flexibility is institutionalized: the organization must be polycentric, the regulatory tools must vary, and the management interventions must be multi-scalar in terms of the levels at which it can be targeted.

The core role of adaptive management is to monitor the complex adaptive system. It's necessary to focus on specific variables at different levels and be able to constantly read the data and consequently react. This manner of regulating requires the development of new tools for governing complex financial systems. Therefore, attention must be paid to the technologically enabled real-time monitoring of complex financial systems.

For this, in summary, any notions of the top-down regulation of financial markets would be abandoned and replaced with a bottom-up relational financial regulation. This new financial regulation would become somewhat experimental in its interventions and open to regulatory innovations and to new technological tools. The necessity to operate with a multilevel approach and the requirement to continuously monitor the factual circumstances involved in the regulation would characterize this approach as quite an interventionist one.

Considering the afore-described attributes of adaptive management, it's possible to estimate that, in case of a crisis of the financial system, the reaction would be very different from the one based on the traditional approach.

The traditional approach is the model according to which financial regulation aims to mitigate the systemic risk, seeking to preserve the existing financial system at all costs (bailing out market externalities). On the contrary, adaptive management would transform the organization and the operation of a financial system in crisis, due to the fact that the previous system could be socially and economically undesirable. Hence, adaptive management could promote change at various levels.

Much attention is paid to the modality in which the financial system fails. As complexity theory revealed when applied in the ecology field, the core is the understanding of the interconnections among the subjects involved, such as banks and investors. For this reason, it's necessary to focus on how information circulates through the system; this data can reveal system stability, strength, and resilience.

The application of this theory to the financial system is deeply connected with the availability of data and the development of quantifiable metrics. The regulation requires a "policy dashboard" monitoring systemic risk [4] and calling for stress-tests of the global financial system in real-time, as we do for the weather.

This "policy dashboard" for the financial system allows us to monitor it, detect its systemic risk and understand how it will evolve. Likewise, the regulatory system is a complex adaptive system (as well as the financial system, or other systems, like the environment, the Internet or health care) and the "policy dashboard" can allow the regulatory system to co-evolve alongside with the system it seeks to regulate.

Legal researchers have begun exploring such "policy dashboards" for politics. These "policy dashboards" allow conscious choices by lawmakers and, eventually, it will also allow them in politics. Quantitative studies about the complex adaptive system behaviors of a regulatory system are also starting to take place [5].

As described before, the financial system and the regulatory system are both complex adaptive systems that co-evolve and—being complex systems—have a close

interconnection. For this reason, the regulatory system should be part of the research and model-building of the financial system [30]: this regulatory approach needs to understand the financial system itself.

The main idea is that it could be very difficult to predict how the regulatory system will evolve. However, the complexity theory allows the creation of abstract models in laboratories, with a focus set on the evolution of the system. This means that complexity theory hands over tools to test the possible effects of the regulation and assess plausible alternatives which might improve regulation effectiveness [4].

Indeed, as a practical application of the complexity theory, adaptive management results as experimental in its interventions and also opens to regulatory innovations and to new technological tools.

6 The Financial System Regulation and the Systemic Risk

As previously noted, the debate about complexity theory and the regulation of the financial system has been deeply affected by the great crisis of 2008. The regulatory system is subjected to a continuous evaluation of the results achieved with the help of technological tools. This constant evaluation of the effects implies continuous learning, which could contribute to the evolution of the regulatory system in relation to systemic risk.

The understanding of the long-term evolutionary dynamics of regulation is very difficult, but one of the purposes of complex system models is measuring the effects of regulations.

This evaluation and assessment activity can be achieved before model building, with the aim to identify possible unintended consequences of prospected regulations. In this scenario, a realistic complexity-based "policy dashboard" can be helpful in empirically assessing reforms before implementing them in real markets [4].

On the other hand, continuous monitoring and testing can facilitate the process of identifying and preventing systemic risks. A recent study made by the Dutch interbank network illustrated how a realistic analysis of the financial system could identify an early warning signal up to 3 years before the actual crisis [4]. Generally speaking, recent techniques such as warning signals of tipping points offer the potential for better monitoring and regulating interconnected economic and financial systems, which may assist in anticipating and managing future crises.

The importance of the evaluation and assessment activity can be easily explained through the leverage example. A crucial factor driving the great crisis was the excessive use of leverage: the banks borrowed too much without proper collateral.

The link between leverage and the crisis was analyzed by John Geanakoplos who described the leverage cycle, in which leverage increases due to competition during times of financial stability. The excessive growth of leverage eventually leads to financial instability and triggers a crash; once the financial system is back to a condition of stability, the cycle repeats itself [14]. Leverage can drive clustered volatility and heavy tails in financial time series [33].

First of all, this situation is a clear example of how the myopic perspective of individual banks can drive systemic risk when too many banks act in unanimity. Secondly, it's possible to say that Basel III, the current regulation, is solely based on the rules of thumb and intuition.

The choice of the proper level for leverage must be made taking into consideration not only the exigence to have sounder individuals but considering all the circumstances, including relationships such as interbank lending, which can provide security in normal times but may amplify the extent of a crash in bad times. This choice must contemplate the fact that bank holding companies are highly complex institutions, formed by thousands of independent entities. Without understanding the network of ownership and control, it's virtually impossible to properly measure leverage. Furthermore, local shocks can have systemic repercussions.

Connectivity might favor disease spreading, with domino effects and bankruptcy cascades. This means that, in many situations, actions ensuring the soundness of one institution (e.g., solvency, liquidity capacity, etc.) may not be consistent with ensuring the soundness of another one [10] and may decrease the stability of the system as a whole.

The complexity theory can help in evaluating as many circumstances as possible, potentially making the regulation of leverage much more scientific than in the past. The complexity theory approach allows us to set some mechanisms that can be activated when the average leverage of the actors in a network reaches a critical threshold. The data analysis concerning the complex system affected by the leverage can provide a solid framework for making sound regulatory decisions.

The complexity approach is capable of developing new indicators, genuinely constructed with a systemic risk approach starting from microscopic data and taking into account the network of mutual exposures among institutions.

7 Complex Systems Models and the Role of Fraud

Complexity theorists show insightful trust in the ability to prevent crises. However, a number of authors argue that relying too heavily on complex system models could be a mistake, due to the fact that individuals can easily distort the model through fraud. Indeed, in the financial crisis of 2008, fraud played an important role [35].

As James Galbraith said, "the existence of a bubble in a stable, regulated market like housing is *prima facie* evidence of fraud" and the presence of fraud can alter the model and, consequently, the ability of the complex system models to prevent crises.

Speaking about the financial crisis of 2008, complexity theorists can argue that the bubble and the crisis itself would have occurred anyway, even without any fraud at all [4]. For example, the housing bubble can be generated by the growth of normal leverage levels, and it can take place without any type of fraud. Furthermore, in controlled laboratories, experimental asset markets bubbles and crashes have been frequently observed, and it was seen how they can legitimately result as the emergent outcome of the circumstances [31].

Analyzing real facts, in the 2008 financial crisis the troubles were generated by the legal use of excessively high leverage, which generated systemic risk.

Consequently, the role of fraud has to be considered as not crucial. Fraud may, of course, increase the instabilities and intensify the signals of the crisis, but it's not the primary driving factor, as these instabilities are an emergent outcome of complex financial networks [3].

Furthermore, fraud can be included in the model. The complex system model, created with the aim of preventing crises, can contain agents who "cheat" the system by not following accepted sets of rules in their behavior [4].

In conclusion, especially after the great crisis, there is a deep interest in using complex system models, with the aim of forecasting the long-lasting effects of the regulation system and to predict a potential collapse of the financial system.

Complexity theory introduces in the debate ideas and concepts that are new for law scholars and also urges the institutions and the scholars themselves to analyze the reality in overall terms, involving at once legal skills, system modeling, economic skills and the knowledge of the technological tools. Proof of this is the implementation, in the financial and regulatory lexicon, of concepts such as tipping points, networks, contagion, feedback, and resilience.

Currently, the use of complexity models and results lingers at an early stage, but it noticeably indicates to scholars that a new multidisciplinary approach is needed in dealing with the issues concerning the regulation of financial markets.

References

1. Arthur, B.W.: Increasing Returns and Path Dependence in the Economy. University of Michigan Press, Ann Arbor (1994)
2. Arthur, B.W.: Complexity and the Economy. Oxford University Press, Oxford (2014)
3. Bardoscia, M., Battiston, S., Cacciola, F., Caldarelli, F.: Pathways towards instability in financial networks (2016). http://arxiv.org/abs/1602.05883
4. Battiston, S., Doyne Farmer, J., Flache, A., Garlaschelli, D., Haldane, A.G., Heesterbeek, H., Hommes, C., Jaeger, C., May, R., Scheffer, M.: Complexity theory and financial regulation. Science **351**, 818 (2016)
5. Boulet, R., Mazzega, P., Bourcier, D.: AI Approaches to the Complexity of Legal Systems, pp. 39–53. Springer (2010)
6. Capra, F.: The Systems View of Life: A Unifying Vision. Cambridge University Press, Cambridge (2016)
7. Carline, A., Murray, J.: Reconceptualizing on-street sex work as complex affective social assemblage. In: Fitzgerald, S.A., McGarry, K. (eds.) Realising Justice for Sex Workers: An Agenda for Change. Rowman & Littlefield, London (2018)
8. Colander, D., Goldberg, M., Haas, A., Juselius, K., Kirman, A., Lux, T., Sloth, B.: The financial crisis and the systemic failure of the economics profession. Crit. Rev. 249–267 (2009)
9. Coveney, P., Highfield, R.: Frontiers of Complexity: The Search for Order in a Chaotic World. Faber & Faber, London (1996)
10. Crockett, A.: Marrying the micro- and macro-prudential dimensions of financial stability. In: Bank of International Settlements Eleventh International Conference of Bank Supervisors Basel (2000). www.bis.org/speeches/sp000921.htm

11. Deleuze, G., Guattari, F.: A Thousand Plateaus. Capitalism and Schizophrenia. University of Minnesota Press, Minneapolis, MN (1987)
12. Fafinsky, S., Finch, E.: English Legal System. Pearson Education ltd. (2007)
13. Freeman, M.: Lloyd's Introduction to Jurisprudence, 9th edn. Sweet & Maxwell, London (2014)
14. Geanakoplos, J.: The Leverage Cycle. Cowles Foundation Discussion Paper. N. 1715R (2010). SSRN: https://ssrn.com/abstract=1539483 or https://doi.org/10.2139/ssrn.1539483
15. Haldane, A.G.: Rethinking the financial network (2009). https://www.bankofengland.co.uk/speech/2009/rethinking-the-financial-network
16. Holland, J.H.: Emergence: From Chaos to Order. Perseus Books, Cambridge, MA (1998)
17. Holland, J.H.: Hidden Order: How Adaptation Builds Complexity. Addison-Wesley, New York, NY (1995)
18. Kambhu, J., Weidman, S., Krishnan, N.: New Directions for Understanding Systemic Risk (2007). https://www.nap.edu/read/11914/chapter/1
19. Kauffman, S.: The Origins of Order: Self-Organization and Selection in Evolution. Oxford University Press, New York (1993)
20. Lippe, P., Katz, D.M., Jackson, D.H.: Legal by design: a new paradigm for handling complexity. In: Banking Regulation and Elsewhere in Law. SSRN Electronic Journal (2014). https://www.researchgate.net/publication/315432799_Legal_by_Design_A_New_Paradigm_for_Handling_Complexity_in_Banking_Regulation_and_Elsewhere_in_Law
21. Luhmann, N.: Operational closure and structural coupling: the differentiation of the legal system. Cardozo Law Rev. **13**, 1419 (1992)
22. Luhmann, N.: Law as a Social System, trans. Klaus A. Ziegert. Oxford University Press, Oxford (2004)
23. Maturana, H.R., Varela, F.J.: The Tree of Knowledge: The Biological Roots of Human Understanding. Shambhala Publications, Boston (1987)
24. Morin, E.: Restricted complexity, general complexity. In: Gershenson, C., Aerts, D., Edmonds, B. (eds.) Worldviews, Science and Us: Philosophy and Complexity. World Scientific, Singapore (2007)
25. Murphy, J.T.: Complexity Theory. Oxford Bibliographies (2017). www.oxfordbibliographies.com
26. Murray, J., Webb, T., Wheatley, S. (eds.): Complexity Theory and Law: Mapping an Emergent Jurisprudence. Routledge (2018)
27. Pagallo, U.: As law goes by: topology, ontology, evolution. In: Casanovas, P., Pagallo, U., Sartor, G., Ajani, G.: AI Approaches to the Complexity of Legal Systems, p. 14. Springer (2010)
28. Prigogine, I., Stengers, I.: Order Out of Chaos: Man's New Dialogue with Nature. Flamingo, London (1984)
29. Ruhl, J.B., Katz, D.: Measuring, monitoring, and managing legal complexity. Iowa Law Rev. **101**, 191 (2015)
30. Ruhl, J.B.: Financial complexity: regulating regulation. Science **352**, 301 (2016)
31. Smith, V.L., Suchanek, G.L., Williams, A.W.: Bubbles, crashes, and endogenous expectations in experimental spot asset markets. Econometrica **56**, 1119 (1988)
32. Sornette, D.: Why Stock Markets Crash (Critical Events in Complex Financial Systems). Princeton University Press (2003)
33. Thurner, S., Doyne Farmer, J., Geanakoplos, J.: Leverage causes fat tails and clustered volatility. Quant. Finance **12**(5), 695–707 (2012)
34. Waldrop, M.M.: Complexity: The Emerging Science at the Edge of Order and Chaos, 2nd edn. Penguin Books, London (1994)
35. Witzling, D.: Financial complexity: accounting for fraud. Science **352**, 301 (2016)

The Emergence of the Order Parameter in the Interpolating Replica Trick for Disordered Statistical Mechanics Systems

Francesco Guerra

Abstract We recall the formulation of the so called replica trick, in the theory of disordered statistical mechanics systems, based on interpolation on the number of replicas, given in previous work. While the customary exploitation of the replica method is based on analytic continuation on the number of replicas, we have proposed a completely different and complementary strategy where interpolation on the number of replicas plays the key role. As examples, we consider the case of the Sherrington-Kirkpatrick mean field model for spin glasses and the case of the Derrida random field model. Firstly we recall the case of an integer number of replicas, which can be explicitly solved, by proving that the annealed free energy of the replicated system can be expressed through variational principles, where definite variational trial functions enter, as functions of special order parameters, depending on the number of replicas. Then we show how to interpolate on the number of replicas, taking now any real positive value. The general scheme is described by an auxiliary system of functions depending on the physical parameters involved and on the "number of replicas" s now considered as a positive real number. The existence and the main properties of these functions are proven through standard interpolation techniques. The case of the usual thermodynamic systems is reached in the limit when the number of replicas s goes to zero. By working firstly in the simple case of the random energy model we prove a very significant fact. While for each integer s the auxiliary functions are naturally expressed through variational principles involving trial functions and variational order parameters depending on the number of replicas s, it is shown that for generic real values of s the whole structure can be embedded in a unique variational principle, where there is a universal trial function and a trial order parameter which do not depend on s. In order to have the values of the auxiliary functions for each real value of s, it is only necessary to put a suitable bound on the order parameter, depending on s, during the variational procedure.

F. Guerra (✉)
Department of Physics, Sapienza University of Rome, Rome, Italy
e-mail: francesco.guerra@roma1.infn.it

National Institute for Nuclear Physics, Section of Rome, Rome, Italy

Historical Museum of Physics and the Enrico Fermi Study and Research Center, Rome, Italy

S. Albeverio et al. (eds.), *Complexity and Emergence*, Springer Proceedings in Mathematics & Statistics 383, https://doi.org/10.1007/978-3-030-95703-2_3

The method extends also to the physically significant zero replica limit. This general strategy is extended also to the Sherrington-Kirkpatrick case, where the universal variational principle assumes the well know ultrametric structure, with a functional order parameter. It is important to notice that, according to this strategy, the universal variational principle, and the order parameter arise only from considerations based on the integer values of the number of replicas. Contrary to the traditional treatment, where replica symmetry breaking is connected with the possibility of multiple analytic continuations from the integer values towards the zero replica value, in the present formulation replica symmetry breaking is found to be a phase transition in the number of replicas. The method can be easily extended to multispecies models, as for example the Hopfield model of neural networks.

Keywords Disordered models · Sherrington-Kirkpatrick · Derrida · Hopfield · Mean field spin glass · Random energy · Neural networks · Interpolation · Functional order parameter · Variational principles

1 Introduction

Let us recall some essential aspects of the replica trick as presented for example in [1].

We consider simple disordered systems of statistical mechanics, as the Sherrington-Kirkpatrick spin-glass mean field model [2], and associated models, and the Derrida Random Energy Model [3].

We consider N spin Ising configurations

$$\sigma : (1, 2, \ldots, N) \ni i \to \sigma_i = \pm 1. \tag{1}$$

There are 2^N Ising configurations on N sites. We will be interested in the limit $N \to \infty$.

For each configuration σ let us introduce (real valued) random variables $\sigma \to \mathcal{K}(\sigma)$. In the simplest cases, we can assume these 2^N variables as Gaussian, with zero averages, and covariances given for example by

$$\mathbb{E}(\mathcal{K}(\sigma)\mathcal{K}(\sigma')) = q^2_{\sigma\sigma'}, \tag{2}$$

where $q_{..}$ are the configuration overlaps defined by

$$q_{\sigma\sigma'} = \frac{1}{N}\sum_{i=1}^{N} \sigma_i \sigma'_i, \tag{3}$$

in the Sherrington-Kirkpatrick model.

The random variables $\mathcal{K}$ can be simply realized in a unique well defined probability space. Let us introduce the normalized independent unit, i.e. mean 0 and variance 1, Gaussian random variables J_0, $J_{ij}, i = 1, 2, \ldots, j = 2, 3, \ldots, i < j$. Then we can define

$$\mathcal{K}(\sigma) = \frac{1}{\sqrt{N}} J_0 + \frac{\sqrt{2}}{N} \sum_{i<j} J_{ij} \sigma_i \sigma_j, \tag{4}$$

and easily verify that the correct covariance properties are valid.

On the other hand, in the case of the Derrida Random Energy Model, we define $q_{\sigma\sigma'} = 1$ if the two configurations are equal and $q_{\sigma\sigma'} = 0$ if they are different, i.e. $q_{\sigma\sigma'} = \delta_{\sigma\sigma'}$.

The random variables $\mathcal{K}(\sigma)$ are exploited to define the energy associated to each configuration σ in the form

$$\mathcal{H}(\sigma) = -\sqrt{\frac{N}{2}} \mathcal{K}(\sigma), \tag{5}$$

where the term $\sqrt{N}$ is introduced for serious thermodynamic reasons, as it will be shown in the following.

In the well known Boltzmann-Gibbs scheme, the partition function is

$$Z_N(\beta) = \sum_{\sigma} \exp(-\beta \mathcal{H}(\sigma)) = \sum_{\sigma} \exp(\beta \sqrt{\frac{N}{2}} \mathcal{K}(\sigma)), \tag{6}$$

where β is the inverse of the temperature.

We have performed the sum over all configurations. Therefore, the partition function does depend only on the random noise present in the $\mathcal{K}(\sigma)$'s.

The (random) free energy $F_N(\beta)$ is defined by

$$-\beta F_N(\beta) = \log Z_N(\beta). \tag{7}$$

The rescaling $\sqrt{N}$ in the definition (5) of the energy is introduced in order to assure a good thermodynamic behavior for the free energy per site, in the limit $N \to \infty$.

In fact, it is not difficult to prove, see for example [4], that the limit

$$\lim_{N\to\infty} \frac{1}{N} \log Z_N(\beta) \tag{8}$$

does exist almost surely in the probability space where all $\mathcal{K}(\sigma)$ are defined. We call $A(\beta)$ this limit, where any random character has been lost.

It turns out that the limit $A(\beta)$ can be calculated also through the quenched averages

$$A(\beta) = \lim_{N\to\infty} \frac{1}{N} \mathbb{E} \log Z_N(\beta), \tag{9}$$

where $\mathbb{E}$ is the average expectation with respect to the noise due to the $\mathcal{K}(\sigma)$'s.

The equality between the probabilistic limit and the quenched limit is due to a moderate statistical fluctuation of the free energy in the limit, which can be easily proved through elementary interpolation methods [4].

There is a deep physical motivation at the basis of the metallurgic terminology. In the partition function $Z_N(\beta)$ we perform only the sum over the σ's, according to Boltzmann prescriptions. Therefore, the noise in the $\mathcal{K}(\sigma)$'s acts as external noise, which is not involved in the thermodynamic equilibrium, but affects thermodynamic equilibrium of the σ'. Then, we take the log, and at the end the average $\mathbb{E}$.

Obviously we can take also the (annealed) average, before taking the log, so that the external noise does participate to the thermodynamic equilibrium

$$\bar{A}(\beta) = \lim_{N\to\infty} \frac{1}{N} \log \mathbb{E} Z_N(\beta). \tag{10}$$

This annealed expression is easily calculated

$$\mathbb{E} Z_N(\beta) = \mathbb{E} \sum_{\sigma} .. = \sum_{\sigma} \mathbb{E}.. = \exp(N(\log 2 + \frac{1}{4}\beta^2)), \tag{11}$$

since for each σ we have

$$\mathbb{E} \exp(\beta \sqrt{\frac{N}{2}} \mathcal{K}(\sigma)) = \exp(\frac{1}{2}\beta^2 \frac{N}{2} \mathbb{E}(\mathcal{K}^2(\sigma))) = \exp(\frac{1}{4}\beta^2 N). \tag{12}$$

The term $\log 2$ comes from the final sum over the σ's.

Of course, the annealed expression $\bar{A}(\beta)$ is not correct in general. In any case it is a rigorous upper bound, uniform in N. In fact, from the concavity of the log

$$\mathbb{E} \log .. \leq \log \mathbb{E}.. \tag{13}$$

we have

$$\frac{1}{N} \mathbb{E} \log Z_N(\beta) \leq \frac{1}{N} \log \mathbb{E} Z_N(\beta) = \log 2 + \frac{1}{4}\beta^2, \tag{14}$$

preserved in the limit

$$A(\beta) = \lim_{N\to\infty} \frac{1}{N} \mathbb{E} \log Z_N(\beta) \leq \log 2 + \frac{1}{4}\beta^2. \tag{15}$$

We are interested in the explicit expression for $A(\beta)$ (given by (8), (9)) in the form of a variational principle.

The paper is organized as follows. In Sect. 2 the concept of replicas is recalled. Then we introduce the auxiliary functions connected with the annealed averages of the replicated free energy, properly normalized, and study their infinite volume limit. These are expressed through variational principles, with appropriate order parameters

and trial functions. The key aspect of the replicated theory, for any integer value for the number of replicas, is the emergence of replica symmetry in the infinite volume limit. Therefore the variational principle can be recast in a simpler replica symmetric formulation. In Sect. 3 we extend our scheme from an integer number of replicas to the case where the number of replicas is formally given by a real number. Next Sect. 4 deals with the random energy model. We show how to embed the model in a general structure where all replicas are involved. We show how the variational principle is emerging. In Sect. 5 we report about the main results of this paper. We consider the Sherrington-Kirkpatrick mean field spin glass model and show how the functional order parameter is emerging together with the trial functional. Our treatment is extremely elementary and pedagogical. Everything is developed from the structure based on real replicas, through simple progressive steps. Finally Sect. 6 is dedicated to some outlook for future development, especially concerning multispecies models and neural networks.

This paper is an extended version of lectures presented at the Lake Como School of Advanced Studies, Villa del Grumello, July 22–27, 2018. We warmly thank the Organizing Committee, Sergio Albeverio, Elisa Mastrogiacomo, Emanuela Rosazza Gianin and Stefania Ugolini, for the kind invitation, the excellent organization, and the very stimulating atmosphere.

2 Replicas and the Auxiliary Function

Let us introduce the concept of *replicas*. For $s = 1, 2, \ldots$ (s positive integer) the s-replicated system has a configuration space which is the s time product space of the original system. Therefore, now the variables are

$$\sigma_i^a = \pm 1,\ i = 1, 2, \ldots, N,\ a = 1, 2, \ldots, s, \tag{16}$$

where the index i denotes the sites, and the index a denotes the replicas. Therefore, now overall there are sN sites.

The energy is now defined as the sum of the energies for each single replica, with the *same* randomness. The Boltzmann factor is therefore factorized, and for the new partition function we have a simple product

$$\begin{aligned} \bar{Z}_{s,N}(\beta) &= \sum_{\sigma^1} \exp(\beta\sqrt{\frac{N}{2}}\mathcal{K}(\sigma^1)) \ldots \sum_{\sigma^s} \exp(\beta\sqrt{\frac{N}{2}}\mathcal{K}(\sigma^s)) \\ &= Z_N^s(\beta), \end{aligned} \tag{17}$$

since (by assumption) every sum on the variables σ's gives the same contribution. Therefore the partition function of the new system is simply the product of identical terms corresponding to the partition function of the original nonreplicated system.

The free energy per site, and its quenched average, of the replicated system is the same as for the original system. In fact, the logarithm of a product is the sum of the logarithms, each with the same contribution

$$\log \bar{Z}_{s,N}(\beta) = s \log Z_N(\beta). \tag{18}$$

Therefore, trivially

$$\frac{1}{sN} \log \bar{Z}_{s,N}(\beta) = \frac{1}{N} \log Z_N(\beta). \tag{19}$$

However, if we take the annealed expressions, as in (10), we have a nontrivial dependence on the number of replicas s. We are induced to introduce an auxiliary function

$$\phi_N(s, \beta) := \frac{1}{sN} \log \mathbb{E}\bar{Z}_{s,N} = \frac{1}{sN} \log \mathbb{E} Z_N^s(\beta), \quad s = 1, 2, \ldots, \tag{20}$$

with a deep motivation made explicit in the following. Obviously, if $s = 1$ we have simply the annealed case, considered above.

It is easy to study the thermodynamic limit $N \to \infty$ of $\phi_N(s, \beta)$, with very interesting results.

The limit, $\phi(s, \beta)$, does exist for any integer $s = 1, 2, \ldots$, and can be explicitely expressed through a variational principle. We only give the general structure. We have to specify the order parameters, and the trial function.

Let us consider firstly the Sherrington-Kirkpatrick model given by (2), (3).

Firstly we calculate the annealed average in the corresponding replicated model. Let us start from

$$\mathbb{E} Z_N^s(\beta) = \mathbb{E} \sum_{(\sigma_i^a)} \exp(\beta \sqrt{\frac{N}{2}} \sum_{a=1}^{s} \mathcal{K}(\sigma^a)) = \sum_{(\sigma_i^a)} \exp(\frac{1}{2}\beta^2 \frac{N}{2}(s + 2 \sum_{(a,b)} q_{a,b}^2)), \tag{21}$$

where we have written explicitely the product Z^s in terms of the sum $\sum_{(\sigma_i^a)}$ on the replicated site variables, have interchanged this sum with the average $\mathbb{E}$, and performed $\mathbb{E}$ by using the standard formula for Gaussian random variables $\mathbb{E} \exp \xi = \exp \frac{1}{2}\mathbb{E}\xi^2$. In the expression $\mathbb{E}(\sum_{a=1}^{s} \mathcal{K}(\sigma^a))^2$, there arise s diagonal terms, each equal to 1 and $s(s-1)$ off diagonal terms each giving rise to a term q_{ab}^2, for each couple of variables (ab). Of course there are $s(s-1)/2$ couples.

By putting this expression in the definition of $\phi_N(s, \beta)$ we have

$$\phi_N(s, \beta) = \frac{\beta^2}{4} + \frac{1}{sN} \log \sum_{(\sigma_i^a)} \exp(\frac{1}{2}\beta^2 N \sum_{(a,b)} q_{a,b}^2). \tag{22}$$

We see that the annealed averages in $\phi_N(s, \beta)$ have a very familiar form of ferromagnetic systems on the σ_i^a Ising variables, $a \in 1, \ldots, s$. By using well known standard techniques, as described for example in [4], we arrive easily at the expression of

the infinite volume limit $\phi(s,\beta) = \lim_{N\to\infty} \phi_N(s,\beta)$ in the form of a variational principle

$$\phi(s,\beta) = \sup_{q_{(..)}} \tilde{\phi}(s,\beta; q_{..}), \tag{23}$$

where the order parameters $q_{(..)}$ and the trial function $\tilde{\phi}(s,\beta; q_{..})$ are defined as follows.

For a given integer s, for each couple of replicas we introduce the system of order parameters $q_{ab} \geq 0$. There are $s(s-1)/2$ order parameters. The case $s = 1$ does not require any order parameter, and the auxiliary function $\phi(1,\beta)$ is given directly in the form

$$\phi(1,\beta) = \phi_N(1,\beta) = \log 2 + \frac{1}{4}\beta^2. \tag{24}$$

For $s = 2$ there is only one order parameter q_{12}. For $s = 3$ three order parameters appear q_{12}, q_{13}, q_{23}, and so on. Notice that here, for the sake of simplicity, we have adopted the notation q_{ab} for the order parameters. The same notation was exploited for the σ overlaps appearing in (21). But we hope that there is no risk of confusion, due to the different context.

The trial function $\tilde{\phi}(s,\beta; q_{..})$ is a function replica symmetric in the $q_{..}$, explicitely given by

$$\tilde{\phi}(s,\beta; q_{..}) = \frac{\beta^2}{4} + \frac{1}{s}\log \sum_{\sigma_1,\dots\sigma_s} \exp(\beta \sum_{(ab)} q_{ab}\sigma_a\sigma_b) - \frac{\beta^2}{4s}\sum_{(ab)} q_{ab}^2. \tag{25}$$

We see that the trial function involves an s-site Ising model, with a disordered two body ferromagnetic interaction given by the trial parameters q_{ab}. If the replicas are permuted its value clearly does not change. We have already said that the variational principle states

$$\lim_{N\to\infty} \phi_N(s,\beta) = \phi(s,\beta) = \sup_{q_{..}} \tilde{\phi}(s,\beta; q_{..}). \tag{26}$$

The variational principle enjoys a remarkable property. In fact, the sup is realized for values of the order parameters, where all $q_{..}$ have the same value $\bar{q} \geq 0$. There is full replica symmetry for the optimal values. This can be shown as follows.

Let us define the value $\tilde{\phi}_{RS}(s,\beta;\bar{q})$ of the trial function $\tilde{\phi}(s,\beta; q_{..})$ in the case when all $q_{..}$ assume the same value $\bar{q}$ by

$$\tilde{\phi}_{RS}(s,\beta;\bar{q}) := \tilde{\phi}(s,\beta; q_{..} = \bar{q}), \tag{27}$$

(the suffix RS standing for "replica symmetric"). An easy direct calculation shows

$$\tilde{\phi}_{RS}(s,\beta;\bar{q}) = \log 2 + \frac{1}{s}\log \int (\cosh(\beta\sqrt{\bar{q}}z))^s \, d\mu(z)) + \frac{\beta^2}{4}(1 - 2\bar{q} - (s-1)\bar{q}^2), \tag{28}$$

where $d\mu(z)$ is the unit Gaussian measure on the real line.

Obviously we have

$$\sup_{\bar{q}} \tilde{\phi}_{RS}(s, \beta; \bar{q}) \leq \sup_{q_{..}} \tilde{\phi}(s, \beta; q_{..}), \tag{29}$$

because constraints make worse a variational trial. But the inverse inequality can be easily obtained through a stability argument arising from the ferromagnetic nature of the interaction. In fact, the following bound holds in general (for $s \neq 1$)

$$\tilde{\phi}(s, \beta; q_{..}) \leq \frac{2}{s(s-1)} \sum_{(ab)} \tilde{\phi}_{RS}(s, \beta; q_{ab}). \tag{30}$$

Therefore

$$\tilde{\phi}(s, \beta; q_{..}) \leq \sup_{\bar{q}} \tilde{\phi}_{RS}(s, \beta; \bar{q}) = \phi_{RS}(s, \beta), \tag{31}$$

where the last equality is nothing but the definition of $\phi_{RS}(s, \beta)$. By putting together all bounds, we have in the infinite volume limit (i.e. $N \to \infty$)

$$\lim_{N\to\infty} \phi_N(s, \beta) = \phi(s, \beta) = \sup_{\bar{q}} \tilde{\phi}_{RS}(s, \beta; \bar{q}) = \phi_{RS}(s, \beta), \tag{32}$$

for all integer values $s = 1, 2, 3, \ldots$.

For a simple sketch of the proof of (30) we refer to [1], where we follow a standard method exploited in the statistical mechanics of disordered ferromagnetic systems in order to "tame the disorder", see also the analogous treatment in [14]. Of course other simple proofs of the replica symmetry for integer values of s are possible.

3 Interpolating on the Number of Replicas

There exists a deep and complex successive development, conventionally called "replica trick", introduced by the pioneers of the study of these systems, as shown for example in [2, 5–7]. The aim of the replica trick is to show that all properties of the system are in some way derived from the treatment at integer values of s. Here we give a formulation of the replica trick in the frame of replica interpolation. The first step is to extend the definition of the auxiliary function $\phi_N(s, \beta)$, from the integers $s = 1, 2, \ldots$ to any real value of s. For the sake of simplicity we consider only the case $s > 0$.

This extension is easily obtained by noticing that the very definition

$$\phi_N(s, \beta) := \frac{1}{sN} \log \mathbb{E} Z_N^s(\beta), \tag{33}$$

originally introduced only for integer values of s, in the replica frame, has a perfect rigorous meaning also for any $s > 0$. For a study of $\phi_N(s, \beta)$, based on large deviation techniques, we refer to [8], see also [9].

For the β derivative, we easily find, through a direct calculation involving also integration by parts on the Gaussian noise,

$$\frac{\partial}{\partial \beta}\phi_N(s, \beta) = \frac{\beta}{2}(1 + (s-1) < q^2_{\sigma\sigma'} >), \tag{34}$$

for an appropriately defined average $<>$, involving two replicas

$$< q^2_{\sigma\sigma'} >= \mathbb{E}(Z^s \Omega(q^2_{\sigma\sigma'}))/\mathbb{E}(Z^s). \tag{35}$$

Here Ω is the Boltzmann-Gibbs average for two replicas

$$\Omega(F(\sigma, \sigma')) := Z^{-2} \sum_{\sigma\sigma'} F(\sigma, \sigma') \exp(\beta \sqrt{\frac{N}{2}}(\mathcal{K}(\sigma) + \mathcal{K}(\sigma')), \tag{36}$$

for any (real valued) function F depending on the spins of the two replicas.

Notice the presence of the term $(s-1)$. Its sign changes at $s = 1$. It is responsible of many notable inversions, as for example superadditivity in N versus subadditivity, and inf versus sup in variational principles.

According to a well established and suggestive tradition, we continue to consider the real number s as the "number of replicas", even when s is not integer.

The hope is that the experience accumulated in the study of $\phi_N(s, \beta)$, and its limit $\phi(s, \beta)$, for integer s, can produce some information in the case of generic values $s > 0$.

The interest in generic values of s is a deep aspect of the "replica trick". As a matter of fact, we find that in the limit $s \to 0$ the auxiliary function ϕ_N reduces to the quenched value

$$\lim_{s\to 0} \phi_N(s, \beta) = \frac{1}{N}\mathbb{E} \log Z_N(\beta), \tag{37}$$

a very important relation which holds also in the thermodynamic limit $N \to \infty$

$$\lim_{s\to 0} \phi(s, \beta) = A(\beta), \tag{38}$$

(with $A(\beta)$ defined as in (8) and (9)). Therefore, the auxiliary function ϕ, for very small values of s, reduces to the expression of the quenched free energy (corresponding to (9)) we are interested in.

Here is a simple intuitive proof. For small values of s we have

$$\mathbb{E}Z^s_N(\beta) = \mathbb{E} \exp(s \log Z_N) \simeq 1 + s\mathbb{E} \log Z_N, \tag{39}$$

where only the dominating term to the first power of s has been considered, and therefore

$$\log \mathbb{E} Z_N^s \simeq \log(1 + s\mathbb{E} \log Z_N) \simeq s\mathbb{E} \log Z_N. \tag{40}$$

After division by s, as requested by the definition (33), in the limit $s \to 0$ only the contribution of the considered dominating terms to the first power of s survives.

It is important to start the discussion of the "replica trick" from the auxiliary function $\phi_N(s, \beta)$ as defined in (33), and not from the apparently similar expression

$$\frac{Z_N^s - 1}{Ns}, \tag{41}$$

which well reproduces $\log Z_N/N$ if the limit $s \to 0$ is taken firstly, but does surely lead to disasters if it is the limit $N \to \infty$ which is taken firstly. Some vigorous apotropaic finger crossing, or equivalent gesture, is needed in this case, as recommended by Michel Talagrand in [8], in the hope that by proceeding formally at the end something significant will be achieved. The rigorous foundation of the replica trick necessarily requires the auxiliary function $\phi_N(s, \beta)$, whose extension to all values $s > 0$ is well founded and well motivated.

By a systematic exploitation of the general interpolation methods [4, 10, 11] it is easy to establish the following important properties of $\phi_N(s, \beta)$, as recalled in [1].

It turns out that $N\phi_N(s, \beta)$ is subadditive in N for $s \geq 1$

$$N\phi_N(s, \beta) \leq N_1\phi_{N_1}(s, \beta) + N_2\phi_{N_2}(s, \beta), \tag{42}$$

for $N = N_1 + N_2$, and superadditive for $s \leq 1$

$$N\phi_N(s, \beta) \geq N_1\phi_{N_1}(s, \beta) + N_2\phi_{N_2}(s, \beta). \tag{43}$$

The proof is immediately obtained through an interpolation argument, according to the strategy introduced in [10], by comparing a system with N sites with two subsystems of N_1 and N_2 sites, with $N = N_1 + N_2$.

The thermodynamic limit $N \to \infty$ follows in the form

$$\phi(s, \beta) = \lim_{N\to\infty} \phi_N(s, \beta) = \inf_N \phi_N(s, \beta), \tag{44}$$

for $s \geq 1$, and

$$\phi(s, \beta) = \lim_{N\to\infty} \phi_N(s, \beta) = \sup_N \phi_N(s, \beta), \tag{45}$$

for $s \leq 1$.

The functions $\phi_N(s, \beta)$ and $\phi(s, \beta)$ are monotone nondecreasing in the parameter s

$$\phi(s, \beta) \leq \phi(s', \beta) \quad for \quad s \leq s', \tag{46}$$

and convex in $1/s$, for any value of β. These properties easily follows as a consequence of Hölder inequality, as shown in [1].

Moreover, we can also prove that the functions $\phi_N(s, \beta)$ and $\phi(s, \beta)$ are convex, not only in $1/s$, but also in s. The proof is very simple if we exploit the deep stochastic variational representations as introduced by Michelle Boué and Paul Dupuis in [12]. Let $W(t)$, $0 \leq t \leq 1$, be the standard $N(N-1)/2$ dimensional Brownian motion, with components $W_{ij}(t)$, so that $W_{ij}(1) = J_{ij}$, where J_{ij} are the unit Gaussian random variables appearing in the definition of the partition function (2), (6). Call $Z(W(1))$ the partition function. Then the following variational representation holds

$$\log \mathbb{E} Z^s = \log \mathbb{E} \exp s \log Z(W(1)) =$$
$$\sup_v \mathbb{E}\big(s \log Z(W(1) + \int_0^1 v(t))\, dt) - \frac{1}{2} \int_0^1 \|v(t)\|^2\, dt\big), \tag{47}$$

where the sup is taken over all $N(N-1)/2$ dimensional processes, with components $v_{ij}(t)$, which are progressively measurable with respect to the filtration generated by the Brownian motion W, and

$$\|v(t)\|^2 = \sum_{(ij)} v_{ij}^2(t). \tag{48}$$

From this representation we immediately derive

$$\frac{1}{s} \log \mathbb{E} Z^s = \sup_u \mathbb{E}\big(\log Z(W(1) + s \int_0^1 u(t))\, dt) - \frac{1}{2} s \int_0^1 \|u(t)\|^2\, dt\big), \tag{49}$$

through a simple rescaling $v(t) = su(t)$. Finally we note that the right hand side is convex in s since $\log Z$ is convex in its arguments, and the other term is linear in s. We see that the convexity in s rests essentially on the convexity of $\log Z$ with respect to the Jij's, a condition of immediate thermodynamic meaning.

The functions $\phi_N(s, \beta)$ and $\phi(s, \beta)$ are convex in β, for any fixed value of s. Here the proof is elementary and does not involve subtle properties based on the Ghirlanda-Guerra identities [13]. As a matter of fact, in general we can see that for any random interaction associated to atoms, $i = 1, 2, \ldots, K, i \to A_i$, with partition function $Z = \sum_i \exp(\beta A_i)$, the expression $\log \mathbb{E}(Z^s)$, for any $s > 0$, is convex in β. Here $\mathbb{E}$ denotes the average with respect to all random content in the A_i's. Through a simple calculation firstly we find for the derivative

$$\frac{\partial}{\partial \beta} \log \mathbb{E}(Z^s) = \frac{1}{\mathbb{E}(Z^s)} \mathbb{E}(s Z^{s-1} \sum_i A_i \exp(\beta A_i))$$
$$= \frac{s}{\mathbb{E}(Z^s)} \mathbb{E}(Z^s \Omega(A)) = s \mathbb{E}'(\Omega(A)), \tag{50}$$

where we have introduced the Boltzmann average

$$\Omega(A) := \frac{1}{Z} \sum_i A_i \exp(\beta A_i), \tag{51}$$

and the deformed $\mathbb{E}'$ average

$$\mathbb{E}'(\cdot) := \mathbb{E}(Z^s \cdot)/\mathbb{E}(Z^s). \tag{52}$$

Now we take a second derivative in β. Through a simple calculation, we find

$$\frac{\partial^2}{\partial \beta^2} \log \mathbb{E}(Z^s) = s\mathbb{E}'(\Omega(A^2) - \Omega^2(A)) + s^2(\mathbb{E}'\Omega^2(A) - (\mathbb{E}'\Omega(A))^2). \tag{53}$$

Since the two terms in the sum are obviously nonnegative, convexity in β is established. It is important to resist to the temptation to resort to integration by parts, for example if the A_i's are Gaussian. In this case the resulting expression, of the Ghirlanda-Guerra type [13], will be very difficult to handle, as seen for example in [8].

4 The Replica Trick in the Random Energy Model and the Emergence of the Universal Order Parameter

In order to explore the potentialities of the "replica trick", now we shift to the laboratory of the Random Energy Model [3]. Here we will show, by following [1] that the new interpretation of the "trick", not based on analytic continuation for $s \to 0$, gives the right order parameter, the right trial function and the right variational principle, for any value of s, starting only from the elementary variational principle at integer values $s = 1, 2, 3, \ldots$.

We show that in this case the replica symmetry is minimally broken. The deep reason for spontaneous replica symmetry breaking arises quite naturally.

Let us recall the expressions of the partition function and the auxiliary function in the Random Energy Model

$$Z_N(\beta) := \sum_\sigma \exp(\beta \sqrt{\frac{N}{2}} J(\sigma)), \tag{54}$$

$$\mathbb{E}(J(\sigma)J(\sigma')) = \delta_{\sigma\sigma'}, \tag{55}$$

so that the $J(\sigma)$'s are independent centered unit Gaussian random variables,

$$\phi_N(s, \beta) := \frac{1}{Ns} \log \mathbb{E}(Z_N(\beta)^s). \tag{56}$$

Recall that at $s = 1$ we have the annealed value

$$\phi(1, \beta) = \log 2 + \frac{1}{4}\beta^2. \tag{57}$$

Firstly let us establish the variational principle for integer values of s. We have by (56)

$$\phi_N(s, \beta) = \frac{1}{Ns} \log \mathbb{E}(Z_N(\beta)^s) = \frac{1}{Ns} \log \mathbb{E} \sum_{\sigma_i^a} \exp(\beta \sqrt{\frac{N}{2}}(J(\sigma^1) + \cdots + J(\sigma^s))). \tag{58}$$

Now we can exchange freely the $\mathbb{E}$ and the $\sum$ (for any fixed $N \in \mathbb{N}$). Therefore we are led to the calculation of

$$\mathbb{E} \exp(\beta \sqrt{\frac{N}{2}}(J(\sigma^1) + \cdots + J(\sigma^s))) = \exp(\frac{1}{2}\beta^2 \frac{N}{2} \mathbb{E}(J(\sigma^1) + \cdots + J(\sigma^s))^2). \tag{59}$$

It turns out that

$$\mathbb{E}(J(\sigma^1) + \cdots + J(\sigma^s))^2 = s + 2 \sum_{(ab)} \delta_{ab}, \tag{60}$$

where the first term s comes from the diagonal terms in the square, while $\delta_{ab} = 1$ if the configurations σ^a and σ^b are equal, and zero otherwise, $1 \le a < b \le s$.

Therefore we have

$$\phi_N(s, \beta) = \frac{1}{Ns} \log \sum_{\sigma_i^a} \exp(\frac{1}{4} Ns\beta^2 + \frac{1}{2} N\beta^2 \sum_{(ab)} \delta_{ab}). \tag{61}$$

Now we split all possible configurations for the σ variables into the sum of K bubbles, $K = 1, 2, \ldots, s$, each made of s_r replicas, with $r = 1, \ldots, K$, $s_r \ge 1$, $\sum_r s_r = s$, in such a way that the σ's are all equal in each bubble, and all different for different bubbles.

The order parameters are therefore $K, s_1, \ldots, s_K$. For each of these specifications the $\sum_{(ab)}$ in the exponent reduces to

$$\sum_{(ab)} \delta_{ab} = \sum_r \frac{1}{2} s_r(s_r - 1) = \sum_r \frac{1}{2} s_r^2 - \frac{1}{2} s. \tag{62}$$

For each system of K bubbles, the sum $\sum_\sigma$ gives $2^N(2^N - 1) \ldots (2^N - K + 1)$. Only the dominating terms in N are relevant, therefore we can take into account only the dominating 2^{NK}.

By collecting all terms, we can write

$$\phi_N(s,\beta) = \frac{1}{Ns}\log \sum_{K,s_1,\dots,s_K} c(N;K,s_1,\dots,s_K)\exp(Ns\tilde{\phi}(s,\beta;K,s_1,\dots,s_K)), \tag{63}$$

where $\tilde{\phi}$ is the trial function

$$\tilde{\phi}(s,\beta;K,s_1,\dots,s_K) = \frac{\beta^2}{4s}\sum_r s_r^2 + \frac{K}{s}\log 2, \tag{64}$$

and the $c(N;K,s_1,\dots,s_K)$ take into account the neglected nondominating terms, as explained above, and the combinatorial factors connected to all possible ways to distribute the replicas into the available bubbles. These combinatorial factors do not contribute to the dominant terms in N, and are irrelevant.

Clearly the infinite volume limit ($N \to \infty$) of $\phi_N(s,\beta)$ will be bigger than each contributing term, and equal to the highest one.

By collecting all information, we have the order parameters $K, s_1, \dots, s_K$, and the trial functional $\tilde{\phi}(s,\beta;K,s_1,\dots,s_K)$ so that

$$\phi(s,\beta) = \sup_{K,s_1,\dots,s_K} \tilde{\phi}(s,\beta;K,s_1,\dots,s_K) = \sup_{K,s_1,\dots,s_K} (\frac{\beta^2}{4s}\sum_r s_r^2 + \frac{K}{s}\log 2). \tag{65}$$

In the trial functional the first term has the meaning of an energy, the second is the entropy. The variational principle is an entropy principle. The entropy is maximum with the constraint of a given energy.

The sup is easily found. In fact we have to consider that the map $s_r \to s_r^2$ is convex. From this and that $1 \le s_r \le s$ we have

$$s_r^2 \le -s + (s-1)s_r, \tag{66}$$

which implies the upper bound on the trial function

$$\frac{\beta^2}{4s}\sum_r s_r^2 + \frac{K}{s}\log 2 \le \frac{\beta^2}{4}(s+1) + K(\frac{\log 2}{s} - \frac{\beta^2}{4}). \tag{67}$$

This bound is linear in K and sharp at the extremal values $K = 1$ and $K = s$. Therefore the sup is reached in one of these extremal values.

At $K = 1$ (one big bubble), we have the only term $s_1 = s$, and the value of the trial function is $\frac{\beta^2}{4}s + \frac{1}{s}\log 2$. At $K = s$ (many small bubbles made by only one replica each), we have $s_r = 1$, $r = 1,\dots,s$, and the value of the trial function is $\frac{\beta^2}{4} + \log 2$.

We see that there are transition points $\beta_c^2(s) = 4\log 2/s$, such that

$$\phi(s,\beta) = \log 2 + \frac{\beta^2}{4}, \tag{68}$$

for $\beta \leq \beta_c(s)$, and

$$\phi(s,\beta) = \frac{1}{s}\log 2 + \frac{\beta^2}{4}s, \tag{69}$$

for $\beta \geq \beta_c(s)$.

The replica symmetry is never broken. The overlaps are all zero in the first region and all one in the second region. Remember that we are considering for the moment only integer values $s = 1, 2, \ldots$.

Now we come to the main point. We show that the variational principle at integer s, gives a strong hint toward what should be the variational principle for all $s > 0$. This is one interesting consequence of the "trick" on interpolating replicas.

Consider the values taken by the trial function

$$\tilde{\phi}(s,\beta;K,s_1,\ldots,s_K) = \frac{\beta^2}{4s}\sum_r s_r^2 + \frac{K}{s}\log 2, \tag{70}$$

for various specifications of the order parameters $K, s_1, \ldots, s_K$.

Let us start from the simple inequality

$$\frac{1}{K}\sum_r s_r^2 \geq (\frac{1}{K}\sum_r s_r)^2 = (\frac{s}{K})^2, \tag{71}$$

so that

$$\frac{1}{s}\sum_r s_r^2 \geq \frac{s}{K}. \tag{72}$$

Therefore, by defining $s/K = m$, so that $1 \leq m \leq s$, we have the estimate

$$\tilde{\phi}(s,\beta;K,s_1,\ldots,s_K) \geq \frac{1}{m}\log 2 + \frac{\beta^2}{4}m. \tag{73}$$

This expression is really remarkable. It suggests to consider the convex trial function for the order parameter m

$$0 < m \to \tilde{\phi}(m,\beta) = \frac{1}{m}\log 2 + \frac{\beta^2}{4}m, \tag{74}$$

independent of s, and such that, at least for integer values of s

$$\phi(s,\beta) = \sup_{1\leq m\leq s}(\frac{1}{m}\log 2 + \frac{\beta^2}{4}m). \tag{75}$$

The trial function is independent of s, only the range of the variational parameter m is taken to depend on s. It is impressive to see that the variational values of $\tilde{\phi}(s,\beta;K,s_1,\ldots,s_K)$ for different integer values of s, but at the same β, do still

have as a common lower bound the same reduced trial function $\tilde{\phi}(m, \beta)$, only the range for m may change, according to s.

Therefore, the order parameter m and the reduced trial function $\tilde{\phi}(m, \beta)$ are suggested, in the present interpretation of the "trick", by the variational values of the complete trial function $\tilde{\phi}(s, \beta; K, s_1, \ldots, s_K)$. The variational parameters $(K, s_1, \ldots, s_K)$, which strongly depend on the value of s, are collapsed to a unique essential variational parameter m.

The suggestion of the "trick" is proficuous. In fact, through some additional work, one can easily prove the following.

Theorem 1 *For the order parameter $m > 0$, introduce the convex trial function*

$$\tilde{\phi}(m, \beta) = \frac{1}{m} \log 2 + \frac{\beta^2}{4} m. \tag{76}$$

Then, we have in the infinite volume limit ($N \to \infty$) for the auxiliary function $\phi(s, \beta)$ for any $s \geq 0$ the following variational principle:

$$\phi(s, \beta) = \sup_{1 \leq m \leq s} \tilde{\phi}(m, \beta), \tag{77}$$

for $s \geq 1$, and

$$\phi(s, \beta) = \inf_{s \leq m \leq 1} \tilde{\phi}(m, \beta), \tag{78}$$

for $0 \leq s \leq 1$.

The inversion from a sup to an inf, by crossing the $s = 1$ line is completely analogous to the mentioned inversion from subadditivity to superadditivity.

Since $\tilde{\phi}(m, \beta)$ is convex, the sup for $s \geq 1$ can be reached only at the boundaries $m = 1$ or $m = s$, and the replica symmetry can not be broken. On the other hand, when $s < 1$, it can happen that the minimum for $\tilde{\phi}(m, \beta)$ is in the interval $s \leq m \leq 1$, and replica symmetry is broken.

Globally the space (s, β), for $s \geq 0$ is split in three regions. For $s \geq 1$ we are always in the replica symmetric case. For $\beta < \beta_c(s)$, with $\beta_c^2(s) = 4 \log 2/s$, we have

$$\phi(s, \beta) = \log 2 + \frac{\beta^2}{4}. \tag{79}$$

From

$$\frac{\partial}{\partial \beta} \phi_N(s, \beta) = \frac{\beta}{2} (1 + (s - 1) < \delta_{\sigma\sigma'} >), \tag{80}$$

we see that here $< \delta_{\sigma\sigma'} >= 0$ in the limit. For $\beta > \beta_c(s)$ we have

$$\phi(s, \beta) = \frac{1}{s} \log 2 + \frac{\beta^2}{4} s, \tag{81}$$

and here $< \delta_{\sigma\sigma'} >= 1$ in the limit.

These results extend the given expression from the case where s is an integer, to any $s \geq 0$. Notice that the line at $\beta_c(s)$ is a first order transition line. The function $\phi(s, \beta)$ is continuous, as it should be, because of the convexity in β. But its derivative in β has a sudden jump.

For $s < 1$ the situation is more complicated. There are two second order transition lines, the first at $\beta_c = 2\sqrt{\log 2}$, the second at $\beta_c'(s) = 2\sqrt{\log 2}/s$. The two merge at $s = 1$, but in general $\beta_c < \beta_c'(s)$ holds.

For $\beta \leq \beta_c$, replica symmetry holds,

$$\phi(s, \beta) = \log 2 + \frac{\beta^2}{4}, \tag{82}$$

and $< \delta_{\sigma\sigma'} >= 0$.

For $\beta \geq \beta_c'(s)$, replica symmetry is restored in the form

$$\phi(s, \beta) = \frac{1}{s} \log 2 + \frac{\beta^2}{4} s, \tag{83}$$

but now $< \delta_{\sigma\sigma'} >= 1$.

In the region $\beta_c \leq \beta \leq \beta_c'(s)$ replica symmetry is broken, and we have

$$\phi(s, \beta) = \beta\sqrt{\log 2}, \tag{84}$$

independently of s.

Now the formula

$$\frac{\partial}{\partial\beta}\phi_N(s, \beta) = \frac{\beta}{2}(1 + (s - 1) < \delta_{\sigma\sigma'} >), \tag{85}$$

gives

$$< \delta_{\sigma\sigma'} >= \frac{1}{1 - s}(1 - \frac{2\sqrt{\log 2}}{\beta}), \tag{86}$$

with a smooth interpolation between the value $< \delta_{\sigma\sigma'} >= 0$ at $\beta = \beta_c$, and $< \delta_{\sigma\sigma'} >= 1$ at $\beta = \beta_c'(s)$.

We can see that replica symmetry breaking is not connected to a difficulty in the analytic continuation of the replica symmetric solution.

In fact let us take $\beta > \beta_c$ and a large value of s, where

$$\phi(s, \beta) = \frac{1}{s} \log 2 + \frac{\beta^2}{4} s. \tag{87}$$

At fixed β, there is no problem in the analytic continuation of this expression to all values of $s > 0$ well inside the region of symmetry breaking, for $s < s_c =$

$2\sqrt{\log 2}/\beta$. However, for $s < s_c$, the equation (87) can no longer be true for a very simple reason.

In fact, at fixed β, the function

$$\frac{1}{s}\log 2 + \frac{\beta^2}{4}s \tag{88}$$

is decreasing, with decreasing s, up to the point $s = s_c$, where there is an inversion, and the function starts to increase with decreasing s.

Notice that the derivative

$$\frac{\partial}{\partial s}(\frac{1}{s}\log 2 + \frac{\beta^2}{4}s) = -\frac{1}{s^2}\log 2 + \frac{\beta^2}{4} \tag{89}$$

is positive in s for $s > s_c$, becomes zero at $s = s_c$, and becomes negative in s for $s < s_c$.

Since $\phi(s, \beta)$ must be increasing in s, we surely have

$$\phi(s, \beta) \leq (\frac{1}{s_c}\log 2 + \frac{\beta^2}{4}s_c). \tag{90}$$

As a matter of fact equality holds here, since for $s = s_c = 2\sqrt{\log 2}/\beta$ we have exactly

$$\frac{1}{s_c}\log 2 + \frac{\beta^2}{4}s_c = \beta\sqrt{\log 2}. \tag{91}$$

We say that in this case that "replica symmetry is minimally broken". Replica symmetry holds everywhere, with the exception of the region where this can not be true, by simple thermodynamic reasons. Then, necessarily

$$\phi(s, \beta) = \phi(s_c(\beta), \beta). \tag{92}$$

Notice that $\phi(s, \beta)$, at each fixed value of β, is convex both in s and $1/s$, as it should be.

This ends our discussion of the "replica trick" for the Random Energy Model. We have seen how to reach the right order parameter, and the right trial function, through a direct inspection of the variational behavior of the trial function for the system at integer values of s.

5 The Emergence of the Order Parameter and the Trial Function in the Sherrington-Kirkpatrick Model

Now we turn to the main topic of this paper. We will try to show that an appropriate form of the variational principle for integer values of s does immediately suggest the right form of the variational principle for all value of s, in particular at $s = 0$. As in the case of the Random Energy Model, the main ingredients of the emerging structure are provided by two facts. The new order parameter, call it symbolically x, and the trial function $\tilde{\phi}(\beta, x)$ must be independent of s. Then the order parameter x must belong to a convex space, and the trial function $\tilde{\phi}$, as a function of x, must be convex in x. The variational principle for $s \geq 1$ should be a $\sup_x$ principle, with some constraints on x which here we denote symbolically by $1 \leq x \leq s$, while for $0 < s \leq 1$ it should be a $\min_x$ principle with constraints on s of the type $s \leq x \leq 1$.

Obviously the two forms of variational principles holding for integer s in the Sherrington-Patrick model, which we have found in (26) and (32), do not satisfy these properties. As a matter of fact, in the general form (26), the qualitative aspects of the order parameters q_{ab} do strongly depend on s, while no kind of convexity is visible. No convexity is also visible in the simplified form (32). As a matter of fact, it is possible to easily verify that the trial function $\tilde{\phi}(s, \beta; q_{ab})$ in (26) is *concave* separately in each of the q_{ab}^2. In a sense, we must transform variational principles where concave trial functions are involved, into equivalent variational principles where *convex* trial functions appear, through an appropriate transformation of the order parameters.

The way out is very simple, but conceptually very deep. We give a synthetic sketch of the general strategy.

Firstly we have to resort to an old result coming from a simple form of interpolation, which was presented for the first time in [15]. Let us recall that the infinite volume limit $\lim_{N\to\infty} \phi_N(s, \beta) = \phi(s, \beta)$, holding in the form of an inf or a sup, as a consequence of subadditivity or superaddivity in N, for $s \geq 1$, or $s \leq 1$ respectively, can be recast, on a general basis, as a Cesàro limit [18] for marginal quantities in the volume N, according to $\phi(s, \beta) = c \lim_{N\to\infty}((N + 1)\phi_{N+1} - N\phi_N)$. For a generic sequence $N \to a_N$, $N = 1, 2, \ldots$, the Cesàro limit is defined as $c \lim_{N\to\infty} a_N = \lim_N N^{-1} \sum_{K=1}^{N} a_K$. Partial averages are involved, so that the Cesàro limit is equal to the standard limit if this does exist, but it can exist in some cases where the standard limit does not exist. A typical case is the alternating sequence $0, 1, 0, 1, \ldots$ which has $1/2$ as Cesàro limit.

Now we follow the strategy outlined in [15] for the case $s = 0$, but easily extended to any s. Starting from the representation given by (2), we can write

$$(N + 1)\phi_{N+1} - N\phi_N = s^{-1} \log \frac{\mathbb{E}Z_{N+1}^s}{\mathbb{E}Z_N^s}$$
$$= \log 2 + s^{-1} \log \mathbb{E}\Omega^s (\exp(\log \cosh(\beta\eta(\sigma))) - B, \quad (93)$$

where we have explicitely performed the sum over σ_{N+1}, which gives rise to the log 2 and the cosh put at the exponent under log for convenience, Ω is a random state depending on the $\sigma_1, \sigma_2, \ldots, \sigma_N$, the "cavity variables", $\eta(\sigma)$ denotes the interaction of the spin associated to the site $N+1$ with the cavity σ variables, given essentially by

$$\eta(\sigma) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} J_{iN+1}\sigma_i. \tag{94}$$

The quantity B in (93) is an additional important term which is not necessary to specify here. Its form, in terms of the relevant order parameter will be clear later.

The interpolating method proposed here is very simple [15]. Introduce an interpolating parameter $0 \leq q \leq 1$, and an auxiliary function $f(q, y)$, where y is a real variable, with given final boundary value $f(1, y) = \log\cosh(\beta y)$. We will try to adjust the function f in such a way that the quantity

$$s^{-1} \log \mathbb{E}\Omega^s(\exp f(q, \sqrt{q}\eta(\sigma)) \tag{95}$$

does not depend on q. If this is the case then

$$s^{-1} \log \mathbb{E}\Omega^s(\exp \log\cosh(\beta\eta(\sigma))) = f(0, 0). \tag{96}$$

By taking the explicit derivative with respect to q, and integrating by parts on the J_{iN+1} Gaussian random variables (see [15]), with some surprise we immediately have that the quantity in (95) really does not depend on q *provided* that the auxiliary function f satisfies the Hamilton-Jacobi-Bellmann equation

$$(\partial_q f)(q, y) + \frac{1}{2}\big(f''(q, y) + x(q)f'^2(q, y)\big) = 0, \tag{97}$$

with final condition

$$f(1, y) = \log\cosh(\beta y). \tag{98}$$

In (5.4), $f' = \partial_y f$ and $f'' = \partial_y^2 f$.

In the Eq. (97) x is a functional order parameter

$$[0, 1] \ni q \to x(q), \tag{99}$$

satisfying the bound $1 \leq x(q) \leq s$, for the case $s \geq 1$, and $s \leq x(q) \leq 1$ for the case $s \leq 1$.

By following the methods in [15] it is possible to show that there are order parameters x such that the representation given in (96) holds, even in cases where (95) does not hold, because of a compensation of terms for different values of q.

Notice that the Eq. (97) is nothing but Parisi equation, and the functional order parameter x is the Parisi functional order parameter [6], extended to any value of

s. Here, they are given for free through interpolation starting from the marginal $(N+1)\phi_{N+1} - N\phi_N$.

Equation (97) with the final condition (98) is easily solved by Gaussian quadratures and limits. In fact, let us start from a piecewise constant order parameter, expressed through constants $q_0 = 1 \leq q_1 \leq \cdots \leq q_K = 1$, and $1 \leq m_a \leq s$, $a = 1, \ldots, K$, for $s \geq 1$, or $s \leq m_a \leq 1$, $a = 1, \ldots, K$, for $s \leq 1$, in the form $x(q) = m_a$ for $q_{a-1} < q \leq q_a$. Then it is immediately seen that the solution of the Eq. (97), with final condition (98), is given recursively, starting from $a = K$ by the equation

$$f(q, y) = \frac{1}{m_a} \int \exp(m_a f(q_a, y + \sqrt{q_a - q})), \tag{100}$$

in the interval $q_{a-1} \leq q \leq q_a$.

A very appealing interpretation of this recursive structure is given in [16], in the frame of the concept of a *multibath* having different temperatures and widely separated timescales.

In any case, it can be immediately shown (see [15]) that f is pointwise continuous in the $L^1(dq)$ norm. In fact, for generic piecewise constant order parameters x, $\bar{x}$, we have

$$|f(q, y; x) - f(q, y; \bar{x})| \leq \frac{\beta^2}{2} \int_q^1 |x(q') - \bar{x}(q')|\, dq', \tag{101}$$

where we have written explicitly the dependence of $f(q, y)$ on x and $\bar{x}$. The continuity in the L^1 norm allows to express pointwise $f(q, y; x)$ in terms of any L^1 order parameter through a limiting procedure.

We have also that the function f is monotone in x, in the sense that $x(q) \leq \bar{x}(q)$, for all $0 \leq q \leq 1$, implies $f(q, y; x) \leq f(q, y; \bar{x})$, for any q and y.

Recently, through ground breaking work of Antonio Auffinger and Wei-Kuo Chen [17], it has been proven that $f(0, 0; x)$ is strictly convex in x, in the sense that if $x(q) = \alpha x_1(q) + (1 - \alpha)x_2(q)$, with $0 \leq \alpha \leq 1$, then

$$f(0, 0; x) \leq \alpha f(0, 0; x_1) + (1 - \alpha) f(0, 0; x_2), \tag{102}$$

where the equality holds only in the degenerate cases, $x_1(q) = x_2(q)$, or $\alpha = 0$, or $\alpha = 1$. The methods exploited in [17] are partly based on [12].

Now we can invoke the powerful broken replica bounds introduced in [11], extended to any value of s, in order to produce the right candidate for the variational trial functional with the correct properties. Let us notice that from (93), (96) we know that there exists an order parameter x such that

$$\phi_N(s, \beta) = \log 2 + f(0, 0; x) - B. \tag{103}$$

Of course, we do not know the value of x, nor the expression of B. However we can try to compare $\phi_N(s, \beta)$ with $f(0, 0; x)$ through an interpolation argument similar to what was done in [11] in the case $s = 0$. We end up with the following

sum rule holding for any x

$$\phi_N(s,\beta) = \log 2 + f(0,0;x) - \frac{\beta^2}{2}\int q\, x(q)\, dq + R, \tag{104}$$

where the error term R shares the following remarkable properties, uniformly in N: in the case $s \geq 1$ it happens that $R \geq 0$ *provided* $x(q)$ is non-increasing in q and bounded by $s \geq x(q) \geq 1$, we call this condition $\mathcal{X}_-(s)$, while in the case $0 \leq s \leq 1$ the error term is $R \leq 0$ *provided* $x(q)$ is non-decreasing in q and bounded by $s \leq x(q) \leq 1$, we call this condition $\mathcal{X}_+(s)$.

Therefore, the whole procedure induces to define

$$\tilde{\phi}(\beta;x) = \log 2 + f(0,0;x) - \frac{\beta^2}{2}\int q\, x(q)\, dq \tag{105}$$

as the trial functional, independent of s and strictly convex in x, as requested.

The sum rule implies

$$\phi_N(s,\beta) \geq \phi(s,\beta) \geq \tilde{\phi}(\beta;x), \tag{106}$$

in the case $s \geq 1$ with $x \in \mathcal{X}_-(s)$, while

$$\phi_N(s,\beta) \leq \phi(s,\beta) \leq \tilde{\phi}(\beta;x), \tag{107}$$

in the case $0 \leq s \leq 1$ with $x \in \mathcal{X}_+(s)$.

Notice that both $\mathcal{X}_-(s)$ and $\mathcal{X}_+(s)$ are convex spaces of functions.

For a moment let us go back to the case of where s is an integer. Through a direct calculation, we immediately see that the replica symmetric trial in (28) can be expressed also as

$$\tilde{\phi}_{RS}(s,\beta;\bar{q}) = \tilde{\phi}(\beta;x_{\bar{q}}), \tag{108}$$

where $x_{\bar{q}}$ is the order parameter in $\mathcal{X}_-(s)$ defined by $x_{\bar{q}}(q) = s$ for $0 \leq q \leq \bar{q}$, and $x_{\bar{q}}(q) = 1$ for $\bar{q} \leq q \leq 1$. Order parameters of the type $x_{\bar{q}}$ are extremals in the convex space $\mathcal{X}_-(s)$.

Then we can check immediately that

$$\sup_{x\in\mathcal{X}_-(s)} \tilde{\phi}(\beta;x) = \sup_{\bar{q}} \tilde{\phi}_{RS}(s,\beta;\bar{q}) = \phi_{RS}(s,\beta) = \phi(s,\beta). \tag{109}$$

The proof is very simple but instructive. Let us consider any function x that is piecewise constant in $\mathcal{X}_-(s)$. It is easy to see that x is given by a convex mixture of averages on the extremals

$$x(q) = \sum_{a=1} p_a x_{q_a}(q), \tag{110}$$

where the weights $p_a \geq 0$ are expressed through the decreasing sequence of the m_a defined by

$$m_1 = s = (p_1 + p_2 + \cdots + p_K)s,\ m_2 = p_1 + (p_2 + \cdots + p_K)s, \ldots \tag{111}$$

Due to convexity we have

$$\tilde{\phi}(\beta; x) \leq \sum_{a=1}^{K} p_a \tilde{\phi}_{RS}(s, \beta; q_a) \leq \phi_{RS}(s, \beta) = \phi(s, \beta). \tag{112}$$

On the other hand (with $\tilde{\phi}_{RS}(s, \beta; \bar{q})$ defined correspondingly to (27):

$$\sup_{x \in \mathcal{X}_-(s)} \tilde{\phi}(\beta; x) \geq \sup_{\bar{q}} \tilde{\phi}_{RS}(s, \beta; \bar{q}), \tag{113}$$

because a constraint is added. By putting together the two inequalities, then Eq. (109) follows.

As a matter of fact this result holds in general, so that we have the Theorem.

Theorem 2 *Define the s independent trial functional* $\tilde{\phi}(\beta; x) = \log 2 + f(0, 0; x) - \frac{\beta^2}{2} \int q\, x(q)$*, strictly convex in the functional order parameter then the infinite volume limits* $\phi(s, \beta)$ *are given by the constrained variational principles*

$$\phi(s, \beta) = \sup_{x \in \mathcal{X}_-(s)} \tilde{\phi}(\beta; x) \tag{114}$$

for $s \geq 1$*, and*

$$\phi(s, \beta) = \inf_{x \in \mathcal{X}_+(s)} \tilde{\phi}(\beta; x) \tag{115}$$

for $0 \leq s \leq 1$.

For the complete proof we have to rely on the extension of the basic results of Michel Talagrand [19] and Dmitry Panchenko [20].

6 Conclusion and Outlook

We have seen how to reach the right order parameter and the right trial functional for the generalized annealed free energy in the case of mean field disordered models, starting from the elementary case of annealed replicated models. We have treated in detail the case of the Sherrington-Kirkpatrick mean field spin glass model, and the case of the Derrida random energy model. Since annealed replicated models can be easily treated, our methods extend also to other interesting models as the multispecies models, the neural networks, the K-SAT problems and similar models, [14, 21]. We plan to report on them on a future occasion.

References

1. Guerra, F.: The replica trick in the frame of replica interpolation. In: Gayrard, V., et al. (eds.) Statistical Mechanics of Classical and Disordered Systems. Springer Nature Switzerland AG (2019)
2. Sherrington, D., Kirkpatrick, S.: Solvable model of a spin-glass. Phys. Rev. Lett. **35**, 1792–1796 (1975)
3. Derrida, B.: Random-energy model: an exactly solvable model of disordered systems. Phys. Rev. B **24**, 2613 (1981)
4. Guerra, F.: An introduction to mean field spin glass theory: methods and results. In: Bovier, A., et al. (eds.) Mathematical Statistical Physics, pp. 243–271. Elsevier, Oxford, Amsterdam (2006)
5. de Almeida, J.R.L., Thouless, D.T.: Stability of the Sherrington-Kirkpatrick solution of a spin glass model. J. Phys. A: Math. Gen. **11**, 983 (1978)
6. Parisi, G.: The order parameter for spin glasses: a function on the interval 0-1. J. Phys. **A13**, L-1101 (1980)
7. Mézard, M., Parisi, G., Virasoro, M.A.: Spin Glass Theory and Beyond. World Scientific, Singapore (1987)
8. Talagrand, M.: Large deviations, Guerra's and A.S.S. schemes, and the Parisi hypothesis. J. Stat. Phys. **126**, 837–894 (2007)
9. Barra, A., Guerra, F., Mingione, E.: Interpolating the Sherrington-Kirkpatrick replica trick. Phil. Mag. **92**, 78–97 (2012)
10. Guerra, F., Toninelli, F.: The thermodynamic limit in mean field spin glass models. Commun. Math. Phys. **230**, 71–79 (2002)
11. Guerra, F.: Broken replica symmetry bounds in the mean field spin glass model. Commun. Math. Phys. **233**, 1–12 (2003)
12. Boué, M., Dupuis, P.: A variational representation for certain functionals of Brownian motion. Ann. Probab. **26**, 1641–1659 (1998)
13. Guerra, F., Ghirlanda, S.: General properties of overlap probability distributions in disordered spin systems. Towards Parisi Ultrametricity. J. Phys. A: Math. Gen. **31**, 9149–9155 (1998)
14. Talagrand, M.: Spin Glasses: A Challenge for Mathematicians. Springer-Verlag, Berlin (2003)
15. Guerra, F.: Fluctuations and thermodynamic variables in mean field spin glass models. In: Albeverio, S., Cattaneo, U., Merlini, D. (eds.) Stochastic Processes, Physics and Geometry, II. World Scientific, Singapore (1995)
16. Contucci, P., Corberi, F., Kurchan, J., Mingione, E.: Stationarization and multithermalization in spin glasses. arXiv:2012.03922v1 [cond-mat.dis-nn]
17. Auffinger, A., Chen, W.-K.: The Parisi formula has a unique minimizer. Commun. Math. Phys. **335**, 1429–1444 (2015)
18. Cesàro, E.: Sur la convergence des séries. Nouvelles annales de mathématiques 3^e série, tome 7, pp. 49–59 (1888)
19. Talagrand, M.: The Parisi formula. Ann. Math. **163**, 221–263 (2006)
20. Panchenko, D.: The Sherrington-Kirkpatrick Model. Springer-Verlag, Berlin (2013)
21. Mézard, M., Zecchina, R.: Random K-satisfiability problem: from an analytic solution to an efficient algorithm. Phys. Rev. E **66**, 056126 (2002)

Information and Complexity, Or: Where Is the Information?

Nihat Ay, Nils Bertschinger, Jürgen Jost, Eckehard Olbrich, and Johannes Rauh

Abstract We demonstrate how a systematic theory of complexity emerges from information theoretical concepts. The complexity of a structure may refer to the difficulty of its description, the encoding of its regularities or the relations between its elements, components or parts. All such measures can be and usually are quantified with the help of information theoretical concepts. We first describe those concepts and then use them to analyze how complexity emerges from interactions between parts or conversely, can be decomposed into individual, joint, or complementary contributions of those parts. We can also use these theoretical concepts to understand the interaction between a system and its environment and the relations between different levels of aggregation in complex systems.

Keywords Information theory · Mutual information · Complexity measures · Information decomposition · Levels and scales

1 Introduction

The concepts of information and complexity seem to be intricately linked. Complexity notions are quantified in information theoretical terms, and a general principle might say that a structure is the more complex, the more information is needed to

N. Ay · J. Jost (✉) · E. Olbrich · J. Rauh
Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany
e-mail: jost@mis.mpg.de

N. Ay · J. Jost
Santa Fe Institute for the Sciences of Complexity, Santa Fe, NM, USA

N. Bertschinger
Frankfurt Institute for Advanced Studies, Frankfurt am Main, Germany

Goethe University, Frankfurt am Main, Germany

S. Albeverio et al. (eds.), *Complexity and Emergence*, Springer Proceedings in Mathematics & Statistics 383, https://doi.org/10.1007/978-3-030-95703-2_4

describe or build it. That principle, however, needs some qualification. One should distinguish between—usually useful—information about regularities of a structure or a process and—often useless—information about random details. The question is not only *information about what?*, but also *where is that information?*, that is, whether and how it is or can be internally stored in a system with limited capacity, at which level of a process information is needed to predict the continuation of a process, and where it can be found in a distributed system. In the latter case, we should, however, not only look for information that is exclusively located somewhere or that is shared between entities, but should also consider complementary or synergistic information, that is, information that only emerges when several sources are combined.

These lecture notes describe what is currently known about these questions, and they develop the underlying theoretical concepts and elucidate them at simple examples. Also, when we can quantify complexity concepts, we can also try to optimize the corresponding complexity measures. This will also be systematically discussed.

These notes are the result of a series of lectures that one of us (JJ) delivered at the Summer School in Como in July, 2018. They present work that we have done jointly during the last few years. JJ thanks Elisa Mastrogiacomo and Sergio Albeverio for organizing a very stimulating school, and the participants and the other lecturers, in particular Luciano Boi, Ivar Ekeland and Frank Riedel, for stimulating discussions.

2 Background: Principles of Information Theory

2.1 Shannon Information

The basic concept is that of the *Shannon Information* [52] of a random variable X, or equivalently, of a probability distribution p, when the possible values x_i of X are realized with probabilities $p_i = p(x_i)$. These probabilities satisfy $0 \leq p_i \leq 1$ for all i, with the normalization $\sum_i p_i = 1$. The Shannon information or entropy then is

$$H(X) = H(p_1, \ldots, p_n) = -\sum_i p_i \log_2 p_i \text{ (bits)}. \tag{1}$$

This is the expected *reduction of uncertainty*, i.e., the *information gain*, if we learn which concrete value x_i of the random variable X from a known distribution p with probabilities $p_i = p(x_i)$ is realized.

This is the basic example: When we have two possible events occurring with equal probability 1/2 (an unbiased coin) we thus gain $\log_2 2 = 1$ bit of information when we observe the outcome.

A fair dice yields $\log_2 6$ bits of information.

2.2 Mutual and Conditional Information

We now consider the situation where we have an additional random variable Y.

In the example of the dice, we could let $Y = 0$ (resp. 1) for an odd (even) result, each with probability 1/2. According to the basic example, we have $H(Y) = 1$ bit. When we know Y, there remain only 3 possibilities for the value of X, each with probability 1/3.

This leads us to the concept of *conditional information*; in this example, the remaining uncertainty about X when knowing Y is

$$H(X|Y) = \log_2 3. \tag{2}$$

Thus, the uncertainty about the value of X is reduced from $\log_2 6$ to $\log_2 3$ bit when knowing Y.

The joint information is related to the conditional information by

$$H(X, Y) = H(X) + H(Y|X) = H(Y) + H(X|Y). \tag{3}$$

Thus, $H(X, Y) \leq H(X) + H(Y)$, and $<$ if X and Y are not independent. In the example, we have $H(X, Y) = H(X)$, since the value of X determines that of Y.

The information gain about X from knowing Y is called the *mutual information* of X and Y,

$$MI(X : Y) = H(X) - H(X|Y). \tag{4}$$

In our example $MI(X : Y) = \log_2 6 - \log_2 3 = \log_2 2 = 1$ bit. From Y, we gain 1 bit of information about X.

The mutual information is symmetric,

$$MI(X : Y) = MI(Y : X). \tag{5}$$

The difference structure is perhaps the most important aspect. In many respects, (4) is more important and fundamental than (1), because we always have some prior knowledge, expressed here through Y, when we observe some X. Thus, the mutual information $MI(X : Y)$ tells us how much we can already infer about X when we know Y. By then observing X, we only gain the additional information $H(X|Y)$.

Summary:

$$H(X) = MI(X : Y) + H(X|Y) \tag{6}$$

$H(X)$ = how much you learn from observing X
$MI(X : Y)$ = how much you learn about X by observing Y
$H(X|Y)$ = how much you learn from observing X when you already know Y.

We can iterate the conditioning process with another random variable Z, to get the *conditional mutual information*

$$MI(X:Y|Z) = H(X|Z) - H(X|Y,Z). \tag{7}$$

$MI(X:Y|Z)$ quantifies how much additional mutual information between X and Y can be gained when we already know Z.

Careful: While always $H(X|Z) \leq H(X)$, we do *not necessarily* have $MI(X:Y|Z) \leq MI(X:Y)$.
Example: The XOR function (exclusive *or*):

x	y	z
0	0	0
1	0	1
0	1	1
1	1	0

where X, Y assume their two values independently with probability 1/2 each. Thus, $MI(X:Y) = MI(X:Z) = MI(Y:Z) = 0$, but $MI(X:Y|Z) = MI(X:Z|Y) = MI(Y:Z|X) = 1$, because knowing the values of two of the variables determines that of the third.

2.3 Maximum Entropy

E. Jaynes' maximum-entropy principle [24]: Take the least informative estimate possible on the given information, that is, don't put any information into your model that is not based on the observed data. Look for p with maximal entropy $H(p)$ under the constraint that the expectation values of certain observables f_α be reproduced,

$$E_p f_\alpha = \sum_i f_\alpha^i p_i \text{ for } \alpha = 1, \ldots, A. \tag{8}$$

The solution is an exponential distribution

$$p_j = \frac{1}{Z} \exp\Big(\sum_\alpha \lambda_\alpha f_\alpha^j\Big) \quad \text{with } Z = \sum_i \exp\Big(\sum_\alpha \lambda_\alpha f_\alpha^i\Big). \tag{9}$$

In particular, when there are no observations,

$$p_j = \frac{1}{n} \text{ for } j = 1, \ldots, n. \tag{10}$$

2.4 Kullback-Leibler Divergence

A reference for the information geometric concepts that will be introduced and used here and in the sequel is [5]. The *Kullback-Leibler divergence* (*KL-divergence* for short) or *relative entropy* for two probability distributions p, q

$$D(p\|q) = \begin{cases} \sum_i p_i \log_2 \frac{p_i}{q_i} & \text{if supp } p \subset \text{supp } q \\ \infty & \text{else} \end{cases} \tag{11}$$

is positive ($D(p\|q) > 0$ if $p \neq q$), but not symmetric, as in general, $D(p\|q) \neq D(q\|p)$.

Example: The mutual information is the KL-divergence between the joint distribution and the product of the marginals,

$$MI(X : Y) = D(p(x, y)||p(x)p(y)). \tag{12}$$

Among all distributions $p(x, y)$ with the same marginals $p(x) = \sum_y p(x, y)$, $p(y) = \sum_x p(x, y)$, the product distribution $p(x)p(y)$ has the largest entropy. This is, of course, a special case of Jaynes' principle. That is, when we only know the marginals, Jaynes' principle would suggest to take the product distribution as our estimate.

Example: The space of all probability distributions on two binary variables is a 3-dimensional simplex. It contains the 2-dimensional subfamily of product distributions. The extreme points of the simplex are the Dirac measures $\delta^{(x,y)}$, $x, y = 0, 1$. Maximization of the distance from the family of product distributions leads to distributions with support cardinality two (perfect correlation or anticorrelation) [4].

The formal way of expressing Jaynes' principle is to project a given distribution onto the product family $\mathcal{E}$ to maximize entropy while preserving the marginals, with π denoting that projection,

$$\begin{aligned} D(p \,\|\, \mathcal{E}) \;&:=\; \inf_{q \in \mathcal{E}} D(p \,\|\, q) \;=\; D(p \,\|\, \pi(p)) \\ &=\; H_{\pi(p)}(X, Y) - H_p(X, Y). \end{aligned} \tag{13}$$

3 Complexity

In this section, we want to introduce and discuss complexity concepts. But what is *complexity*? Some possible answers (see [5, 6] for a systematic discussion): **Complexity** is

1. the minimal effort or the minimal resources needed to describe or generate an object. Examples of such complexity concepts include algorithmic complexity

(Kolmogorov [31], Chaitin [14], Solomonoff [57]); computational complexities; entropy (Shannon [52]), or entropy rate (Kolmogorov [31], Sinai [56]).
2. the minimal effort or the minimal resources needed to describe or generate the *regularities* or the *structure* of an object. Examples of such complexity concepts include Kolmogorov minimal sufficient statistics and related notions, stochastic complexity (Rissanen [46]), effective complexity (Gell-Mann and Lloyd [18]), excess entropy [53], also known as effective measure complexity [21], forecasting complexity [64], also introduced as statistical complexity by Crutchfield, Young, Shalizi [15, 51].
3. the extent to which an object, as a whole, is more than the sum of its parts (Aristotle [1]), that is, the extent to which the whole cannot be understood by the analysis of the parts of the system in isolation, but only by also considering their interactions.

In order to systematically explore these aspects, we start with the most basic concept, that of algorithmic complexity [14, 31, 57] (see [33] for a systematic exposition). This concept expresses 1) in its purest form.

3.1 Algorithmic Complexity

The *algorithmic complexity* of an object, such as a number or a piece of text, is the length of the shortest computer program that generates or produces the object as output.[1] Typically, one cannot compute this complexity, but only provide an upper bound by producing a computer program, but does not know whether this is the shortest possible one.

From a conceptual perspective, the basic premise is that irregular or random structures have the highest algorithmic complexity, because they do not admit a short description. In other words, we want to characterize the complexity of a structure by the difficulty of its description. That is, we ask the question: How much can the description of a structure be simplified by utilizing regularities?

- Very simple structures need not be simplified any further.
- Random structures cannot be simplified.
- *Computational complexity* (see for instance the expositions in [38, 39]): Running time of shortest computer program that can generate the structure: A simple structure is produced quickly, whereas for a random one, everything has to be explicit in the program, and so, it does not need to run for a long time either.
- Random structures are not of interest for themselves, but only as members of an ensemble; it therefore suffices to describe the latter (Gell-Mann and Lloyd [18]).

[1] To make the complexity of different objects comparable, one needs to agree on a predetermined programming language; usually, one assumes some universal Turing machine, and changing that Turing machine will introduce an additive constant in the upper bounds.

3.2 External and Internal Complexity

The question that arises from the above concept of algorithmic complexity is how to compute it, that is, how to find the shortest description of a given structure. Quite apart from the fact that this depends on the choice of the device we use to evaluate it (in theory: some universal Turing machine, and the choice of that Turing machine then introduces an additive constant), in practice, we have only bounded means to represent a structure. Thus: What do we want to know? We want to

1. know a rich and complex structure,
2. but represent it most efficiently.

More formally, we want to

1. maximize *external complexity*,
2. but minimize *internal complexity*.

This perspective was introduced in [25]. For an application in pattern classification, see for instance [3].

3.3 Optimization Principles

Organisms live in and interact with a complex environment, see for instance [61] (for a measure theoretical approach, see [7]), and need to maintain their own autopoiesis [37]. A modern society consists of several complex subsystems that follow their own rules, but need to interact with each other [35, 36]. With the concept of Shannon information, we can formulate some abstract principles that either maximize or minimize some kind of complexity (we follow [26] here). The basic versions, however, lead to trivial results, as we shall now see.

1. Gain as much information as possible: Look at random patterns
2. Avoid surprises: Look at blank screen
3. Try to predict future sensory inputs as accurately as possible on the basis of the current ones (and perhaps try to bring yourself into a state where this is possible [17])
4. Try to manipulate the environment such that the results of own actions are as accurately predictable as possible [32].
5. Maximize

$$\begin{aligned} & MI(S_{t+1} : E_t) \qquad - MI(S_{t+1} : E_t | S_t) \\ = \; & H(S_{t+1}) - H(S_{t+1}|E_t) - H(S_{t+1}|S_t) + H(S_{t+1}|E_t, S_t) \end{aligned} \tag{14}$$

   to establish the strongest possible correlation between the current state E_t of the environment and future sensory data S_{t+1}, but such that this correlation can already be predicted from the current input S_t [9]

To proceed further, let us discuss some questions.

1. Q: Why should a system model an external probability distribution?
   A: To make predictions on the basis of regularities
2. Q: How can this be achieved in an environment that is vastly more complex than the system itself?
   A: Detect regularities
3. Q: How to detect regularities?
   A: Because of 2), the system is forced to compress.

These answers have some consequences in various fields:

- Psychology: Use heuristics [19, 54, 55]
- Cognition: External versus internal complexity [25]
- Statistics: Avoid overfitting
- Statistical learning theory: Start with models with few parameters and gradually increase as you learn (Vapnik-Chervonenkis) [59, 60].

3.4 Correlations in Time Series

We can also use the information theoretical notions to evaluate the complexity of a time series in terms of the correlations that it exhibits. A time series $X_t, t \in \mathbb{N}$ could possibly have

- No regularities: $H(X_t|X_{t-1}) = H(X_t)$
- the Markov property: $H(X_t|X_{t-1}, X_{t-2}, \dots) = H(X_t|X_{t-1})$, or
- Long term correlations, as in texts, genetic sequences, …

To evaluate this, we quantify how much new information is gained when one already knows n consecutive symbols and then sees the $(n+1)$st. (Grassberger [21]).

For which n is this largest? When n is small, one perhaps cannot predict much, and if n is large, one may be able to guess the rest anyway.

The larger this n, the more complex the sequence.

For genetic sequences, $n \sim 14$ [47], for amino acid sequences (proteins) $n \sim 5$.

In literature analysis, such a principle can be used to evaluate the complexity of language [16].

A more sophisticated concept is the genon concept of molecular biology [30, 48, 49].

3.5 Complementarity

Instead of trying to predict the environment, one can also let the environment do the computation itself (see [26]).

If you want to catch a ball, you do not use Newtonian mechanics to compute the trajectory, but simply run so that the ball appears under a constant angle. The environment computes the trajectory, and you only need to sample. This outsourcing of computation represents one mechanism for the compression mentioned in Sect. 3.3.

More generally, embodied cognition has emerged as a new paradigm in robotics [43].

3.6 Hierarchical Models and Complexity Measures

In this section, we follow [5, 6]. Returning to Jaynes' approach, we could maximize entropy while preserving marginals among subsets of variables. For instance, for a distribution on 3 variables, we could prescribe all single and pairwise marginals.

Assume that we have a state set V that consists of the possible values of N variables. We then consider the hierarchy

$$\mathfrak{S}_1 \subseteq \mathfrak{S}_2 \subseteq \ldots \subseteq \mathfrak{S}_{N-1} \subseteq \mathfrak{S}_N := 2^V, \tag{15}$$

where $\mathfrak{S}_k$ is the family of subsets of V with $\leq k$ elements, from wich we get the set of probability distributions $\mathcal{E}_{\mathfrak{S}_k}$ with dependencies of order $\leq k$. For instance, $\mathcal{E}_{\mathfrak{S}_1}$ is the family of distributions that are simply the products of their marginals. In particular, for a probability distribution in this family, there are no correlations between the probabilities of two or more of the variables. In $\mathcal{E}_{\mathfrak{S}_2}$, we then allow for pairwise correlations, but no triple or higher order ones.

We point out that one can also consider other families of subsets of V and the corresponding probability distributions. For instance, when V is the ordered set of integers $\{1, \ldots, N\}$, one could consider the family of those subsets that consist of uninterrupted strings of length $\leq k$. This will be our choice when we discuss the excess entropy below.

We let $\pi_{\mathfrak{S}_k}$ be the projection on $\mathcal{E}_{\mathfrak{S}_k}$, $p^{(k)} := \pi_{\mathfrak{S}_k}(p)$. For instance, $p^{(1)}$ is the product distribution with the same marginals as p.

We have the important Pythagorean relation

$$D(p^{(l)} \parallel p^{(m)}) = \sum_{k=m}^{l-1} D(p^{(k+1)} \parallel p^{(k)}), \tag{16}$$

for $l, m = 1, \ldots, N-1, m < l$. In particular,

$$D(p \parallel p^{(1)}) = \sum_{k=1}^{N-1} D(p^{(k+1)} \parallel p^{(k)}). \tag{17}$$

If we take configurations with dependencies of order $\leq k$, we get the **Complexity measure** [6] with weight vector $\alpha = (\alpha_1, \ldots, \alpha_{N-1}) \in \mathbb{R}^{N-1}$

$$C_\alpha(p) := \sum_{k=1}^{N-1} \alpha_k \, D(p \parallel p^{(k)}) \tag{18}$$

$$= \sum_{k=1}^{N-1} \beta_k \, D(p^{(k+1)} \parallel p^{(k)}), \tag{19}$$

with $\beta_k := \sum_{l=1}^{k} \alpha_l$.

$p^{(k)}$ is the distribution of highest entropy among all those with the same correlations of order $\leq k$ as p.

Thus, we consider a weighted sum of the higher order correlation structure.

Examples:

- Tononi-Sporns-Edelman complexity [58]: $\alpha_k = \frac{k}{N}$ addresses the issue of the interplay between differentiation and integration in complex systems (for an analysis of system differentiation from an information theoretical perspective, see also [28])
- Stationary stochastic process X_n: Conditional entropy

$$h_p(X_n) := H_p(X_n \mid X_1, \ldots, X_{n-1}).$$

Entropy rate or *Kolmogorov–Sinai entropy* [31, 56]

$$h_p(X) := \lim_{n\to\infty} h_p(X_n) = \lim_{n\to\infty} \frac{1}{n} H_p(X_1, \ldots, X_n), \tag{20}$$

Excess entropy (Grassberger [21])

$$E_p(X) := \lim_{n\to\infty} \sum_{k=1}^{n} (h_p(X_k) - h_p(X))$$

$$= \lim_{n\to\infty} \left(H_p(X_1, \ldots, X_n) - n h_p(X) \right) \tag{21}$$

$$= \lim_{n\to\infty} \underbrace{\sum_{k=1}^{n-1} \frac{k}{n-k} \, D(p_n{}^{(k+1)} \parallel p_n{}^{(k)})}_{=:E_p(X_n)}, \tag{22}$$

where we choose $\mathfrak{S}_k$ as the sequences of integers $j+1, j+2, \ldots, j+\ell$ with $\ell \leq k$. The excess entropy measures the non-extensive part of the entropy, i.e. the amount of entropy of each element that *exceeds* the entropy rate.

3.7 Interactions Between Levels

The question of emergence, that is, how a higher level that is (at least partially) autonomous from lower levels, arises in many disciplines. For example, classical mechanics arises from an underlying quantum structure, but the laws of classical mechanics are causally closed, in the sense that for computing trajectories of Newtonian particles, we do not need information from the quantum level. Likewise, human genetics rests on the laws of Mendel and does not need to consider an underlying biochemical level. In other fields it is often not so clear, however, to what extent laws operate autonomously at a certain level without needing permanent or at least regular access to some lower level. For instance, does it suffice for understanding macroeconomic processes to consider relations between macroeconomic variables, or is an input from the microeconomic level essentially needed? Or can one understand social dynamics without access to the psychic and mental states of the participating individuals? For a general discussion of the issue of emergence from the perspective developed in the present contribution, see for instance [29].

Here, we describe the approach of [41, 42] (and refer to [41] for references to earlier work). We consider a structure

$$\begin{array}{ccc} \widehat{X} & \xrightarrow{\psi} & \widehat{X}' \\ {\scriptstyle\pi}\uparrow & & \uparrow{\scriptstyle\pi} \\ X & \xrightarrow{\phi} & X' \end{array}$$

with basic level X, X' and higher level $\widehat{X}$, $\widehat{X}'$; an arrow $Y \to Y'$ represents a discrete time step where X, X' form a Markov process, with transition kernel ϕ, which can be observed at the higher level $\widehat{X}$, $\widehat{X}'$ in a lossy fashion.

The higher level could result from averaging or aggregating the lower level. Think of $\widehat{X}$ as a coarse-graining of X given by an observation map π.

We can propose several criteria for the upper process being closed in the sense that it depends on the lower process only through some initialization.

I **Informational closure**: The higher process is informationally closed, i.e. there is no information flow from the lower to the higher level. Knowledge of the microstate will not improve predictions of the macrostate.

$$MI(\widehat{X}' : X|\widehat{X}) = 0 \tag{23}$$

where the conditional mutual information

$$MI(\widehat{X}' : X|\widehat{X}) = H(\widehat{X}'|\widehat{X}) - H(\widehat{X}'|X) \tag{24}$$

measures the reduction in uncertainty about $\widehat{X}'$ when knowing X instead of only $\widehat{X}$.

II **Observational commutativity**: It makes no difference whether we perform the aggregation first and then observe the upper process, or we observe the process on the microstate level, and then lump together the states.
Kullback-Leibler divergence between the lower and the upper transition kernel from X to $\widehat{X}'$ is 0, for some initial distribution on X.

$$I \Rightarrow II, \text{ and in deterministic case also } II \Rightarrow I. \tag{25}$$

(In *I*, probabilities at $\widehat{X}$, in *II* at X)

III **Commutativity**: There exists a transition kernel ψ such that the diagram commutes (Görnerup-Jacobi, 2010)

$$II \Rightarrow III, \text{ and in deterministic case also } III \Rightarrow II. \tag{26}$$

II: Transition kernels satisfy $\Psi = \Pi\Phi\Pi^T$
III: Transition kernels satisfy $\Psi\Pi = \Pi\Phi$

IV **Markovianity**: $\widehat{X}$, $\widehat{X}'$ forms again a Markov process (Shalizi-Moore, 2003).

$$I \Rightarrow IV, \text{ but } IV \nRightarrow III. \tag{27}$$

V **Predictive efficiency**: A more abstract formulation is that an emergent level corresponds to an efficiently predictable process, that is, one that can be predicted in its own terms, without permanent recourse to a lower level.

3.7.1 A Test Case: The Tent Map

We now evaluate the preceding concepts at the example of the tent map, following [40] (see also [2] for background).

$$T(x) = \begin{cases} 2x & \text{if } 0 \le x \le 1/2 \\ 2 - 2x & \text{else} \end{cases}$$

The tent map is a basic example of a chaotic dynamical iteration, because at every step differences between values can get doubled, and therefore, after several steps, even very tiny differences between initial values can become macroscopically large. The folding at $x = 1/2$ ensures that nevertheless the unit interval is mapped to itself. Thus, some differences also get reduced. Understanding this interplay between amplification and reduction of differences is surprisingly subtle, as one may also see in the following.

For a threshold value $\alpha \in [0, 1]$ we define the symbolic dynamics

$$\phi_\alpha : \quad X \to \hat{X} = \{0, 1\}$$
$$\phi_\alpha(x) := \begin{cases} 0 & \text{if } 0 \le x < \alpha \\ 1 & \text{else} \end{cases}$$

The sequence $x_n = T^n(x)$, for an initial value $x \in X$, yields the derived symbol dynamics $s_n = \phi_\alpha(x_n) \in \{0, 1\}$.

The probability of finding s_n in the state 0 is the probability that x_n lies in the interval $[0, \alpha]$ (which is α for the tent map).

We consider the symbolic dynamics derived from consecutive time steps

$$(s_{n+m}, s_{n+m-1}, \ldots, s_n)\,,$$

with $k \in \mathbb{N}$

$$s_k(x) = \begin{cases} 0 \text{ if } T^k(x) < \alpha \\ 1 \text{ if } T^k(x) \ge \alpha \end{cases}.$$

For comparison, we take a random sequence $\xi_n \in [0, 1]$ (uniformly, i.i.d.), and consider the corresponding symbolic dynamics

$$\sigma_n := \begin{cases} 0 & \text{if } \xi_n < \alpha \\ 1 & \text{if } \xi_n \ge \alpha. \end{cases}$$

The question now is: Are there systematic differences between the symbolic sequence s_n derived from iterations of the tent map and σ_n?

For $\alpha = 1/2$, they look the same (in fact, we simply have a Bernoulli sequence: the values 0 and 1 occur with equal probability $1/2$; $p(0) = p(1) = 1/2$). If we don't know x, s_n looks as random as σ_n. The transition probabilities are

$$p(0|0) = p(1|0) = p(0|1) = p(1|1) = 1/2.$$

We next consider $\alpha = 2/3$. Put $x_n := T^n(x)$.
$\sigma_n = 0$ and $\sigma_n = 1$ occur independently with probabilities $2/3$ and $1/3$.

When $s_n = 1$, that is, $2/3 < x_n \le 1$, then $0 \le x_{n+1} < 2/3$, that is $s_{n+1} = 0$. Thus, there is no transition from 1 to 1. For the state $s_n = 0$, both transitions are equally likely: when $0 \le x_n \le 1/3$, we have $0 \le x_{n+1} \le 2/3$, that is, $s_{n+1} = 0$, while for $1/3 < x \le 2/3$, we get $s_{n+1} = 1$. Thus, for s_n,

$$p(0|0) = p(1|0) = 1/2,\ \ p(0|1) = 1,\ \ p(1|1) = 0$$

while for σ_n

$$p(0|0) = p(0|1) = 2/3,\ \ p(1|0) = p(1|1) = 1/3.$$

This leads us to the concept of *forbidden sequences*. While for the threshold $\alpha = 1/2$, the symbolic dynamics of the tent map cannot be distinguished from that of a random

sequence, and is Markovian, in contrast, for the threshold $\alpha = 2/3$, the sequence 11 does not occur, and the symbolic dynamics is different from a random one, but still Markovian.

For other thresholds, we can also get longer forbidden sequences and non-Markovian symbolics.

Even from a random sequence ξ_n, we can derive non-Markovian symbolic dynamics.

Let $x^1, x^2 \in [0, 1]$; we consider the symbolic rule

$$s(x^1, x^2) = \begin{cases} 0 & \text{if } x^1 \leq x^2 \\ 1 & \text{if } x^2 < x^1. \end{cases}$$

For our random sequence, take $x^1 = \xi_n$, $x^2 = \xi_{n+1}$. Thus, we draw the points x^1, x^2 randomly and independently.

The state probabilities are again $p(0) = p(1) = 1/2$, but the transition probabilities now depend on the history. The more 1s we have seen, the less likely it is to see another 1, because then ξ_n is expected to be very small, hence most likely, $\xi_{n+1} > \xi_n$.

We now analyze the *information flow* of this example. The information flow between the micro-level corresponding to state x_n and the coarse-grained level s_n is the conditional mutual information

$$MI(s_{n+1} : x_n | s_n) = H(s_{n+1}|s_n) - H(s_{n+1}|s_n, x_n)\,.$$

Since s_{n+1} is fully determined by x_n, the second term vanishes,

$$MI(s_{n+1} : x_n | s_n) = H(s_{n+1}|s_n)\,,$$

i.e., the information flow = conditional entropy on the coarse grained level, which has a local minimum at $\alpha = 2/3$.

Instead of drawing information from below, the upper level system relies on its memory.

4 Information Decomposition

We finally turn to the concept of information decomposition. To motivate it, we start with the *transfer entropy* [50][2]

[2] Such a principle had already been introduced by the econometrician Granger [20] who wrote "We say that Y_t is causing X_t if we are better able to predict X_t using all available information than if the information apart from Y_t had been used." In the econometric literature, this principle was applied only in linear settings. As [8], explained, the transfer entropy can be seen as an operationalization of this principle in a general context.

$$TE(Z \to X) := MI(X_+ : Z_-|X_-) \tag{28}$$

where the subscript—refers to the past and $+$ to the future. $TE(Z \to X)$ quantifies the amount of information contained in Z about the future of X that cannot be obtained from its own past.

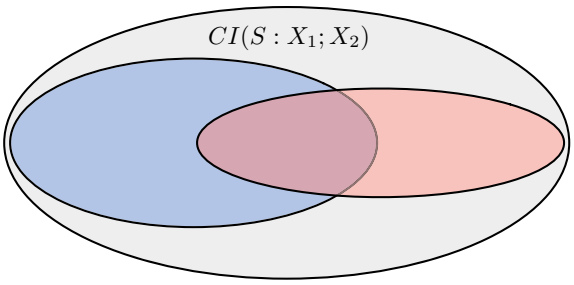
Problem: $X_+ = \mathrm{XOR}(X_-, Z_-)$:
Here, the information in Z_- is only useful together with that of X_-. The transfer entropy cannot distinguish this situation from one where X_- does not contribute and Z_- determines X_+ by itself.

This problem is addressed by information decomposition. It was started by Williams and Beer [63] (but their measure $I_{\min}$ of shared information does not distinguish whether different random variables carry the same information or just the same amount of information), and continued by Harder, Salge, Polani [23], Griffith and Koch [22], Bertschinger, Rauh, Olbrich, Ay, Banerjee, Jost [12, 13, 44, 45], and taken up by many other people (see for instance the references in [34]), with applications in different fields, like neuroscience [62]. There is no optimal solution, but that of Bertschinger, Rauh, Olbrich, Jost, Ay [13] (called the *BROJA* decomposition in the community) is currently the most widely accepted.

To describe our approach, we consider three random variables X_1, X_2 and S. The (total) mutual information $MI(S : X_1, X_2)$ quantifies the total information that is gained about S if the outcomes of X_1 and X_2 are known. How do X_1 and X_2 contribute to this information? For two explanatory variables, we expect four contributions to $MI(S : X_1, X_2)$:

$$\begin{aligned} MI(S : X_1, X_2) = {} & SI(S : X_1; X_2) && \text{shared information} \\ & + UI(S : X_1 \backslash X_2) && \text{unique information of 1} \\ & + UI(S : X_2 \backslash X_1) && \text{unique information of 2} \\ & + CI(S : X_1; X_2) && \text{complementary or synergistic information.} \end{aligned}$$

Here, $UI(S : X_1 \backslash X_2)$ is the information that X_1 has, but X_2 does not have, $SI(S : X_1; X_2)$ is the information that both of them have individually. Perhaps the most interesting term is the last, $CI(S : X_1; X_2)$, the information that only emerges if X_1 and X_2 pool their knowledge. This term is best illustrated in the **XOR** example discussed below.



We consider some examples. **AND**

x_1	x_2	s	$p(x_1, x_2, s)$
0	0	0	1/4
1	0	0	1/4
0	1	0	1/4
1	1	1	1/4

Here, x_1 and x_2 jointly determine s, but cannot be fully recovered from s.

When 1 has the value $x_1 = 0$, she can exclude $s = 1$, and analogously for 2.

Thus, when they both see 0, they share the information that $s = 0$.

The mechanism loses some information. When X_1, X_2 are i.i.d.,

$$H(X_1, X_2) = 2 \text{ bits},$$

but

$$H(S) = MI(S : X_1, X_2) = -\frac{1}{4}\log\frac{1}{4} - \frac{3}{4}\log\frac{3}{4} \approx .811 \text{ bits}.$$

In general, we may have both correlations between the input variables and relations created by the mechanism that computes S.

We next recall **XOR** from Sect. 2.2:

x_1	x_2	s
0	0	0
1	0	1
0	1	1
1	1	0

Neither 1 nor 2 can determine the value of S by herself, but the value of the other is needed for that. This is a clear case of synergistic information only.

Our approach: Unique and shared information should only depend on the marginal distribution of the pairs (S, X_1) and (S, X_2). This idea can be explained from an operational interpretation of unique information: Namely, if X_1 has unique information about S (with respect to X_2), then there must be some way to exploit this information. More precisely, there must be a situation in which X_1 can use this information to perform better at predicting the outcome of S.

In this interpretation, 1 possesses unique information about S compared with 2, if there exists a reward function for which 1 can achieve a higher expected reward based on her value x_1 and her knowledge of the conditional distribution $p(s|x_1)$ than if she knew and utilized instead the conditional distribution of 2.

Thus, unique and shared information depend only on pairwise marginals. Only the synergistic information includes higher order dependencies. In that sense, synergy becomes a measure of higher order interactions, in the sense of information geometry.

From a conceptual perspective, and independently of the way the different terms in the decomposition are quantified, it is important to understand synergy, in order to clarify discussions that have become quite sterile, like the relative importance of genes and environment in biology. For a perspective in this direction, see [27].

References

1. Aristoteles: Philosophische Schriften 5. Metaphysik. Nach der Übersetzung von Hermann Bonitz bearbeitet von Horst Seidl, Felix Meiner, Hamburg (1995)
2. Atay, F., Jalan, S., Jost, J.: Randomness, chaos, and structure. Complexity **15**, 29–35 (2009)
3. Avdiyenko, L., Bertschinger, N., Jost, J.: Adaptive information-theoretical feature selection for pattern classification. In: Computational Intelligence, pp. 279–294. Springer International Publishing (2015)
4. Ay, N.: An Information-geometric approach to a theory of pragmatic structuring. Ann. Prob. **30**(1), 416–436 (2002)
5. Ay, N., Jost, J., Lê, H.V., Schwachhöfer, L.: Information Geometry. Springer, Ergebnisse der Mathematik (2017)
6. Ay, N., Olbrich, E., Bertschinger, N., Jost, J.: A geometric approach to complexity. Chaos **21**, 037103 (2011). https://doi.org/10.1063/1.3638446
7. Ay, N., Loehr, W.: The umwelt of an embodied agent a measure-theoretic definition. Theory Biosci. **134**(3), 105–116 (2015)
8. Barnett, L., Barnett, A., Seth, A.: Granger causality and transfer entropy are equivalent for Gaussian variables. Phys. Rev. Lett. **103**(23), 38701 (2009)
9. Bertschinger, N., Olbrich, E., Ay, N., Jost, J.: Autonomy: an information theoretic perspective. Biosystems **91**, 331–345 (2008)
10. Bertschinger, N., Rauh, J., Olbrich, E., Jost, J.: Shared information—new insights and problems in decomposing information in complex systems. Proc. ECCS 251–269 (2012)
11. Bertschinger, N., Rauh, J., Olbrich, E., Jost, J., Ay, N.: Quantifying unique information. Entropy **16**, 2161–2183 (2014). https://doi.org/10.3390/e16042161
12. Bertschinger, N., Rauh, J., Olbrich, E., Jost, J.: Shared information–new insights and problems in decomposing information in complex systems. In: Proceedings of the European Conference on Complex Systems, Brussels, Belgium, 2–7 September 2012, pp. 251–269
13. Bertschinger, N., Rauh, J., Olbrich, E., Jost, J., Ay, N.: Quantifying unique information. Entropy **16**, 2161–2183 (2014)
14. Chaitin, G.: On the lengths of programs for computing finite binary sequences. JACM **13**, 547–569 (1966)
15. Crutchfield, J.P., Young, K.: Inferring statistical complexity. Phys. Rev. Lett. **63**, 105–108 (1989)
16. Efer, T., Heyer, G., Jost, J.: Shakespeare unter den Deutschen (C. Jansohn, Hrsg.)
17. Friston, K.: The free-energy principle: a unified brain theory? Nat. Rev. Neurosc. **11**, 127–138 (2010)
18. Gell-Mann, M., Lloyd, S.: Information measures, effective complexity, and total information. Complexity **2**, 44–52 (1996)
19. Gigerenzer, G., Todd, P.: Simple Heuristics that Make us Smart. Oxford University Press (1999)
20. Granger, C.: Investigating causal relations by econometric models and cross-spectral methods. Econometrica **37**, 424–438 (1969)
21. Grassberger, P.: Toward a quantitative theory of self-generated complexity. Int. J. Theor. Phys. **25**, 907–938 (1986)
22. Griffith, V., Koch, C.: Quantifying synergistic mutual information. In: Prokopenko, M. (ed.) Guided Self-Organization: Inception, vol. 9, pp. 159–190. Springer, Berlin/Heidelberg, Germary (2014)

23. Harder, M., Salge, C., Polani, D.: A bivariate measure of redundant information. Phys. Rev. E **87**, 012130 (2013)
24. Jaynes, E.T.: Probability Theory: The Logic of Science. Cambridge University Press (2003)
25. Jost, J.: External and internal complexity of complex adaptive systems. Theory Biosc. **123**, 69–88 (2004)
26. Jost, J.: Sensorimotor contingencies and the dynamical creation of structural relations underlying percepts. In: Proceedings of the Strüngmann Forum: Where's the Action? The Pragmatic Turn in Cognitive Science. MIT Press, im Druck (2016)
27. Jost, J.: Biological information. Theory Biosci. (2020)
28. Jost, J., Bertschinger, N., Olbrich, E., Ay, N., Frankel, S.: An information theoretic approach to system differentiation on the basis of statistical dependencies between subsystems. Phys. A **378**, 1–10 (2007)
29. Jost, J., Bertschinger, N., Olbrich, E.: Emergence. A dynamical systems approach. New Ideas Psych. **28**, 265–273 (2010)
30. Jost, J., Scherrer, K.: Information theory, gene expression, and combinatorial regulation: a quantitative analysis. Theory Biosci. **133**, 1–21 (2014)
31. Kolmogorov, A.N.: Three approaches to the quantitative definition of information. Probl. Inf. Transm. **1**, 4–7 (1965)
32. Klyubin, A., Polani, D., Nehaniv, C.: Empowerment: a universal agent-centric measure of control. Proc. IEEE CEC **2005**, 128–135 (2005)
33. Li, M., Vitányi, P.: An Introduction to Kolmogorov Complexity and Its Applications. Springer (1997)
34. Lizier, J., Bertschinger, N., Jost, J., Wibral, M.: Information decomposition of target effects from multi-source interactions: perspectives on previous, current and future work. Entropy **20**, 307 (2018). https://doi.org/10.3390/e20040307
35. Luhmann, N.: Soziale Systeme. Suhrkamp, Frankfurt (1984)
36. Luhmann, N.: Die Gesellschaft der Gesellschaft, 2 Bde. Suhrkamp, Frankfurt (1997)
37. Maturana, H., Varela, F.: Autopoiesis and Cognition. Reidel, Boston (1979)
38. Moore, C., Mertens, S.: The Nature of Computation. Oxford University Press (2011)
39. Papadimitriou, C.: Computational Complexity. Addison Wesley (1994)
40. Pfante, O., Olbrich, E., Bertschinger, N., Ay, N., Jost, J.: Closure measures for coarse-graining of the tent map. Chaos **24**, 013136 (2014)
41. Pfante, O., Bertschinger, N., Olbrich, E., Ay, N., Jost, J.: Comparison between different methods of level identification. Adv. Complex Syst. **17**, 1450007 (21 Seiten) (2014)
42. Pfante, O., Bertschinger, N., Olbrich, E., Ay, N., Jost, J.: Wie findet man eine geeignete Beschreibungsebene für ein komplexes System? In: Jahrbuch der Max-Planck-Gesellschaft, im Druck (2016)
43. Pfeifer, R., Bongard, J.: How the Body Shapes the Way we Think. MIT Press (2007)
44. Rauh, J., Banerjee, P., Olbrich, E., Jost, J., Bertschinger, N.: On extractable shared information. Entropy **19**(7), 328 (2017)
45. Rauh, J., Bertschinger, N., Olbrich, E., Jost, J.: Reconsidering unique information: towards a multivariate information decomposition. In: Proceedings of 2014 IEEE International Symposium on Information Theory (ISIT), Honolulu, HI, USA, 29 June–4 July 2014, pp. 2232–2236
46. Rissanen, J.: Stochastic Complexity in Statistical Inquiry. World Scientific (1989)
47. Sadovsky, M.: Genes, information and sense: complexity and knowledge retrieval. Theory Biosci. **127**(2), 69–78 (2008)
48. Scherrer, K., Jost, J.: The gene and the Genon concept: a functional and information-theoretic analysis. Mol. Syst. Biol. **3**, 87, Epub 2007 Mar 13: EMBO and Nature Publishing Group
49. Scherrer, K., Jost, J.: Gene and Genon concept: coding versus regulation. Theory Biosc. **126**, 65–113 (2007)
50. Schreiber, T.: Measuring information transfer. PRL **85**, 461–464 (2000)
51. Shalizi, C.R., Crutchfield, J.P.: Computational mechanics: pattern and prediction, structure and simplicity. J. Stat. Phys. **104**, 817–879 (2001)

52. Shannon, C.: The Mathematical Theory of Communication (1948), reprinted in: Shannon, C., Weaver, W. The Mathematical Theory of Communication. Blahut, R., Hajek, B. (eds.), pp. 29–125. Univ. Illinois Press (1998)
53. Shaw, R.: The Dripping Faucet as a Model Chaotic System. Aerial Press, Santa Cruz (1984)
54. Simon, H.A.: A behavioral model of rational choice. Quart. J. Econom. **69**(1), 99–118 (1955)
55. Simon, H.A.: Rational choice and the structure of environments. Psych. Rev. **63**, 129–138 (1956)
56. Sinai, J.: On the concept of entropy for a dynamical system, (Russian) Dokl. Akad. Nauk SSSR **124**, 768–771 (1959)
57. Solomomoff, R.: A formal theory of inductive inference, Part I. Inf. Control **7**, 1–22. Part II. Inf. Control **7**, 224–254 (1964)
58. Tononi, G., Sporns, O., Edelman, G.M.: Measures of degeneracy and redundancy in biological networks. PNAS **96**, 3257–3267 (1999)
59. Vapnik, V.N.: The Nature of Statistical Learning Theory. Springer (1995)
60. Vapnik, V.N.: Statistical Learning Theory. Wiley-Interscience (1998)
61. von Uexküll, J.J.: Umwelt und Innenwelt der Tiere. In: Mildenberger, F., Herrmann, B. (eds.) Klassische Texte der Wissenschaft. Springer, Spektrum (2014)
62. Wibral, M., Priesemann, V., Kay, J.W., Lizier, J.T., Phillips, W.A.: Partial information decomposition as a unified approach to the specification of neural goal functions. Brain Cogn. **112**, 25–38 (2017)
63. Williams, P., Beer, R.: Nonnegative decomposition of multivariate information. arXiv:1004.2515v1 (2010)
64. Zambella, D., Grassberger, P.: Complexity of forecasting in a class of simple models. Complex Syst. **2**, 269 (1988)

Complex Systems: From the Presocratics to Pension Funds

Frédéric Patras and Victor Planas-Bielsa

Abstract Complexity, as investigated in biology, chemistry, mathematics, physics and the humanities requires new modes of thinking, beyond the mechanistic and reductionist standards. We investigate it mostly from a philosophical point of view, tracing back some of the relevant questions to Greek philosophers. Examples are taken from probability, economics and finance to feature typical phenomena. We detail in particular the one of defined benefit pension funds that points out at various important issues in contemporary finance and economics.

Keywords Complex systems · Presocratic philosophies · Ionian school · Atomism · Logical atomism · Defined benefit pension fund · Card shufflings

1 Introduction

The notion of complexity, and the related one of complex systems, have a wide acception and many fields of application: to quote only a few, living beings in their globality in biology, energy levels and other properties of large molecules in chemistry, chaotic dynamical systems in mathematics and physics, various forms of social networks in social sciences, and so on. Whereas it is difficult to encompass all these forms of complex systems under a single generic pattern or set of axioms, it is generally agreed that one of their main common features is that they cannot be treated

[1] We refer to the other contributions in the present volume for an illustration of these ideas in various fields of application.

F. Patras (✉)
Laboratoire J.-A. Dieudonné, University Côte D'Azur, CNRS, UMR 7351, Parc Valrose 06108, Nice Cedex 02, France
e-mail: Frederic.Patras@unice.fr

V. Planas-Bielsa
Département de Biologie Polaire, Centre Scientifique de Monaco, Principality of Monaco, Quai Antoine 1er 98000, Monaco

S. Albeverio et al. (eds.), *Complexity and Emergence*, Springer Proceedings in Mathematics & Statistics 383, https://doi.org/10.1007/978-3-030-95703-2_5

with a reductionist approach: that is, splitting the system into elementary components whose deterministic interactions would allow to explain its global behaviour.[1]

This insight paves the way to various philosophical and epistemological theories. In many respects, one can argue that complex systems could and even have to produce their own epistemology. For instance, the idea of chaos in dynamical systems born with the work of Poincaré has led to numerous reflections on causality, determinacy and more generally on the meaning that has to be given to deterministic mathematical models so sensitive to initial conditions that the unavoidable uncertainty on them makes the evolution of the system unpredictable (at least in the usual sense of the word).

This process of interactions between science and philosophy, where the unforeseeable progress of science guides the development of philosophy, which in turn provides tools to understand conceptually the scientific revolutions, is typical of what is often refered to as *historical epistemology*. The theory was born with G. Bachelard[2] in the first part of the 20th century and is currently revived by Ian Hacking,[3] Lorraine Daston,[4] Hans-Jörg Rheinberger[5] and many others.

Using this method in the context of complex systems makes sense. In the present work we will however follow another approach.[6] Building on the insight that considering complexity requires a rethinking of the very foundations of science, we will argue that another approach is possible, namely by addressing directly the question of the interplay between complexity and simplicity. Here, by simplicity we mean the hope of science and theoretical knowledge to reconduct science to simple, fixed, permanent elements, a hope that cannot be accounted for only by the idea of reductionism. The interplay between simplicity and complexity is evolving fast, however some of its features remain constant and various problems raised by the Greek thinkers at the very beginning of philosophy are still meaningful today. This will be the subject of the first section, with, as a modern illustration of the philosophical problems discussed, an example taken from 20th century mathematics: a

[2] Lecourt [1], Brenner [2].

[3] Hacking [3].

[4] Daston and Galison [4].

[5] Rheinberger [5].

[6] This text is composed of two distinct contributions. The lectures given by the first author on the philosophy of complexity at the 2018 School *Complexity and Emergence: ideas, methods, with a special attention to economics and finance* form the content of the first three sections. The last section, on pension funds, is instead a joint work: it is based on a working paper written by the two authors in 2007 for the Monaco Hedge Funds Research Institute that was directed by the second author. The paper introduced a model aiming at describing some of the risks embedded in defined benefit pension funds. Soon after, the 2007–2008 financial crisis occurred and our interests shifted. The project remained dormant and the paper unpublished—after an unsuccessful submission at the Journal of Applied Corporate Finance we had realized that the model needed a serious upgrade to be used in practice. However, on the one hand the questions that we raised appear to be timely again (see the references in Sect. 5). On another hand, in spite of its quantitative shortcomings (structural models are difficult to implement quantitatively), we also believe that, as it is, it illustrates nicely various risk and complexity related phenomena.

convergence phenomena for Markov chains, due originally to Poincaré but revisited recently.

The second section will deal with the ideas of model and causality. Feynman diagrams are a classical example in modern physics of entities whose ontological status in uncertain. Are they simple computational tools, or do they reflect actual physical phenomena? To address this kind of questions, epistemological tools are lacking in contemporary philosophy of science. As in the first section, we appeal to Greek philosophy, not so much to find solutions and answers, but other ways of thinking. Concretely, we survey Aristotle's physics and more specifically his theory of causes -only one of which fits modern, post-galilean science: formal causes. The section concludes with two examples of applications of Aristotle's ideas: quantum field theory and the Black-Scholes paradigm in mathematical finance.

The third section treats various forms of atomism. Russell and Wittgenstein's logical atomism is historically one of the most interesting, in spite of well identified drawbacks. We advocate the meaningfulness of the notion of mathematical atomism to account for various phenomena, some of them internal to mathematics and related to the notion of axiomatic systems, others related to applications of mathematics. We develop from this point of view the example of Asset Backed Securities (ABS). The section concludes with a plea for considering mathematics as a whole as a dynamical system. We also argue that historical epistemology and philosophy of concepts in Cavaillès' sense support these views.

The fourth and last section deals with one of the key problems of contemporary economics: the funding of pensions and the ageing of populations. We focus on a specific issue and present a toy model for the quantitative appraisal of some of the risks embedded in Defined Benefit Pension Funds (DBPF). From the epistemological point of view, the lesson to be learned from the model is that mathematical atomism (the decomposition of a problem into a family of "atomic models" of its elementary components) tends structurally to overlook the effects of interactions between these components. This is particularly obvious for DBPF for which the risks due to the correlation between the sponsor's firm value and the fund assets seems to have been for long overlooked.

2 Presocratic Philosophy Revisited

2.1 Heraclitus and the Philosophy of Nature

Presocratic philosophy is classically divided into various schools:

- Thales (c. 625 BC, c. 546 BC), the "first philosopher", and the Ionian school with Anaximander, Anaximenes and Heraclitus (c. 535 BC, c. 475 BC), that started to develop the theoretical analysis of nature,
- the school of Pythagoras (c. 570 BC, c. 495 BC),

- Parmenides (end of 6th century BC, beginning of 5th) and the Eleatic school with Zeno, maybe Empedocles.

Whereas pythagoreanism stands apart due to its stance on the role of numbers and arithmetic, the Ionian and Eleatic schools correspond to two radically different ways of thinking about the world, about nature.

The Ionian school features movement, interaction between elementary components (fire, earth, water, …). The most interesting of its members, at least in the context of this article, is Heraclitus, the so-called "obscure philosopher". He relates the idea of movement and changes to the question of the very possibility of theoretical knowledge. The world is in a permanent flux and the stability of what surrounds us is misleading. Of his writings only few fragments remain,[7] most of which are classical, such as:

> We step and do not step into the same river,

or,

> You cannot step twice into the same river.

The conclusion drawn from these fragments, due also to Heraclitus, is that:

> All the objects of the senses are in a perpetual flux and cannot be the subject matter of science.

What is still meaningful in Heraclitus is this idea that dynamics, time evolution of phenomena can be intrinsically an obstacle to the building of theoretical knowledge.

In modern terms, what is at stake is the key ontological difference between objects conceived as stable, permanent, sometimes eternal beings, and the moving reality of real beings such as Heraclitus' river.

2.2 Parmenides and the Modern Idea of Science

Following a tradition in the history of philosophy that has its roots in Plato, Parmenides is the great opponent to Heraclitus. Whereas changes, flux, movement are keywords for Heraclitus' philosophy, the One, the unit, unity are the central ones in Parmenides'.[8] The world undergoes a perpetual movement, a constant evolution, but the use of concepts allows us to grasp a unity behind the flow. Heraclitus' river is again a good example: it is constant as a conceptual reference, but in perpetual evolution as an object of the senses. Unity is therefore the ultimate principle of science because it is the key to the constitution of units of signification. To think of an object, a thing, and to name it, means to gather together, in a unit, in a totality, the diversity of its positions in space and time, its possible changes of form.

---

[7] Heraclitus. Fragments. In Voilquin [6].

[8] On the One and the unit in Parmenides, see Patras [7].

With Parmenides, the idea that there is an opposition between the logical thinking and the empirical one, between science and doxa, science and opinion emerges. This opposition is important because it excludes from the realm of science any discourse that would not fit the constraints of logic –whatever is meant by logic. It is also clear that living systems, for example, hardly fit in this conception of science, whereas mathematics or (Greek) astronomy that study fixed, eternal entities, would be paradigmatic examples of sciences in this framework. By many respects, consciously or not, our conception of science remains largely dependent on these early views.

Today, when we face the problems raised by Heraclitus and Parmenides, several answers are possible. One one hand, we know that very often the dynamics of phenomena is driven by permanent laws (think of gravity, electromagnetism, quantum mechanics …). As such, the corresponding changes are intelligible and can be the subject matter of science. Classical physics relies on this ground. Quantum mechanics raises already several issues: as far as its basic laws and principles are concerned, they are fixed and therefore plainly "scientific". However, the uncertainty intrinsic to the very notion of quantum states and measurements is already more problematic to address, and there are still lively debates on the interpretations of quantum mechanics -De Broglie-Bohm versus Copenhagen for example.

Of a very different nature is the question whether or not there is a science of facts that do not obey fixed laws or that obey laws that we cannot expect to discover in the present state of knowledge. Economics and finance provide examples: equilibrium theory, risk neutral pricing and other similar key notions and principles are, structurally, only an approximation of the "real" behaviour of markets. Biology, ecology and living organisms raise still other problems of the same kind.[9]

Science is a difficult notion to define and, depending on the given meaning, many theories can or cannot be considered as sciences, from mathematics to the humanities. In this context, Heraclitus' questioning is still meaningful: discussing on the very possibility of theoretical knowledge in relation to dynamics, changes, transformations, allows to avoid restricting the debates to the Eleatic view, implicitly dominant whenever science is discussed.

2.3 Democritus and Atomism

The opposition between Eleatic and Ionian philosophers is a very deep and structuring one. On one side, we have logical requirements, paving the way to mathematical ones. On the other side, we would also like to have global views on nature, being able to understand global structures, dynamical variations and changes. Whereas classical mathematics and physics are largely Eleatic, the mathematical and natural phenomena studied in the context of complex systems (in the broadest sense of the term) suggest a synthesis of the two approaches.

[9] Bailly and Longo [8].

From this point of view, Democritus (c. 460 BC, c. 370 BC) is an interesting philosopher, and it would be tempting to suggest that philosophical investigations on complexity and emergence first started when his philosophy was discussed and challenged. He is sometimes considered as a presocratic, although younger than Socrates. His philosophy, atomism, could have started as a reflexion on Parmenides', but leading to very different conclusions:

> The metaphysical problem [faced by Democritus] is the same that arose for Anaxagoras and Empedocles, following the Eleatic criticism of change. How to reconcile the immutability and the eternity of being with the reality of movement and change, the "way of truth" with that of "opinion"?[10]

For Parmenides, there is nothing excepted pure being, eternal, perfectly homogeneous, finite and perfect, excluding movement and transformation. Democritus' idea was that these features make sense for the ultimate components of matter, the atoms. However, from the existence of a plurality of atoms follows also the possibility of motions, interactions, movement and evolution. The main problem that emerges from atomism is ultimately how to account for the phenomena starting from atoms?

> Without being augmented with a principle of internal stability (such as, for example, the stoicists' *pneuma*), [Democritus' theory of atoms] does not seem able to account for the cohesion of bodies. Leibniz would say that it is lacking a *vinculum substantiale* holding the atoms together. The difficulty is characteristic of any theory that explains the "complexes" by mere aggregation of the "simples" –be it Greek atomism, the monadic composite or (in the 20th century) the logico-atomistic constructions of the world.[11]

There is indeed a huge gap between the idea of elementary components of matter and large scale phenomena as we observe them. In modern science, this phenomenon can sometimes become a well-identified problem: for example, decoherence, the disappearance of quantum effects showing up in large quantum systems, is still only very partially understood. This raises considerable tecnhological problems when trying to build quantum computers with a large number of qbits, one of the current technological challenges.

2.4 Markov Chains

The problem of emergence of patterns out of elementary "blocks" and their interactions is manifold: there is a wide variety of situations where such phenomena occur. An interesting cognitive fact is that we are used to such phenomena, and we know some implicit rules of their emergence out of experience, although without being able most often to explain the underlying reasons. In that sense, we have a prescientific understanding of these phenomena. For example, we are not surprised by the group flight of tens or hundreds of birds creating a moving cloud.

---

[10] Gil [9].

[11] F. Gil, op. cit.

Limit theorems in probability theory are an important source of examples of patterns emerging out of elementary schemes. It occurs often that we are aware of the existence of these patterns and of their concrete meaning and practical consequences. A striking illustration dates back from the beginning of the 20th century. It is not an example of complex system: it is infinitely much simpler than the behaviour of interacting systems such as birds' clouds. However, in spite of its simplicity, it is underlying the long term behaviour of many dynamical systems and is also showing on a very concrete example that we have a fairly deep spontaneous intuition of relatively complex phenomena that most of us would fail to explain if they were asked to.

The example dates back to Poincaré's treatise on probabilities[12] and originated one of the main research lines in probability and statistics: the evolution of random processes (time-dependent random variables) and random chains (sequences of random variables). The problem is the following: start with a deck of cards. We know intuitively and by experience that if we mix the deck randomly enough, by repeated shufflings, no information will be available after the mixing, and no one will be able to take advantage of the ordering of the cards in the deck. Mathematically, this means that iterated random shufflings of the deck create a random distribution that is close to the uniform one (the distribution where all orderings have the same probability).

Poincaré's analysis of the problem is essentially[13] as follows. To mix a deck, one splits it randomly into two decks and then mix the two decks by shuffling the cards. This latter operation amounts to randomly selecting a card in one of the two decks, put it on a new deck, and repeat this operation till the mixing is complete. Let us detail the simplest possible case which is already not completely trivial: a deck of two cards. The general case can be treated exactly along the same lines with some knowledge of elementary linear algebra.[14]

Starting from a deck of two cards 1, 2 with 1 on the top (we write the configuration 12), splitting and shuffling gives the following four possibilities denoted by arrow diagrams[15]:

$$12 \longmapsto (12, \emptyset) \longmapsto 12$$

$$12 \longmapsto (1, 2) \longmapsto 12, 21$$

$$12 \longmapsto (\emptyset, 12) \longmapsto 12.$$

[12] Poincaré [10].

[13] We treat the case of so-called perfect shuffles. Poincaré's analysis is actually more general, allowing essentially for arbitrary non trivial mixings.

[14] The reader is invited to read Poincaré's treatise, astonishingly modern and insightful –the modern accounts of these phenomena are actually very similar to Poincaré's.

[15] Pairs in the middle denote the two decks resulting from the splitting, for example $(1, 2)$ means that the first deck is the card 1, the second the card 2, and so on. On the right are the possible outcomes of the shuffling. For example $12, 21$ are the two decks that can be obtained by shuffling 1 and 2.

Assuming that all mixing paths are equally likely, we get that the outcome of a random shuffle starting from the deck 12 is 12 with probability 3/4 and 21 with probability 1/4.

It is convenient to encode this process by a transition (or Markov) matrix:

$$M = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{pmatrix}$$

so that, starting with a probability distribution: $P(12) = p$, $P(21) = 1 - p$, we get after a perfect shuffle the new distribution $Q(12) = 1/2\ p + 1/4$, $Q(21) = 3/4 - 1/2\ p$ since:

$$M \begin{pmatrix} p \\ 1-p \end{pmatrix} = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{pmatrix} \begin{pmatrix} p \\ 1-p \end{pmatrix} = \begin{pmatrix} 1/2\ p + 1/4 \\ 3/4 - 1/2\ p \end{pmatrix}$$

The key idea of Poincaré is that the time evolution of probability distributions is governed by the spectral analysis of the matrix M, which has a first eigenvalue 1, corresponding to the uniform distribution $U(12) = 1/2,\ U(21) = 1/2$ (that is, the uniform distribution is stable under perfect shuffles: it is called the equilibrium distribution). The second eigenvalue is $1/2$ with eigenvector

$$\begin{pmatrix} 1/2 \\ -1/2 \end{pmatrix}.$$

In general, for a deck of N cards, one can show that the spectrum of the Markov matrix describing perfect shuffles is $1, 1/2, \ldots, 1/2^{N-1}$ with the uniform distribution spanning the eigenspace associated to the top eigenvalue 1. A group theoretical method to obtain this result is indicated below.

Coming back to the situation where one starts from the deck 12 and applies k perfect shuffles, the resulting distribution Q_k is obtained by computing

$$M^k \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} + 1/2^k \begin{pmatrix} 1/2 \\ -1/2 \end{pmatrix}$$

so that, $Q_k(12) = 1/2 + 1/2^{k+1}$ and $Q_k(21) = 1/2 - 1/2^{k+1}$, which converges exponentially fast to the uniform distribution U.

This phenomenon is called convergence to equilibrium of Markov chains and applies in a wide variety of situations. The case of perfect shuffles that we just described has actually two mathematical interpretations in the literature, with two different epistemological implications. We sketch only the mathematical ideas and refer the interested reader to the literature.

The first one is essentially the one due to Poincaré. Assume that the random evolution of a discrete system is described by a Markov matrix such as M. Then, under relatively mild "mixing" conditions, this matrix has an isolated eigenvalue 1 whose eigenspace is associated to an invariant equilibrium distribution P. The modules

of the other eigenvalues are then strictly less than 1, from what one deduces the exponential convergence of the system to the equilibrium distribution. This property can be used for example to devise Monte Carlo methods (probabilistic methods to approximate numerically random distributions).

The second interpretation is more recent and in the spirit of mathematical structuralism: treating the problem by appealing to general properties of algebraic structures. It applies to a small class of random systems but is grounded on another large class of (group-theoretical) phenomena. The process of splitting a deck of cards into two subdecks and shuffling the resulting two decks is typical of a general combinatorial principle: in many situations in combinatorics (cards, words, but also partially ordered sets, finite topological spaces …), such a splitting is encoded by a coalgebra structure (the structure dual to the one of algebra) formally defined here by

$$\Delta(x_1 \dots x_n) = \sum_{i=0}^{n} x_1 \dots x_i \otimes x_{i+1} \dots x_n,$$

(where $x_1 \dots x_n$ stands for a deck of cards labelled $x_1, \dots, x_n$) whereas the mixing is encoded by a product defined recursively by

$$x_1 \dots x_n \times y_1 \dots y_m = x_1(x_2 \dots x_n \times y_1 \dots y_m) + y_1(x_1 \dots x_n \times y_2 \dots y_m).$$

With two cards, we recover our earlier computations in algebraic form: $\Delta(12) = 12 \otimes \emptyset + 1 \otimes 2 + \emptyset \otimes 12$, whereas $12 \times \emptyset = 12 = \emptyset \times 12$ and $1 \times 2 = 12 + 21$, so that $\times \circ \Delta(12) = 3 \cdot 12 + 21$ (three times the configuration 12, one time 21). The splitting and the mixing define together a bialgebra or Hopf algebra structure.[16] These ideas were first emphasized by Rota and coauthors,[17] giving rise to the use of Hopf algebra techniques in combinatorics, a very active approach in the field for at least 20 years.

Commutative Hopf algebras such as the one that we just defined abstracting the definition of perfect shuffles are naturally associated to groups. Groups and commutative Hopf algebras are "dual" notions: commutative Hopf algebras can be though of as algebras of functions on groups. The spectral analysis of Markov transitions associated to perfect shuffles appear then as a special case of spectral phenomena occurring when studying, at the level of functions, power maps $x \longmapsto x^k$ on groups.[18]

In the end, what we would like to emphasize with these examples (discrete random dynamical systems, group theory, iterations and power maps, exponential convergence to equilibrium …) is that we have a pre-theoretical and intuitive understanding of many phenomena. Certain are based on our daily experience of the world, others are more complex: the convergence to the uniform distribution of cards by iterated shufflings implicitely grounds the way we play card games and the idea of how to

[16] These structures are studied in detail in Reutenauer [11], Cartier and Patras [12].

[17] Joni and Rota [13].

[18] The general theory was developed in Patras [14–16]. The application of these techniques to card shufflings is more recent and was obtained in Diaconis et al. [17].

play "fairly", but the underlying intuition is most likely not based only on experience, but also on prescientific views on probability and randomness.

The two general mathematical explanations we have presented of the convergence to equilibrium phenomenon for deck of cards under iterated shufflings point out at two different orders of phenomena, both deeply grounded in our pre-mathematical understanding of the world. The convergence of Markov chains can be related to the very idea of randomness. For example, we know that if we move in a forest alternating random forward moves and random turns, we will be lost pretty fast, although explaining this basic fact through a theoretical model is certainly not straightfoward -one would now use models such as Lévy flights.[19] Group and other composition laws, power maps, are still another family of basic intuitions with different epistemological and phenomenological roots. Following Dedekind's approach for example,[20] power maps of functions could be the intuition grounding the construction of natural numbers. These ideas are still discussed in the Philosophy of Mathematics literature, among others in relation to structuralism and the so-called Benacerraf dilemna.[21]

Philosophically, these insights resort largely to Husserl's views as exposed in his book on the epistemological crisis in modern science[22]: even the more sophisticated scientific constructs would rely in the end on our fundamental intuitions and experiences of the world. According to his views, disentangling science from its intuitive roots would be a dangerous and counterproductive attempt.

3 Models and Causes

The atomistic and other reductionist approaches are bottom up: we start from elementary components, a dynamics or interaction rules, and try to grasp what happens at higher levels. This approach is typical, for example, of modern mathematical finance. Here, in the paradigmatic approach, mark-to-market valuation and risk neutral probabilities, the elementary components are all the available market data: stock prices, interest rate curves, prices on futures on commodities, swaps and swaption prices …. From these data, that are assumed to account for all the available knowledge on financial and economic entities, one should be able to account also for the long term behaviour of complex assets. The example of how this strategy, applied to Residential Mortgage-Backed Securities (RMBS) and other ABS, resulted in the 2007–2008 financial crisis is relatively well-known—we will come back to these questions later.

---

[19] Lévy flights are used to model various natural phenomena, a popular one being sharks foraging. Application fields include finance, earthquakes…See e.g., also for references on the subject, Humphries and Sims [18].

[20] Dedekind [19].

[21] Benacerraf [20].

[22] Husserl [21].

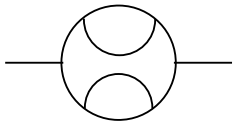


Fig. 1 The shape of a given Feynman diagram in ϕ^3 theory

Here, we will consider the other possible approach: starting from the phenomena as we observe them and trying to understand their structure, their behaviour. Modern, post galilean, physics has taught us how to proceed: namely by expressing everything in mathematical language. There is no reason to depart from this programmatic idea –however another question should still be addressed: what is intended and expected exactly by a mathematical modeling? For example, should the model account for the quantitative and dynamical properties of phenomena (in the sense that one could read in the model the underlying "reasons" for its happening), or should it simply describe them and give rise to the best possible predictions?

3.1 The Example of Feynman Diagrams

These seemingly innocuous questions are not so easy to answer, even in very abstract and theoretical frameworks. A classical example is provided by Feynman diagrams in perturbative quantum field theory (QFT). QFT is one among the most. If not the most surprising scientific achievements ever. The so-called standard model, describing the elementary particles and their interactions, is predicting phenomena at an amazing level of precision. The underlying calculations are based on Feynman diagrams[23] such as (Fig. 1).

For those not familiar with them, they can be thought of by analogy with Taylor series expansions of functions (or, better, with the solutions of differential equations obtained by Picard iterations, but we stick here to the more familiar example of Taylor expansions)

$$f(x) + f'(x)(y-x) + \frac{f''(x)}{2!}(y-x)^2 + \frac{f^{(3)}(x)}{3!}(y-x)^3 + \dots.$$

Whereas the components of a Taylor series aim at approximating a function using its successive derivatives, Feynman diagrams expansions aim at expressing the quantities relevant in the analysis of particle physics in terms of the fundamental interactions between these particles (the diagrams parametrize the terms of the perturbative expansion). The diagrams are built out of vertices with incoming and outgoing edges representing these interactions (for example 3 edges pro vertex in ϕ^3 theory as in the Figure above).

[23] For an epistemology-minded introduction, we refer to Brown [22].

> Feynman diagrams are now iconic. Like pictures of the Bohr atom, everyone knows they have something important to do with physics. Those who work in quantum field theory, string theory, and other esoteric fields of physics use them extensively. In spite of this, it is far from clear what they are or how they work. Are they mere calculating tools? Are they somehow pictures of physical reality? Are they models in any interesting sense? Or do they play some other kind of role?[24]

In the same article, J. R. Brown notices that whereas they clearly are efficient calculation tools, going beyond this general statement is difficult:

> If you ask me how to get from Toronto to Montreal, I could respond in two ways: (1) I could tell you to drive north until you reach the main highway, then turn right and continue on for about five hours, or (2) I could give you a map and tell you where you presently are on it. Both ways provide the information to get you successfully to Montreal. The map in the second method is clearly a model; the instruction in the first method is clearly not.

He argues then that Feynman diagrams "are a lot like (1) in spite of appearing a lot like (2). In other words, they are not pictures or descriptions of reality, nor are they models in any reasonable sense". Other physicists would probably disagree with his views—arguing, for example, that one should not look for more than efficient computation tools: these would be the ultimate "models" and there would be no reason to look for an explanation beyond them.

To restate these ideas, even very classical physical models such as the standard model of particle physics raise ontological problems. Feynman diagrams are just a particularly meaningful example: it is not clear whether they are mere computational tools or correspond, at least partially, to actual physical phenomena and therefore exist as models of actual physical processes.

3.2 Aristotle on Causality

To go on with the program of the first section, we would like to analyse these questions by going back, again, to Greek philosophy instead of appealing to the current debates in science and epistemology. The underlying idea is to broaden the spectrum of point of views that can be used when trying to understand contemporary science, some problems raised by ancient philosophers keeping some relevance in spite of the context in which they were stated.

Modern philosophy, as we know it, was really born with Plato and Aristotle. Consciously or not, we are still much more dependent on the way they thought about what science is and should be than many would believe. Modernity has kept and developed certain of their ideas, but lost contact with other ones.

Plato, as far as theoretical knowledge is concerned, followed largely Parmenides. Aristotle instead had quite different views, and the ones he had on Physics, although often sharply criticized since Galileo, could still have some meaningful features. The classical post-galilean views on Aristotle are not without a ground: he featured

[24] Brown op. cit.

a clear cut distinction between cosmology—the sky and the stars having perfect mathematical movements of which there can be a science—and the sublunar world that would be understandable only qualitatively. There would be therefore no room for modern physics and modern science in a philosophy of science filled with qualities and substances and void of quantitative and experimental methods.

Although largely true, this understanding misses many interesting ideas. As Galileo himself observed, one should not indeed confuse Aristotle's general approach of science and epistemology and the applications he made of them to physics and natural sciences as they existed five centuries B.C.

What Aristotle wanted to understand, as a philosopher, is once again related to presocratic philosophies: the fact that there is a contradiction between the constant changes that physical beings are enduring and the permanence of knowledge. However, according to him, we cannot exclude from the field of theoretical knowledge movement and changes, contrary to what Parmenides and Plato had suggested. His Physics is therefore in the end mainly concerned with the essence of movement, of becoming -in opposition to the Eleatic study of the essence of beings, of permanence, of the underlying substance.

> Aristotle's physics is not "physics" in modern sense. Or, it is not the main sense of physics in his work. It is enough to think of the fact that its object is what, in a programmatic sense, modern physics avoids to consider a theme of inquiry. The subject matter of Aristotle's text is the $\varphi\acute{\upsilon}\sigma\iota\zeta$ [nature] and the things that belong to it. And, as the latter is characterized as having in itself the principle of movement, the meaning and the structural conditions of movement form the content of the aristotelician tradition.[25]

Moreover, and this point is also essential, movement for Aristotle does not only mean mechanical movement, but all the transformations that we can observe:

> What we indicate with the term "movement" translates the two notions that Aristotle uses often indifferently and as synonymous, that is $\kappa\acute{\upsilon}\eta\sigma\iota\sigma$ and $\mu\varepsilon\tau\alpha\beta o\lambda\acute{\eta}$, contains in itself the various forms of movement: generation and corruption, alteration, increase and decrease, translation; that is, using the categories as a reference scheme, movement according to substance, quantity, quality and position.[26]

Understood in that way, Aristotle's physics is another attempt, quite different from atomism but equally meaningful, to go beyond the presocratic apory and the opposition between Ionian and Eleatic philosophies. One of the problems that Aristotle faced was to understand general notions such as the infinite; space; the vacuum; time; the continuum. All these questions have been central to the 20th century mathematics and science; they all have an intrinsically ontological and metaphysical dimension that one shouldn't ever forget. Here, we will emphasize another side of his physics, namely his theory of causality.

Aristotle distinguishes four types of causalities[27]: material, formal, efficient and final and, "since there are four causes, the physicist has to study all of them and,

---

[25] Ruggiu [23].

[26] Ruggiu [23].

[27] Aristotle, Physics [194b16–195b30, 198a14–b9].

considering all of them together, has to search, as a physicist, for the "reasons", that is, matter, form, movement, finality.[28] The distinction between material and formal goes back to a key distinction in his work, namely the distinction between matter and form. He always insists that the subject matter of physics in neither matter nor form, but the interplay between the two. Form abstracted from matter is instead the subject matter of mathematics.

Formal causality refers therefore to ideas, structures, platonism, mathematics. Material causality would instead be typical of Ionian philosophy, the phenomenological understanding of matter (fire, water, earth …). Efficient causality relates to the movement and likely also to earlier philosophers like Anaxagoras. Final causality refers at last to finality –the kind of causality that would be typical of early attempts to understand, for example, living systems.

3.3 Aristotle's Causes in Modern Science

Excepted for formal causality, these ideas seem very far away from modern science. Aristotle's work hints at the fact that we should maybe have broader views on science, its philosophy and its goals. This point of view has been defended recently by Francis Bailly and Giuseppe Longo in their book, *Mathématiques et sciences de la nature*. They refer explicitly to Aristotle and his theory of causality, some of their analysis echoing our previous developments:

> Physics and biology, in contrast to very abstract paradigms still dominating in the foundations of mathematics, are constituted respectively around the concepts of matter and life, seemingly so concrete although they cannot be defined internally in these disciplines. They also present the difficulty of appealing all the time and essentially to the requirements of rational coherence, largely mathematized in physics.[29]

They discuss then explicitely how some of Aristotle's causalities could translate in the framework of classical quantum mechanics.

We prefer to consider here the more fundamental framework of QFT. A striking but seemingly unnoticed fact is indeed that many textbooks of QFT follow spontaneously a pattern that fits largely an aristotelician-type analysis of foundations (although most likely without any intention of the authors to follow such a pattern).

Formal causality relates to the mathematical consequences of fundamental principles of invariance (or symmetry), which translate into physical principles and physical quantities.[30] Time translation invariance of the theory leads to the conservation of energy. Space translation invariance leads to conservation of momentum. Rotational invariance to the conservation of angular momentum.

---

[28] Op. cit. [198a23].

[29] Bailly and Longo [8].

[30] The mathematical framework being Noether's principle: roughly stated, symmetries translate into conservation laws and conserved quantities.

Material causality is embodied instead in the definition of concrete theories and quantum fields. This amounts to specify the underlying "matter" and its properties. For example, photons, electrons and their interaction rules, in quantum electrodynamics together with the choice of the physical constant giving the strength of the interaction.

Efficient causality deals with movement and dynamics. The quantization of the evolution equations of classical theories leads to the Schrödinger equation and other equations that allow in the end to describe the free motion and the interaction of particles. The Feynman diagrams we discussed in the first section appear when expanding perturbatively their solutions.

These three moments of the construction of theories of quantum fields are of course intimately related and cannot be disentangled: it is actually classical in physics to use conservation laws to derive equations of motion. They do however correspond to three different moments of the analysis: symmetry principles; the definition of "objects" (particles as quantum fields); the study of transformations and motion.

It is an interesting exercise to analyze other fields through this filter of causality. In mathematical finance, the other field that we have chosen to illustrate epistemological problems related to complexity, the foundational model is Black-Scholes'. Once again, looking at the model different moments emerge in its constitution.

Formal causes arise from the mathematical translation of assumptions on the behaviour of financial markets, in particular the absence of riskless profits (no arbitrage opportunity principle or "no free lunch" under the hypothesis of efficient markets and perfect information).

Material causes include the existence of stocks, risk free assets, but also transaction rules (possibility of short selling), markets behaviour (liquidity of assets, transaction costs …). In a subtler way, they also include the financial analog of physical constants in QFT: for instance implied volatility for vanilla call and put options.

Efficient causality would refer instead to the dynamics of assets. The basic assumption here is the lognormal behaviour of stock prices, which is usually grounded theoretically on the central limit theorem and the idealized view of many independent agents cooperating to asset price formation.

4 Mathematical Atomism

Epicure's philosophy had already raised the problem of emergence of global patterns out of local interactions. This problem, central to atomists' philosophy, admits many variations, in various contexts.

4.1 Logical and Mathematical Atoms

Logical atomism is one of the best known. In many respects it is based on the same assumptions as classical atomism: the idea that there are elementary components (of thought, of sensation …) from which our knowledge of the world would be assembled. Logical atomism was conceived in the early 20th century by Russell[31] and Wittgenstein[32] and was very influential inside the Vienna Circle and for the edification of analytic philosophy. The opposition that developed among philosophers of science such as Cavaillès or Lautman against the Vienna circle was largely rooted in their anti-reductionist stance against this conception of science and knowledge.[33]

In philosophy, logical atomism *stricto sensu* was abandonned relatively rapidly, partially because of the criticisms of Wittgenstein himself who realized that language and therefore thinking is not the mere atomistic description of the structure of the world and that the formation of meaning obeys to much more complex rules. However, the first Wittgenstein is an heraldic figure and his initial views remain influential and underly a deep trend in logic and philosophy of language.

Another form of atomism is more relevant to discussions around reductionism, complex systems and mathematical practice, namely *mathematical atomism*. We use this name to denote the widespread temptation to think that the scientific description of phenomena using mathematical models can always be obtained as the sum or conjonction of atomic models, each taking in charge a particular feature of the problem. The problem of mathematical atomism is not so much the fact that the atomic models can be wrong, than the fact that the emergence of patterns out of their interactions often requires new ideas, new methods.

The 2007–2008 financial crisis[34] provides an illustration of the drawbacks of mathematical reductionism, entangled with other modelling and practical problems stemming from financial markets and finance.

The crisis made evident that complex financial products such as RMBS were much more difficult to price than expected. Practitioners did then value them using a standard approach as far as interest rates, inflation and similar financial quantities were involved, and Monte-Carlo simulations for the long-term behaviour of the other parameters. This approach was mixed to a mark-to-market one, parameters implied

---

[31] "Bertrand Russell (1872–1970) described his philosophy as a kind of "logical atomism", by which he meant to endorse both a metaphysical view and a certain methodology for doing philosophy […]. According to logical atomism, all truths are ultimately dependent upon a layer of atomic facts, which consist either of a simple particular exhibiting a quality, or multiple simple particulars standing in a relation. The methodological view recommends a process of analysis, whereby one attempts to define or reconstruct more complex notions or vocabularies in terms of simpler ones […]. Russell's logical atomism had a profound influence on analytic philosophy in the first half of the 20th century; indeed, it is arguable that the very name "analytic philosophy" derives from Russell's defense of the method of analysis". Klement [24]. See also Russell [25].

[32] Wittgenstein [26].

[33] See e.g. Castellana [27] and Benis Sinaceur [28].

[34] For an in-depth and technical analysis of the role of mathematical models in the context of the crisis, we refer to Brigo et al. [29].

from the existing prices of contracts being used to derive the value of the newly issued ones.

After the crisis, it became clear that the financial industry, whose behaviour had been driven by various reductionist paradigms, had overlooked the key ingredients of mortgage valuation. One key issue was the intrinsic contradiction between the short term views of mark-to-market methods (suited for traders whose aim is to optimize the value of their portfolios at high reporting frequencies) and the very long term behaviour of the underlying contracts (typically 20 years, or more). Concretely, the financial techniques used were based on a double transposition. First, the methods in use to create and manage vanilla derivatives on stocks (call, puts …) had been extended to the management of credit risk (bonds, corporate or sovereign loans …). This first step raised already serious problems as default risk (that is, the risk encoding defaults on the repayment of interests or notional on bonds, bankruptcies …) is of a quite different nature from the one of the risks embedded in the random evolution of stock prices.[35] Then, the same methods were extended further to pools of products with embedded credit risk such as mortgages, student loans, and so on. Still another layer of abstraction was under development (CDO squared, based on pools of pools of contracts) when the industry collapsed with the crisis.

The conflict we alluded to (between short term and long term views) was reflected in the discrepancy between two financial communities with different cultures, backgrounds and paradigms: say, derivative issuers and traders on one side; mortgage issuers and retail bankers on another. What the short term views failed to understand was:

- The key role of the housing market. As far as prices raised, borrowers in difficulty could resell their houses with a benefit and repay their mortgages. When the market fell, prices collapsed and liquidity dried.
- The difficulty of modelling the evolution of interest rates and inflation on the long term. Market-implied solutions (based on the traded forward values of rates) do account only for the present views of markets and not on a serious modelling of their long term dynamics.
- Various risks were embedded implicitely in RMBS, difficult to model and take into account: evolution of the labour market, possibility for the borrowers to renegociate their loans or repay them earlier …
- Lastly, the question of the quality of the loans, that could be addressed at "low" levels (as occurs in retail banking) but not at the "macro" level of large pools of mortgages.

Although these issues may seem to have little to do with mathematical reductionism, they do indeed. Most of the ingredients used in the overall pricing method were based on relatively sound and robust principles. For example, it is sound to link the interest rate served on a loan to the risk that the borrower will not be able to repay it and will default on the scheduled paiements. Using long term forward rates on sovereign

[35] On the pricing of credit derivatives, see e.g. Bielecki and Rutkowski [30].

bonds and other information available on bond markets makes sense for an insurance company managing dynamically a portfolio of pension funds.

On a behavioural side, the key mistake was to use these practices and models outside of the domain where they were born and had been conceived. There is certainly a tendency in human beings to use their existing knowledge as a proxy for knowledge still to be developed,[36] and to trust their beliefs outside the area where they can be considered as representing safely reality. On a purely mathematical side, the key mistake was the idea that standard stochastic models that could make sense separately for the various involved parameters would still hold when combined with each other: wrong ways risks resulted from the correlation between all the parameters. This phenomenon will be studied in detail in the last section of the article on a simple toy model.

4.2 The Principle of Reason

Mathematical atomism has also philosophical roots and a technical background. The starting point, in the modern area, could be a principle stated by Leibniz: the principle of reason -*principium rationis*, *der Satz vom Grund*. The German philosopher Martin Heidegger dedicated a long essay to the question,[37] and we will implicitely follow part of his analysis.

The principle of reason can be stated simply as "every effect has a cause". Of course, this may look like a tautology, but it is not when one looks at the true meaning of the sentence—and actually at all its possible meanings. We know from Aristotle that "cause" has several meanings, and as much can be expected from the principle of reason. For example, it can be interpreted as a cognitive principle: we should always be able to trace back a phenomenon to intelligible and rational principles and grounds. This is how the principle is often understood. One can go a step further and ask the description of causes to be mathematical—as it happens for example in physics. The problem is then that the mathematical model is often confused with the explanation of the phenomena, without a proper questioning of the limits or adequacy of the model. This phenomenon is worsened when mathematical models are used without an adequate training in mathematics, the latter helping to understand the ground of the underlying hypotheses.

The example of QFT is interesting from this point of view: we are able to describe very precisely particle physics with a complex mathematical apparatus. However, whether or not this means that we truly understand what a particle, what the world truly is, unclear: why has the theory this form? Why are the physical constants the ones we observe and not different ones? Why is the mathematical theory plagued with infinites that one has to remove through a complex process without a clear

[36] This tendency is closely related to the notion of paradigm in the work of Kuhn and more generally in the context of historical epistemology. On the latter, see Brenner [2].

[37] Heidegger [31].

physical meaning, called renormalization, to get sensible results? On another hand nobody can claim to understand what a particle is without learning first its definition in the context of classical non relativistic quantum mechanics (where plenty of ontological problems already arise); then the definition of quantum fields, of Feynman propagators for free particles; at a higher level, the role of ghost fields, and so on. There is no proper understanding without mathematical models, but these models should not terminate the quest for a proper understanding. In other words, formal, mathematical "causes" cannot be easily disentangled from the other ones –this is one of the interesting sides of the philosophy of complexity. Complexity, complex systems, are indeed particularly interesting from this point of view since, by their very nature, they raise the problem of relating effects with causes, explanations, models, in situations where these relations are all but evident.

4.3 Mathematics as a Dynamical System

Turning back to mathematics, Heidegger observed that the latine word "principium" and the German "Grund-Satz" do not say exactly the same thing and that the corresponding Greek word would be "axiom", with still another meaning. For Aristotle and till recently, axioms were essentially propositions that hold true because expressing an obvious content. Another important idea of Aristotle, put in action in Euclide's treatise, the *Elements*, was that mathematics (arithmetics and geometry at the time) can be obtained by looking at elementary objects (planes, lines, points …), construction rules, and their interactions governed by a small set of fundamental axioms and principles. This is probably the first example of highly sophisticated intellectual construction built on elementary components. The Euclidean model would lead to Hilbert's views on axiomatics and later to the Vienna circle, and Russell's and Wittgenstein's theses on logical atomism, with the consequences that we have briefly described on 20th century philosophy of science.

Euclide's *Elements*, due maybe to their beauty and deepness, have created indeed the illusion that mathematics can be generated from basic sets of axioms, whereas the underlying generation process is highly complex and can certainly not be accounted for by a mere invocation of axioms and logical principles. The structure of groups, with all its theoretical ramifications and applications, has for example little to do with the axioms of set theory on which it is supposed to be founded. Another way to state these ideas is that it is true that modern mathematics can be presented as meaningless symbols interacting through a limited number of rules -the axioms of set theory for example. But this presentation will never be able to account for the way they progress, for their meaning, and why they have so many applications. The view that axiomatics would be the right way to account for all the structure and meaning of mathematics has however been for long popular as, in biology, the idea that DNA would contain all the information on living, biological systems. The idea is still popular in some circles of mathematical philosophy, although loosing momentum in view of its scarse implications when it comes to analyse actual mathematics.

Here we suggest mathematics as a whole should be understood, at least metaphorically, as a very sophisticated dynamical system. Its growth, its evolution are governed by a mix of internal constraints; internal goals that emerge spontaneously from its progress; external motivations, like the ones coming from physics, chemistry, biology, economics, and lastly by esthetical requirements and metaphysical views. In other terms, mathematics deserve to be explored from the point of view of dynamics, complexity and emergence.

We do not pretend to develop this program in the context of the present article, but we will stress its possible meaningfulness from the point of view of the development of mathematics and philosophy of mathematics during the 20th century and the beginning of the 21st. A possible path would be to build the analysis of mathematical progress jointly on the two traditions of historical epistemology and philosophy of concepts on one side, the problematics and mathematical developments surrounding complex systems on another side. Following such a path would also go along with the current reappraisal of historical epistemology in the philosophy of sciences to which we have already alluded.

To explain why such as approach makes sense on theoretical grounds, we will largely follow here the account of the French philosophy of mathematics given by H. Benis Sinaceur.[38] According to an epistemological vein running from Brunschvicg and Bachelard to Cavaillès, Lautman and, more recently, Desanti, Granger or Vuillemin, mathematical concepts live and develop. The internal logic of the objects and theories governs the dynamics, but in an unpredictable way:

> In mathematics, links are made across a complex network scattered with concepts connected to each other by organic links of different kinds. This 'organism' is not stable. It evolves constantly under the influence of local changes, which have repercussions on the configuration of the whole. The development of the concept is more important than the concept itself. With mathematics we are dealing with a 'conceptual progression'. The concept lives, and develops.[39]

To describe the logic underlying this life and development, Cavaillès and others used the Hegelian term 'dialectic'.[40]

> The dialectic is a logic, but it is not a formal logic [...]. It expresses the, so to speak, substantial link between the necessity and the unpredictability of mathematical development.[41]

Sinaceur's analysis is focused on the history of epistemology but still conveys implicitely a strong thesis: we have to take into account the legacy of these theories when trying to understand contemporary science. The mathematics studied by Cavaillès and Lautman -and by later philosophers of the same tradition- are different from the present ones in many respects. However, their 'organic' conception of

---

[38] H. Benis Sinaceur, op. cit.

[39] H. Benis Sinaceur, op. cit.

[40] Interestingly, logical atomism was deliberately and explicitely devised by Russell against hegelianism. See his essay on the subject "The Philosophy of Logical Atomism" in his Collected Papers, vol. 8, op. cit.

[41] H. Benis Sinaceur, op. cit.

mathematical development is still meaningful: we should just adapt it to the new phenomenology of mathematical progress, discoveries and problems. Conversely, advances in mathematics, in the study of living systems as a whole, in random growth processes, can help to revisit the ideas of organic links between concepts, of organic growth, of a network of ideas and theories or of mixing necessity and unpredictability.

5 Pension Schemes as Complex Systems

Although they are not solved, the nature of the problems raised by the valuation of ABS is now relatively well understood. The funding of pensions raises very similar difficulties that have been for long underestimated due to a poor understanding of the joint dynamics of the involved parameters.[42] The aim of this section is to show how simple models allow to capture some of these phenomena.

5.1 Defined Benefit Pension Funds

Whereas in defined contributions plans the employer is only committed to serve determined contributions to the pension plans of its employees, in defined benefit plans[43] it is committed to abund a fund that will serve pensions. Their amount is determined in advance, hence the terminology "defined benefit". The fund is invested in assets, typically bonds and stocks. The employer, called the sponsor of the fund, carries the investment risks but can also benefit from surpluses of the fund. An important concern with defined benefit plans is their possible underfunding and its consequences on the survival of the sponsor and the future paiement of pensions.

Whether considering defined benefit pension funds (DBPF) or defined contributions ones, there is a long list of parameters that are interacting to contribute making any modelling of the system of pensions and any analysis of its long term reliability extremely difficult. One can quote: the evolution of interest rates and inflation; the evolution of stock markets and world growth; the ageing of populations and the calculation of future mortality rates, and so on.

We focus here on a single issue: the wrong way risk created by the investment of a DBPF in stocks. Many features of the problem are now well-known. First of all, DBPF management has for long relied on accounting rules and portfolio management practices that did not take into account the very particular life insurance-like features

[42] As mentioned earlier, the results in this section were announced in a working paper written by the two authors in 2007. We thank Lionel Martellini for pointing out to us the timeliness of these questions and publications addressing them, in particular Martellini and Milhau [32] and Inkmann et al. [33].

[43] Defined benefit plan assets amounted to 7.9 trillion in the U.S. at the end of 2013. Inkmann, Blake and Shi, op. cit.

of pension funds. The accounting rules were often based on high expected returns on the equity pension fund portfolio; this tended to give a fully inadequate picture of the level of funding.

These problems tend to be solved progressively, following the introduction of regulations, better rules of practice, modern accounting standards[44] and a better education of trustees (the managers of the fund, in the UK terminology) to the concepts, tools and methods of modern quantitative finance. However, the quantitative treatment of pension liabilities still remains a domain where the theory has not crystallized into a stable set of paradigms, as illustrated by the ongoing debates among practitioners and academics in all the domains involved: among others, accounting, corporate finance, insurance or regulation. Here, we will address a specific problem: how the asset allocation of a DBPF affects the borrowing capacities of the sponsor and its probability to go into bankruptcy. This is part of the general problem of understanding how corporate and DBPF management are entangled, and how this entanglement should be dealt with.[45]

We follow the line of researches inaugurated by R. Merton and his collaborators.[46] Merton pointed out that, knowing that the U.S. stock market incorporate shortfalls and surpluses of pension funds into its estimates of company value, the most important issue in the field is related to the risk induced by the very structure of the asset allocation of pension funds assets. Pension funds have debt-like liabilities and hold equity-like assets. However the corresponding impact of the risk profile of the firm is not taken into account as it should be, among others because of the accounting of the pension-related debt. Notice that, from the various stakeholders point of view, investing in a company with a DBPF is a leveraged bet on the equity market, most often without a clear view on the size of the leverage.[47]

On the technical side, our approach relies on the so-called structural methodology to assess the default risk of a company and on the corresponding credit risk tools linking the default probability of the firm and its funding costs with its capital structure.[48] The conclusions of our simulations may be summarized as follows: from a quantitative risk-management approach, the impact of the asset allocation of the pension fund, in particular the effect of the dependence of its assets value on the stock market is certainly much higher than what one would naively expect. It is doubtful that the risk induced for the sponsor by the fund investments into equity is correctly

---

[44] See for instance Chapman et al. [34] or Cowling et al. [35].

[45] Similar problems can be raised for government-funded plans, which are a burden on a state's finances that can hinder its growth, lead to an increased deficit and other similar consequences. The idea that a state cannot fail on its long terms commitments is largely an illusion.

[46] See for instance Jin et al. [36] or Merton [37].

[47] Industry managers and shareholders have progressively become aware of the problem. One could insist on reallocating the DBPF assets towards equity volatility immune assets but this practice can have a negative impact on the long term in a context of low returns on bonds. Deriving tools to decide what is the best possible asset allocation from a risk management point of view is actually one of the main problems.

[48] See T. Bielecki and M. Rutkowski, op. cit. and L. Martellini and V. Milhau, op. cit.

priced by the market. Also, corporate managers may not fully appreciate the extent of their exposure to the stock market.[49]

5.2 Wrong Way Risk

Let us start with a brief description of the problem featured, as well as of its quantitative implications, both for the sponsor (the firm) and for the beneficiaries of a given scheme. The debt D_t of the firm decomposes into two components: the debt associated with the DBPF accounted deficit Df_t, and another component, $Di_t := D_t - Df_t$ that we call the industrial debt. The pension deficit Df_t, in turn, relies on two components: the fund assets A_t, which are invested in bonds, equity and other financial instruments to which the fund liabilities L_t have to be subtracted, that is the discounted value at time t of the future pensions cash flows: $Df_t = L_t - A_t$. The equity component of the pension assets is usually large and makes the fund's deficit behavior share many features with the behavior of the market capitalization of a firm.[50]

Below, we address the effect on the risk profile of the firm of the correlation of the sponsor's corporate value with the equity component of the pension assets. The sponsor is more likely to default when the stock market and the economy behave poorly. However, in such a situation, and because of the equity allocation of the fund, the pension deficit will also deteriorate, enhancing mechanically the default probability of the sponsor. This kind of effect is usually referred to as wrong way risk; a company with pension assets highly invested into equity could be expected to have higher funding costs than a company with pension assets highly invested into bonds, everything else being equal.

5.3 A Merton Model for Defaults

Let us start with classical assumptions. Our research is conducted within the so-called structural approach, introduced by Merton, Black, Cox and others.[51] A corporate valuation interpretation of the structural model relies on the idea that a firm defaults when the value of its assets falls below the value of the debt. In this interpretation, the process V_t introduced below would therefore stand for the corporate value at t and K for the expected value of the debt at a given maturity T.[52]

---

[49] From a historical perspective, this phenomenon is illustrated by the spring 2005 credit crisis that followed the downgrades in the U.S. Automotive sector—a direct consequence of the depreciation of the U.S. pension funds assets after the 2001 stock market crisis.

[50] The observation was used by Merton to revisit the WACC computations in the presence of a DBPS.

[51] Black and Cox [38], Merton [39].

[52] Stated in this way, this is, of course, a too strong assumption for various well-documented reasons. For instance, there is a considerable uncertainty on corporate assets valuation; the safety covenants

Concretely, one introduces a lognormal process

$$dV_t = \mu V_t dt + \sigma V_t dB_t \tag{1}$$

and a threshold K. Here B_t is a standard Brownian motion. The firm defaults at time T if and only if $V_T \leq K$, where K can be computed from p_T, the default probability at time T of the sponsor[53]:

$$p_T := P(V_T \leq K). \tag{2}$$

In practice, it makes sense to assume that the threshold K of the structural model should, at first order, vary in proportion to the debt. This will be one of our assumptions, so that, for example, if the pension fund deficit accounts for 60% of the firm's liabilities, a 10% increase of the deficit would result into a 6% increase of the threshold value K. Random variations of the fund deficit related to the stock market volatility should therefore impact the default threshold K and ultimately the default probability and credit spread of the firm.

Since our goal is to study the variations of the default probability of the firm when the asset allocation of the DBPS varies, we will use as a benchmark the case when the fund is fully invested in bonds and will study how the firm's default probability changes when the asset allocation varies.

We first have:

$$V_T = V_0 \exp^{(\mu - \frac{\sigma^2}{2})T + \sigma B_T}, \tag{3}$$

and:

$$p_T = P(V_T \leq K) = P(B_T \leq \sigma^{-1}(\log(\frac{K}{V_0}) - (\mu - \frac{\sigma^2}{2})T))$$

$$= N(\sigma^{-1}(\log(\frac{K}{V_0}) - (\mu - \frac{\sigma^2}{2})T)),$$

where N stands for the cumulative Gaussian distribution. Solving for K gives the value of the threshold corresponding to the market implied default probability p_T:

$$K = V_0 \exp\left[(\mu - \frac{\sigma^2}{2})T + \sigma N^{-1}(p_T)\right].$$

---

triggering default are strongly related to the value of the debt, but bankruptcy rarely occurs as the mere effect of the asset values being less than the debt, since a firm would normally do various attempts to restructure its debt before such a phenomenon occurs, and so on.

[53] The quantity p_T can be computed from the corporate credit spread s (the spread of corporate bonds relative to the risk free rates), the expected recovery rate and the corresponding hazard rate h, where $p_T = 1 - \exp^{-hT}$. If credit spreads are not available, the hazard rate may be approximated from the knowledge of the company's rating (and the corresponding default probability). See T. Bielecki and M. Rutkowski, op. cit.

We assume, for simplicity, a constant risk free interest rate r. Since we do not want to enter into considerations that would be irrelevant for our purposes, such as the composition of the fund, the number of beneficiaries already receiving a pension, the refunding of the scheme by the sponsor, we assume that the scheme's liabilities L_t behave as a risk free asset.

We write α for the proportion at $t = 0$ of the fund's assets A_0 invested in stocks, $A_0^s = \alpha A_0$, and A_0^b for the fund's assets invested in bonds assumed to be risk free. Assuming that no rebalancing occurs we get:

$$L_t = \exp^{rt} L_0,$$

$$A_t^b = \exp^{rt} A_0^b = (1 - \alpha) \exp^{rt} A_0.$$

Assuming that the equity component A_s^t of the fund assets follows a lognormal process with drift μ' and volatility τ, we get:

$$A_t^s = \alpha A_0 \exp^{(\mu' - \frac{\tau^2}{2})t + \tau B_t'},$$

where B_t' is the time t value of a Brownian motion. We further assume that the two Brownian motions B_t and B_t' have correlation ρ: $dB_t dB_t' = \rho dt$. The total T value of the fund's assets is then given by:

$$A_T = A_0(\alpha \exp^{(\mu' - \frac{\tau^2}{2})T + \tau B_T'} + (1 - \alpha) \exp^{rT}).$$

When the fund assets are totally invested in bonds ($\alpha = 0$), we get as benchmark value of the fund deficit Df_T:

$$Df_T^{bench} = A_0 \exp^{rT} - L_0 \exp^{rT} = Df_0 \exp^{rT}.$$

In general, $D_T = Df_T + Di_T$ and Df_T depends of α:

$$Df_T(\alpha) = A_0(\alpha \exp^{(\mu' - \frac{\tau^2}{2})T + \tau B_T'} + (1 - \alpha) \exp^{rT}) - L_0 \exp^{rT}.$$

We finally assume that no refinancing takes place and that Di_T behaves deterministically: $Di_T = Di_0 \exp^{rt}$.

Recall that, according to our Merton-type assumptions, the threshold K behaves proportionally to D_T. We get finally:

$$K(\alpha) = K \frac{D_T(\alpha)}{D_T^{bench}}$$

where K is computed in the benchmark case ($K = K(0)$) and, for the default probability dependency on α,

$$p_T(\alpha) = P(V_T \leq K(\alpha)).$$

Our following numerical results are based on a Monte-Carlo solution of this equation.

5.4 Quantitative Results

5.4.1 First Scenario

Sponsor with strong fundamentals: $\mu = 8\%$, $\sigma = 15\%$, $\mu' = 8\%$, $\tau = 20\%$, $r = 3.5\%$, $h^{bench} = -\log(1 - p_T(0)) = 50bp$, $Di_0 = 150M\$$, $L_0 = 500M\$$, $A_0 = 400M\$$.

The following table expresses the dependency of the 1Y default probability on the correlation between the sponsor and the stock market and on the proportion of the fund's assets invested in bonds. The first (0,0) entry is the benchmark 1Y default probability obtained under the assumption that the fund assets are fully invested in risk free assets.

$\frac{\rho}{\alpha}$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
0.1	0.005	0.006	0.007	0.007	0.008	0.009	0.01	0.011	0.012	0.012	0.013
0.2	0.007	0.008	0.01	0.012	0.013	0.015	0.017	0.019	0.021	0.023	0.025
0.3	0.009	0.012	0.015	0.018	0.021	0.024	0.027	0.03	0.033	0.037	0.041
0.4	0.014	0.017	0.021	0.026	0.03	0.034	0.039	0.043	0.048	0.053	0.057
0.5	0.019	0.024	0.03	0.035	0.041	0.046	0.053	0.058	0.063	0.069	0.074
0.6	0.027	0.033	0.04	0.047	0.053	0.059	0.066	0.073	0.079	0.085	0.092
0.7	0.036	0.044	0.051	0.058	0.066	0.073	0.08	0.088	0.096	0.103	0.108
0.8	0.046	0.055	0.063	0.071	0.079	0.088	0.096	0.103	0.11	0.118	0.124
0.9	0.057	0.066	0.075	0.086	0.094	0.102	0.11	0.117	0.125	0.132	0.139
1	0.07	0.08	0.089	0.098	0.107	0.115	0.124	0.131	0.139	0.147	0.154

Numerical results show that the effect of the stock market volatility on the 1Y default probability can be significant: even without taking into account the sponsor/stock market correlation, the default probability moves from 50 bp in the benchmark hypothesis (fund assets fully invested into bonds) to 270 bp when 60% of the fund assets are invested into equity. Moreover, taking into account the sponsor/stock market correlation strongly enhances the default probability.

For example, under the assumption of a fund asset portfolio invested at 60% in equity, the effect of the correlation on the 1Y default probability is of around 60 bp if $\rho = 10\%$, 130 bp if $\rho = 20\%$, 200 bp if $\rho = 30\%$, and so on.

5.4.2 Second Scenario

Sponsor with weaker fundamentals and higher DBPS deficit under low returns and growth hypothesis: $\mu = 1\%, \sigma = 25\%, \mu' = 2\%, \tau = 20\%, r = 3.5\%, h^{bench} = 500$ bp, $Di_0 = 250$ M\$, $L_0 = 500$ M\$, $A_0 = 300$ M\$.

$\frac{\rho}{\alpha}$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	0.049	0.049	0.048	0.049	0.049	0.049	0.049	0.048	0.049	0.049	0.049
0.1	0.05	0.05	0.052	0.053	0.054	0.056	0.056	0.057	0.058	0.059	0.06
0.2	0.051	0.053	0.056	0.058	0.06	0.061	0.064	0.066	0.068	0.07	0.071
0.3	0.053	0.057	0.059	0.063	0.066	0.069	0.072	0.075	0.078	0.081	0.083
0.4	0.056	0.061	0.065	0.069	0.073	0.076	0.08	0.084	0.088	0.091	0.095
0.5	0.06	0.065	0.07	0.075	0.08	0.085	0.089	0.094	0.099	0.102	0.106
0.6	0.064	0.069	0.076	0.081	0.087	0.093	0.098	0.103	0.108	0.112	0.118
0.7	0.067	0.075	0.082	0.088	0.095	0.101	0.107	0.113	0.119	0.124	0.129
0.8	0.073	0.081	0.088	0.095	0.103	0.109	0.115	0.122	0.129	0.0.135	0.14
0.9	0.078	0.087	0.095	0.103	0.11	0.118	0.126	0.132	0.138	0.145	0.15
1	0.083	0.092	0.102	0.11	0.119	0.126	0.134	0.14	0.148	0.154	0.161

The results show, once again, a strong enhancement of the default probabilities. However, precisely because of the weaker fundamentals of the sponsor in the second scenario and the higher underfunding of the pension fund, its exposure to the equity component of the pension fund portfolio is lesser than the exposure of the sponsor in the first one. It follows that the wrong way risk due to the investments of the pension fund is relatively weaker in that situation.

5.5 Financial Conclusions

Investing in equity is a very tempting solution for DBPS managers, in view of the long term higher expected returns that can be achieved on the stock market. However, this strategy may have devastating effects on the sponsor, enhancing its credit spreads and, under bad market and/or idiosyncratic conditions, leading the company to bankruptcy. Our computations, undertaken under conservative modeling assumptions, show that the effect of the DBPS investment strategies may be much greater than one would probably naively expect. This is particularly the case if the correlation of the sponsor's corporate value to the equity market is high.

Corporate managers that want to cooperate with pension funds asset managers to achieve together determinate risk objectives (which should be in the interest of all the sponsor's stakeholders and may occur e.g. when discussing the refunding of the fund's deficit) may do so in two ways. The first one is reducing the exposure to the stock market by switching from equity to bonds in the DBPS portfolio. The second is to reduce the correlation of the sponsor to the equity component of the portfolio, a result that may be achieved by switching investments, for example to stocks with a different exposure to economic cycles than the sponsor, to other classes of assets or to foreign stock markets.

5.6 Epistemological Lessons

Pension funds and more generally pensions are an enlightening example of complex systems. Trying to decompose these systems into elementary components will most often fail to account for their joint dynamics on which the behaviour of the system is ultimately based. The case study we have chosen to develop, featuring the effect of correlation between a DBPF sponsor and the stock markets investments of the fund is only one among many phenomena that could be analyzed in relation to pensions. For instance, our computations indicate short term effects (1Y) of this correlation, whereas more important problems can be expected to arise on the long term, due to the structure of pensions payoffs. In practice, the problem of modelling pension schemes is not only theoretical: parameters such as correlation are extremely difficult to calibrate on existing data—even more when they are supposed to account for long term phenomena. These questions can typically not be solved by brute mathematical force and require a delicate blend of technical knowledge, experience and … cartesian good sense.

The fact that this kind of phenomena took so long to be identified seems very surprising in retrospect: the articles by Merton and collaborators on the problems raised by pension plans date from 2004–2006. Insurance companies are used to handle these questions and have been doing so for long using relatively robust accounting and actuarial methods. The shift that could be observed during the last 20 years is based on the replacement of these classical techniques by financial ones inspired by a mark-to-market philosophy backed by the use of mathematical models originating in derivatives trading and related areas. Analyzing this shift is not so easy. Some robustness has been lost by trading long term views on financial markets and economy by short term ones. Mark-to-market techniques increase the volatility of valuations and prices and enhance wrong risks effects: a bad economical environment and bearish stock market conditions will simultaneously deteriorate the valuation of a firm and increase the underfunding of its DB pension plan, leading potentially to feedback effects that can put the survival of the sponsor at risk. On another hand, another kind of robustness has been gained. For example, discounting future pensions payoffs using discretionary rates to compute their present value, as it had been done earlier, was most certainly putting the plans at risk.

At a technical level, the example we treated is based on the use of default probabilities and therefore, implicitely, credit spreads and hazard rates: notions whose mathematical theory (as we used it) has been developed relatively recently, in the context of credit derivatives. The understanding of the effects of correlation between assets has also experienced deep advances due to the problems raised by multiname credit derivatives such as ABS, CDOs, RMBS and the like. Being able to use now these techniques to analyze problems is certainly useful and leads to quantitative assesments that would be impossible otherwise.

Our conclusions are the therefore mitigated. Whereas mathematical atomism is certainly dangerous when used blindly, a good blend of it with global views may pave the way to a good modelling of the phenomena. Or, using a direct language, math-

ematics cannot be avoided, but cannot be relied on too much and blindly. Progress in finance and economics in domains such as the one of pensions funding has to go on two legs: creating models; criticizing them and fixing boundaries to their validity. By experience, this second leg tends to be amputated. To paraphrase Aristotle: This is ignorance not to be able to understand what a mathematical argument proves and what it does not.[54]

References

1. Lecourt, D.: L'épistémologie historique de Gaston Bachelard. Vrin, Paris (2002)
2. Brenner, A.: Quelle épistémologie historique? Kuhn, Feyerabend, Hacking et l'école bachelardienne. Revue de métaphysique et de morale **49**(1), 113–125 (2006)
3. Hacking, I.: Historical Ontology. Harvard University Press (2002)
4. Daston, L., Galison, P.: Objectivity. Zone Books (2007)
5. Rheinberger, H.-J.: On Historicizing Epistemology: An Essay. Stanford University Press (2010)
6. Voilquin, J.: Les Penseurs grecs avant Socrate: De Thalès de Milet à Prodicos. Garnier (1964)
7. Patras, F.: The Essence of Numbers. Lecture Notes in Mathematics. History of Mathematics Subseries. Springer, (2020)
8. Bailly, F., Longo, G.: Mathématiques et sciences de la nature. La singularité physique du vivant. Hermann (2006)
9. Gil, F.: Démocrite, Encyclopedia Universalis (1988)
10. Poincaré, H.: Calcul des probabilités. Gauthier-Villars (1912)
11. Reutenauer, C.: Free Lie Algebras (London Mathematical Society Monographs). Clarendon Press (1993)
12. Cartier, P., Patras, F.: Classical Hopf Algebras and Their Applications. Springer (2021)
13. Joni, S.A., Rota, G.C.: Coalgebras and bialgebras in combinatorics. Stud. Appl. Math. **61**(2), 93–139 (1979)
14. Patras, F.: Homothéties simpliciales. Ph.D. Thesis, Université Paris 7 (1992)
15. Patras, F.: La décomposition en poids des algèbres de Hopf. Ann. Inst. Fourier **43**(4), 1067–1087 (1993)
16. Patras, F.: L'algèbre des descentes d'une bigèbre graduée. J. Algebra **170**(2), 547–566 (1994)
17. Diaconis, P., Pang, C.A., Ram, A.: Hopf algebras and Markov chains: two examples and a theory. J. Algebraic Comb. **39**(3), 527–585 (2014)
18. Humphries, N.E., Sims, D.W.: Optimal foraging strategies: Lévy walks balance searching and patch exploitation under a very broad range of conditions. J. Theor. Biol. **358**, 179–193 (2014)
19. Dedekind, R.: Was sind und was sollen die Zahlen?, 1st edn. Auflage, Vieweg, Braunschweig (1888)
20. Benacerraf, P.: What numbers could not be. Philos. Rev. 47–73 (1965)
21. Husserl, E.: Die Krisis der europäischen Wissenschaften und die transzendentale Phnomenologie: Eine Einleitung in die phänomenologische Philosophie (The Crisis of European Sciences and Transcendental Phenomenology: An Introduction to Phenomenological Philosophy) (1936)
22. Brown, J.R.: How do Feynman diagrams work? Perspect. Sci. **26**(4), 423–442 (2018)
23. Ruggiu, L.: Introductive Essay to Aristotle's Physics in Aristotle, Physics. Rusconi (1995)
24. Klement, K.: Russell's Logical Atomism. The Stanford Encyclopedia of Philosophy. Spring (2020 Edition)

[54] Aristotle's original sentence in his Metaphysics Γ reads: This is a crude ignorance not to distinguish what requires a demonstration and what does not. He was probably arguing against Heraclitus, Democritus and their followers.

25. Russell, B.: Collected Papers of Volume 8, The Philosophy of Logical Atomism and Other Essays: 1914–1919. Slater, J.G. (ed.). Allen and Unwin, London (1986)
26. Wittgenstein, L.: Tractatus Logico-Philosophicus. Routledge and Kegan Paul Ltd., London (1922)
27. Castellana, M.: Alle origini della nuova epistemologia. Il Congrès Descartes del 1937. Il Protagora 17-18, Lecce (1992)
28. Benis Sinaceur, H.: From Kant to Hilbert: French philosophy of concepts in the beginning of the twentieth century. In: Ferreirós, J., Gray, J. (eds.) The Architecture of Modern Mathematics: Essays in History and Philosophy, pp. 349–376. Oxford University Press (2006)
29. Brigo, D., Bielecki, T., Patras, F. (eds.): Credit Risk Frontiers: Subprime Crisis, Pricing and Hedging, CVA, MBS, Ratings and Liquidity. Wiley-Bloomberg Press (2011)
30. Bielecki, T., Rutkowski, M.: Credit Risk: Modeling, Valuation and Hedging. Springer (2004)
31. Heidegger, M.: Le Principe de raison, trad. A. Préau. Gallimard, Paris (1962)
32. Martellini, L., Milhau, V.: Capital structure choices, pension fund allocation decisions and the rational pricing of liability streams J. Pension Econ. Finance 1–21
33. Inkmann, J., Blake, D., Shi, Z.: Managing financially distressed pension plans in the interest of beneficiaries. J. Risk Insur. **84**(2), 539–565 (2017)
34. Chapman, R.J., Gordon, T.J., Speed, C.A.: Pension, Funding and Risk. Institute of Actuaries and Faculty of Actuaries (2001)
35. Cowling, C.A., Gordon, T.J., Speed, C.A.: Funding Defined Benefit Pension Schemes. Institute of Actuaries and Faculty of Actuaries (2004)
36. Jin, L., Merton, R.C., Bodie, Z.: Do a firm's equity returns reflect the risk of its pension plan? J. Financ. Econ. **81**(1), 126 (2006)
37. Merton, R.C.: The Real Problem with Pensions. Harvard Business Review (2004)
38. Black, F., Cox, J.: Valuing corporate securities: some effects of bond indenture provisions. J. Finance **31**, 351367 (1976)
39. Merton, R.C.: On the pricing of corporate debt: the risk structure of interest rates. J. Finance **2**, 449–470 (1974)

From Complex Dynamics to the Architecture of the City

Ferdinando Semboloni

Abstract The paper presents a theory for the architecture of the city. The aim of the theory, based on the imitation of cities that grow spontaneously, is to propose a method for designing a city. In doing so, we apply the concept of pattern derived from Alexander's design method. Using the pattern method, we conceive the city as a tool to solve the problem of communication between its inhabitants. We then explain the basic center-area pattern as a practical way of solution. The hierarchical combination of these patterns results in the urban structure. The method is shown with an example of a city design made through a Netlogo computer platform. With the help of the computer, the road network is then designed and its characteristics are studied using graphical topological surveys that allow the definition of squares and urban axes, the main components of the urban space. Based on centrality measures, we establish the use of the land and the shape of urban blocks. The conclusions underline the need for a comprehensive approach to the conception of the city.

Keywords Town design · Complex systems · Zipf's law · Pattern language

1 Introduction

Understanding cities as a result of a process of self-organization is a recent achievement of urban science [5, 39]. The main sign of this process is the persistence of Zipf's law [54] in many aspects of urban life, starting with the distribution of the population in cities. The internal organization of cities has also been recognized as self-similar [17], according to the definition of a fractal object.

As large infrastructure projects are realized according to a plan, the spontaneous rapid expansion of cities and new knowledge emerging from the field of complexity raise the question of the role of urban planning and design. Indeed, it makes no sense, one might say, to plan what is already managed by a spontaneous process, even if planning, being a political process, is just another aspect of society's self-regulation.

F. Semboloni (✉)
Department of Architecture, University of Florence, Florence, Italy
e-mail: ferdinando.semboloni@unifi.it

S. Albeverio et al. (eds.), *Complexity and Emergence*, Springer Proceedings in Mathematics & Statistics 383, https://doi.org/10.1007/978-3-030-95703-2_6

Self-organization questions the existence of an external order. Therefore it can represent a challenge for architecture which, as the name suggests, aims to establish the principles for the construction. That echoes the dispute between creationists and evolutionists and is probably the reason why the science of complexity has received so little attention in the practice of urban planning and design.

Contemporary architecture is a global phenomenon and architects belong to different cultures. It is therefore difficult to give a general explanation of the current trends in a nutshell. At the origins of modern architecture, there was a want to provide people with a healthy, sunny home, well integrated with the natural context. But now, one of the features that immediately catches the eye is the quest for an architectural artifact to differentiate itself from the urban context, with strange shapes, heights, and materials. We are witnessing a plethora of languages, from deconstructivism to digital architecture. Their intention seems to amaze the user, as in the Baroque period, with the difference that in that period, however, the ways of surprising people were much more subtle and intellectual, while now they primarily rely on technology. As if these unusual, and surprising architectures were trying to talk to each other by screaming louder and louder without considering the urban fabric to which they belong and the other common buildings that surround them. That is probably one of the reactions of architecture to the complexity of the world.

There are other trends in architecture that don't appear so dramatically at odds with their surroundings, but rather, realize buildings to make people happy. Some architects use a classical language arising from the tradition of the European city they want to imitate [29]. Other consider architecture as an anthropological expression, emerging from the culture of people in search of quality and beauty. I am referring to Paul Oliver's work on vernacular architecture [35], or *architecture without architects*, according to Bernard Rudofsky [41]. In the research of Christopher Alexander [1] on elementary forms of construction, there is not a style, nor the need to impress people, but only an effort to make people happier to live in harmony in a place. My idea builds on a similar understanding: I propose to approach the design of the city by imitating the existing reality and replicating the urban characteristics resulting from the process of self-organization. This approach, that I will explain in the following, make it possible to develop new applications of the science of complexity whose role becomes that of finding the invariant characteristics of self-similar objects, to reproduce them in a project.

Considering the city as an object is a classical position of architecture [48]. The reference is to Alberti's famous statement that the city is a large house and the house a small city. This statement was used by Alberti to show that rooms in a house, like buildings in a city, have a specific function. But it is also used to mean that architecture may claim its right to control the entire urban form. This desire finds its expression in the ideal and Utopian city [19]. But a city is a complex object that emerges from the interaction of many entities. Stephen Marshall calls it a collective entity [31]. How can it be considered a unitary object? That is the fundamental problem that I will try to address with imitation.

The text presents the theory for the design of a city and is structured as follows. After a brief review of the main theories on urban systems, the dichotomy between

designed city and spontaneous development is presented. Building on it, we propose an urban design that imitates the characteristics of the spontaneously growing city. Then we analyze the architectural knowledge based on the transformation of objects, and the related concept of pattern that includes problem and solution, to move on to the concept of imitation of invariant time-independent characteristics. By using the pattern method, we conceive the city as a tool to solve the problem of communication among its inhabitants. We then explain the basic center-area pattern as a practical way of solution. The hierarchical combination of these patterns results in the urban structure. The method is shown with an example of a city design realized through a computer Netlogo platform. With the support of the computer, we then design the road network and analyze its features by means of graph topological surveys that allow defining squares and urban axes, the main components of urban space. Based on centrality measures, the use of land and finally the shape of the built urban blocks are established.

2 Basic Classic Theories on Urban Systems

Since the idea is to mimic the essential characteristics of cities, let's recap some main notions of theories about their organization. One of the main concepts behind the explanation of cities is that of a market area: a city is a market place where people gather to sell or buy goods.

Let us consider the case of an agricultural firm which sells goods with different transportation cost. The farther the land is from the market, the higher the transportation cost and the lower is the rent offered for the use of the land. The farmer who bids the highest wins the auction for the land and cultivates it. These are the basic assumptions of von Thunen's theory for explaining the decrease in land rent (1824), a revised form of Ricardo's theory of marginal rent. From this theory, it follows the distribution in concentric circles of agricultural production around the urban center, depending on the difficulty of transporting the product. The theory was applied, one century and a half later, by Alonso [3] to the internal organization of the city, based on land rent.

Customers too, have to pay transportation costs to buy goods in the market. The price of the good therefore increases as the distance from the market growths. This assumption underlies the market area concept which is the core of the Central place theory of Christaller (1933) [12]. The basic hypotheses are (1) goods with different purchase prices and frequencies, (2) a maximum distance that the customers are willing to travel to buy the good at the market price, and (3) a minimum number of customers below which, due to fixed costs, it is not convenient to sell or produce the good. Under these hypotheses, the theory demonstrates that a hierarchically organized network of centers is formed with hexagonal market areas of extension depending on the frequency of purchase of the good and that these market areas overlap.

Only later in 1949, searching for the macroscopic laws of phenomena consisting of a multitude of events, Zipf, a quantitative linguist, formulated his law on the

distribution of population among cities [54]. Zipf's law for cities, known as the Rank-size-rule, is an empirical law that describes the population of a city belonging to a geographical area, as inversely proportional to its position (called rank) in descending ordering of cities by population size. This law is essentially a power law, similar to that formulated by Pareto about the distribution of income among individuals. It provides a rule about the distribution of population between cities which is the macroscopic result of complex dynamics.

This brief review would not be complete without mentioning Thomas Schelling's contribution to segregation [43]. Schelling starts from a simple assumption concerning the tolerance of an individual to live next to another individual different in ethnic characteristics or wealth. Applying this rule many times and for many individuals, the emergence of homogeneous areas is observed. The theory of segregation integrates the results of the bid rent theory proposed by Alonso-von Thunen because being close to similar people may result in economic advantages.

The theory of segregation is dynamically formulated using the complex systems method. Others define the final state without saying how to get there. For this reason, there have been studies to show how a dynamic system could reach the state of equilibrium hypothesized by the theories, especially for Central place, and Rank-size-rule. The Zipf's law, which seems to be the most mysterious, due to its validity for many phenomena, has had dynamic explanations from various viewpoints and empirical studies that have confirmed it for urban systems. Similarly, the other theories have also been confirmed by empirical studies.

Zipf's law remains the most important reference for this study since it is the only one that describes the macroscopic and invariant characteristics of many urban phenomena. But it says nothing about the organization of urban space. To this end, the Central places theory is much more useful. Below we will use these two theories and others as a guide to city design. So that we can move on to the main topic: design and self-organization in cities.

3 Cities: Design and Self-organization

Cities are strange objects. Observed from above, most of them appear as a spontaneous artifact, like a settlement of termites, grown naturally, with subsequent additions, more or less integrated with the previous ones. That is quite obvious when one thinks of the many human generations who have collaborated in the construction (or destruction) of a city that has existed for two thousand years, let's say more, during which the city could be in decline due to external adversities. But looking closely, every object, or part of the city, appear as realized according to some project or model. Even a hut is built with a plan. Thus cities seem to be the result of spontaneous development when viewed from afar and designed when viewed from a close distance, or of the superimposition of continuous, local growth processes and punctual planned changes [4].

We then start from these two opposite but also complementary [28] points of view for the construction of a city. The first: the designed city, a top-down approach and the second: the self-organization paradigm in which the city, according to the theory of complex systems, emerges from repeated interactions. Both have positive aspects: the design approach emphasizes the application of a unitary principle, while the complex system approach deepens the study of the functioning and evolution of the city.

4 The City Designed

The design and construction of an entire city is a rare event because the decision to create a new city is mainly a political one. The city usually hosts the administration of the surrounding region. The foundation of a new town can be decided by a higher level authority that intends to reorganize the territory under its jurisdiction, or by a new community that invades a region on which wants to establish its new authority. It is also connected to the colonization of territory, as in Spanish South America, or to the discovery of new regions as in North America or Australia. It is also worth remembering the experience of the sacred cities in India built with the method of Vaastu Shastra, the architectural treatises of medieval India [47]. But the construction of a new town is also motivated by the control of territory for defense, and for this reason, military towns are the most widespread example. The designed city has a simple and repetitive structure which does not change as the growth of the urban structure is not allowed. Let's consider these two aspects.

The most widespread military colonization experiment in the Western world can be considered that of the Roman Empire which built around 170 new cities, including London, Paris and of course Rome. Those cities were founded according to the rules for setting up military camps, including organizing the surrounding countryside based on similar principles (centuriation). These were so efficient that they were replicated in the United States with the Jefferson Grid in 1785.

Designing a city includes the effort to declare what the elements of a city are and how they must be combined to achieve the desired result: a functioning settlement. One of the most informative sources is the method of building a Roman camp. The elements of the Roman camps are essentially three: the walls, the center, and the roads between the two. Let's analyze them.

Walls are a fundamental element of a city. Drawing the walls is one of the first steps for founding the city. According to legend, when Romulus decided to build Rome, he began by drawing the line of the walls that separate the city from the rest of the world. The city is the inside, the rest is the outside. Although this act might have been initially motivated by defensive reasons, the idea of separation is recurrent in the planned cities. For example, in the garden city [23] there are no walls, but the village has clear boundaries with the countryside. Along walls, there are the gates which connect the city with the outside.

The second main element is the center. While walls are like a logical operator that decides whether something is in or out, the center is like a fuzzy logical operator; the relationship to the center is measured by distance and an individual can be near, far or very far from the center. The center, where the consul's tent was set up, identifies the power in the Roman castrum.

Between the center and walls, there is the roads network. It allows communication between each point and the center, and the accessibility of each point from any other point. In the Roman camp, there is a hierarchy in the road network. The main roads, cardo and decumano, connect the central square where they cross, with the four doors in the walls, while the smaller streets provide access to any point of the camp.

The street layout also allows the regular location of the buildings. That is the main reason for the town's gridiron structure. The gridiron is, in fact, efficient in two respects: easy location of each element thanks to only two measurable distances and modular organization of buildings even of different sizes. The alternative radial structure has the advantage of keeping the distance from the boundary to the center invariant; it allows efficient communication with the center, but it is inefficient for buildings that need to be adapted to the variable shape of the blocks.

The Roman camp plan was therefore inspired by order as mandated by military requirements and thus responded to the vision of a society based on a rigid hierarchy, as opposed to the organic nature of spontaneous development. Similar commitments towards social order were also in the projects of Utopian cities, since the earliest Filarete's Sforzinda. But what unites these experiences, even more, is the prefiguration of a static organism that should not have to be grown. In the case of the Roman camp, made up of men only, this was logical. But even later, the ideal city is considered as a place without expansion which must rather occur in satellite towns.

The ideal city and the fortification models (Fig. 1) represented for a long period a guide for the design of cities. In some sense, these cities are the result of a similar approach: that of special places detached from the normal way of life but responding

Fig. 1 Palmanova a Venetian designed town

to special principles that are order and discipline in case of a military city, and ethical principles in case of the ideal city. The design of a new town refers to the project of a different society or the affirmation of a new political power in society.[1]

The most widespread achievements of the planned city in recent times have their origin in Howard's idea of a garden city. It inspired the construction of New towns designed to decrease the density of metropolitan areas and generate a more balanced regional development, especially after the World War II. On this occasion, manuals were produced for designing whole urban arrangements [27]. The method involved a radial organization of the streets, an urban center with commerce, offices and services, secondary centers all around, alternating with green areas and an industrial zone. An organization which is similar to that of Palmanova (Fig. 1). Unless buildings that were not aligned with the road layout, as established by the modern movement in architecture. After the end of the building of New towns in the Western world, this phenomenon has found new life with the recent urban expansions in China [8, 50].

Aside from the new town's experience, during the period of urban spatial expansion caused by rapid population growth, the city project consisted of the addition of new buildings to the existing urban fabric. The aim was to solve the most pressing problems as social housing. Urban design was handled by urban planning and architecture [14]. Urban planning aimed to regulate growth according to norms and standards. While architectural projects proposed large buildings ideally containing a whole city. Both these strands have come together in an attempt to generate an easy-to-build modular city. That did not result in the production of beauty and more lively urban space as had happened in previous centuries when rich people were also interested in producing beautiful things to make a profit, see for example Venice, but in a mere increase of the number of buildings. Therefore the growth machine [33] produces the city through a self-organized development, even if it is formally planned.

5 The Spontaneous Self-organizing City

Discussions about designed cities seem unimportant compared with the extent of new settlements consisting of informal neighbourhoods, or widespread sprawl in the surrounding area. Research on complex systems as well as the awareness of the negative impacts of urban sprawl have stimulated the study of cities as complex systems in spontaneous evolution (Fig. 2). Urban evolution is the result of the interaction of a multitude of elements, while the conventional design approach considers a single organizational principle.

The theory of complex systems applied to cities has produced many insights that explain both microscopic urban dynamics and the macroscopic characteristics of urban systems such as the power law distribution of the size of cities. Moreover, the study of complex systems and self-organization deserve credits for emphasizing the spontaneous aspects of social organization and criticizing policies based on the idea

[1] Figure 1 is taken from GoogleMaps.

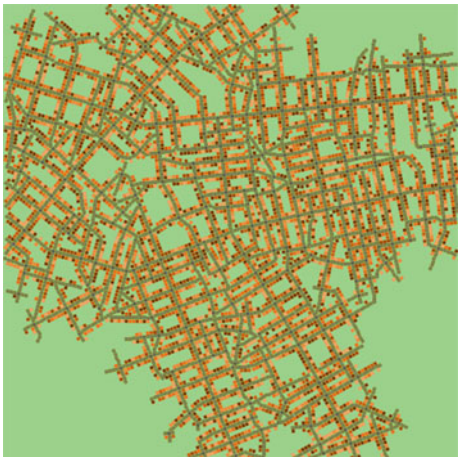


Fig. 2 The self-organised development of a city simulated by the computer

that central control can achieve better results than lack of control and spontaneous organization.

The theory of complex systems has been applied to the prediction of the effect of an urban plan with the idea of better evaluating alternative solutions, see for instance [53]. These are major advances in urban science, that I have been involved in. But the way to define the alternatives is left to the creative action, to the mediation between different interests, or to the participation of citizens who are asked to suggest solutions. Furthermore, the problem of uncertainty remains, which makes prediction difficult for complex systems. They are chaotic by definition, that is, strongly dependent on their initial conditions so that prediction becomes almost impossible.

Urban transformations originate more from a political will to change than from the prediction of some future state. We then return to the idea that only through an act of social will it is possible to define the future physical form and obtain configurations more suited to the wished goals.

6 The Artificial Self-organized City

So what is the purpose of studying the complexity of these systems? Do we have to throw away our studies just because the predictions don't hold? We need a complexity-based approach since any urban project, once completed, becomes part of the dynamic system. The buildings created by a rationality that starts from abstract principles resist for a short time the wear and tear of dynamics, especially that of complex systems, if these principles are unrelated to reality, or even in opposition to it. A clear example is the experimentation of modern architecture, especially in the field of social housing, sometimes demolished because it was counterproductive for the established goals.

The Utopian approach is rooted in the idea that actions and therefore decision-making must counteract the current lifestyle and social order. This approach is often responsible for creating problems in decision making because it sets goals that are difficult to achieve and in opposition to reality. On the contrary, I propose an approach to urban design based on the imitation of the existing world. This is not a replica, but the use of the macroscopic features of reality to produce a new part modified by its major flaws. Even the slums could inspire the cities of the future [24].

But how to imitate a self-organizing reality and why? If the world is self-organizing, it is not necessary to make an effort to create it. Like a tree: just dig a hole to put a seed in it and wait for the tree to grow with its characteristic fractal shape. Indeed, from a macroscopic point of view, the social system is self-organized, while if we focus on details we realize that there are many actions planned and intended as a voluntary act to generate something that depends on someone's will. If we accept imitation, the problem is not how to generate a world that depends on a single idea or, on the contrary, let all reality be self-generated, but how to learn from complexity to improve our decision-making and our ability to transform reality [45].

The problem we are facing is urban design as an example of the science of artificial [46]. But with an additional difficulty: it involves designing objects that would otherwise be produced spontaneously. There are many examples of this human activity, from artificial intelligence to city design. The reasons humans want to replicate are different, but the problem they face is similar: building an object that has to compete with similar objects in a dynamic world.

Designing an artificial city is also a way to study, or to understand the urban system from a design point of view. This is what Tony Garnier did with his "Cité Industrielle" [20]. Other architects have also proposed new city structures to reproduce urban features or represent extreme aspects of urban life. Such as Frank Lloyd Wright who conceived both the Illinois Sky City, the mile-high skyscraper, and Broadacre City, the low-density suburb with an acre of land for each family. We will discuss these aspects in the next section.

7 The Scientific and Architectonic Way of Knowledge

Architecture traditionally deals with the construction of artificial objects, and to do that it has developed a knowledge suited to the purpose. This type of knowledge differs from that of the natural sciences which traditionally do not intend to change what they observe, but only better understand the dynamics and then use the laws they discover in human activities. Time is therefore one of the main dimensions in scientific analysis. The designer, on the other hand, is interested in creating an object that satisfies the expected requests of a future user. That is why he/she is not very interested in dynamics, but rather in discovering the basic elements to compose a project which, once completed, will modify the world, adding a new object to those already existing.

This method resembles the radical constructivism in epistemology, according to which knowledge of the world is a human and social construction [49]. However, while constructivists think the world is independent of human minds, architecture sees the world as a product of the architect's activity. The knowledge of architecture consists of exploring the possible future state of the world. When an architect looks at an object, he/she is interested in the different combination of the single elements to create a new object, for example, a building, or a garden that will change the state of the place. This change can develop in many directions because there is no single solution to an architectural problem. Therefore, the knowledge of architecture can be applied to produce many solutions that can be considered as the possible future states of the object.

The set of these solutions, along with the current state of the object, represent the architectural knowledge. An architecture competition that results in so many different solutions is like collective learning of the problem. Even the Utopian project as the city of 3 million inhabitants of Le Corbusier can be considered not only as a project of something that could be realized but also as a mean to understand, by contrast, the contemporary urbanization model. In other words, as one of the alternative states, albeit very different from what a city could assume in its development. From this point of view, an Utopian project could also be used and partially implemented, as in the case of Tony Garnier's industrial city, which he used as a model for some projects carried out in Lyon.

8 From Object to Patterns

The need for an object arises from a problem, of which this object represents the solution. Whenever the same problem occurs, the same object will be the solution. But the problem will never be the same again and neither will the object. Therefore knowledge must consist of problems and solutions expressed in general terms that coexist together in architectural knowledge. We call "pattern" a re-usable form of a solution to a design problem, according to the definition of Christopher Alexander [2].

A table, for example, results from two basic patterns: the top from the pattern of having a flat surface on which to do the operations and the legs from the pattern of supporting it horizontally at a certain height. Many objects are not as simple as tables, but they have their internal dynamics, even if partly designed. Such as cities which are objects created by many individual actions. In this case, the problem is to identify the patterns on which people's collective knowledge is rooted, the use of which makes it possible to design the entire object.

However, we are aware that in the case of a complex object such as a city, the relationship between problem and solution is not linear but dialectical. A problem arises from a solution that already exists. The solution exists because it was there before. In other words, the evolution of a city is path-dependent. Natural conditions stimulate the creation of a settlement. The realized infrastructure becomes second nature to further development [30]. The design approach reduces complexity to a linear relationship, to imitate and use the solutions resulting from complex dynamics.

9 Imitation and Time

Studying complex systems can help understand how from collective intelligence solutions to problems emerge. This way we can create a book of solutions to imitate. For example, biomimicry, which mimics organisms or ecosystems, is an expanding field of research [37]. Through imitation, which considers the complex dynamics of these systems as a source of knowledge, this study constructs the rationality that emerges from reality and therefore is no longer in opposition to it.

The process of imitation has existed for a long time in art, but to produce an additional art piece it is necessary to imitate by inventing something new, interpreting the laws of classical art as in the Renaissance. Imitation is not making a copy but finding a synthesis of the main characteristics of objects belonging to the same set.

Furthermore, imitation does not create a theatrical environment for the city, however entertaining it may be, such as neo-vernacular outlets or urban procedural models [51] which tend to automatically reproduce many urban configurations especially as a game setting for the computer. The research is interesting, but the result depends on the goal of creating realistic-looking scenarios for games. They are also fully automated, while architectural knowledge and design are craftsmanship's activities that can use a computer but still rely on human intelligence to make strategic decisions.

The process is similar to the imitation of nature in Aristotle's sense. But what we consider nature is the spontaneous and self-organized processes of the cities. Our nature is the many examples of cities around the world, including all historical cities, usually formed by a spontaneous process, of which we imitate stable macroscopic characteristics [11]. The more the project resembles the macroscopic characteristics of the context, the greater its probability of long-lasting survival, successfully integrated with the existing reality.

Imitation implies an evaluation of the system to be reproduced. The structure of the city is a consequence of the economic system. It may be unacceptable in some of its aspects. Imitation does not imply the uncritical acceptance of everything that exists. We imitate aspects that work well and reject or modify what does not work. The criterion is not that of a novelty for which new is better, but that of functionality for the maximum number of people, taking into account those who are disadvantaged.

According to the proposed design approach, a project is an improved reproduction of what spontaneous dynamics is capable of achieving, with a difference. Observing the world as a set of systems emphasizes their dynamics which are understood through their changes, using time as the reference variable of the various states. In turn, knowing the world for changing it, means identifying the elements that can lead to building a different world. But this knowledge must abstract from time, considering the timeless elements. While the reality is continuously changing, the project ignores the dynamics, because it must propose a new stable state to be realized. The problem arises after the implementation as the dynamic, expelled, becomes the context in which the project, once completed, must live and exist over time.

The relationship with time can be better understood by remembering the method for maximizing entropy. The maximization of entropy in statistical physics results in the exponential distribution obtained with the maximum number of alternative

states of the particles. This distribution is stable and therefore time-independent, and could become one of the system requirements to imitate. Statistical physics does not consider alternative states in their individuality, but by the general macroscopic statistical law, they produce. The process of imitation must follow the reverse path, that is from generality to individuality, to propose a concrete alternative to be implemented, choosing one which respects the general law, in that case, the exponential distribution.

Once the invariant elements have been identified, we realize that these can be at various scales. A city can be made up of neighborhoods and these in turn of blocks and therefore of buildings. Spontaneous processes bring out general aspects from microscopic elements, such as individual buildings. But when it comes to designing a city in a unified way, it is necessary to start from the general organization and then go into detail. It simplifies a comprehensive approach to designing the object that might otherwise consist of additional parts. This is why we start establishing the general pattern to which the city belongs, answering the crucial question: for what problem is the city the solution?

10 Communication Is the Main Problem that a City Faces and Solves

There are many ways to understand how and why a city works:

- as an organism with its metabolism: materials that enter are processed and leave;
- as a safe area where their inhabitants can live and work;
- as a market area that serves the surrounding region;
- as a flow of people, goods and information.

These explanations refer to the city as a pattern, offering the solution to some problem. Choosing the problem means giving an idea of the essence of the city. The identification of the problem cannot be univocal, since it expresses a point of view. Establishing the key issue is a political act as it is a way of seeing the city that also includes a vision of the future state of the city.

Movement, not stillness, is the normal state of individuals. The movement also produces interaction and therefore communication which is the essence of a social system. On the other hand, communication requires movement. The problem is twofold: communication between many individuals gathered and living in one place and communication with the outside of the place. The city is the solution to both problems.

The city becomes a communication system [32], an urban web [42], that connects individuals there living. We get the result by distinguishing the space of flows and communication from that of living. Flows run in a continuous space extended throughout the entire city, from which the living spaces, the blocks, emerge like many small islands. Moreover, the city organization includes business or commercial centers that facilitate and stimulate a concentration of flows. In so doing, the city reduces the randomness of communication events and helps contact within the city [34].

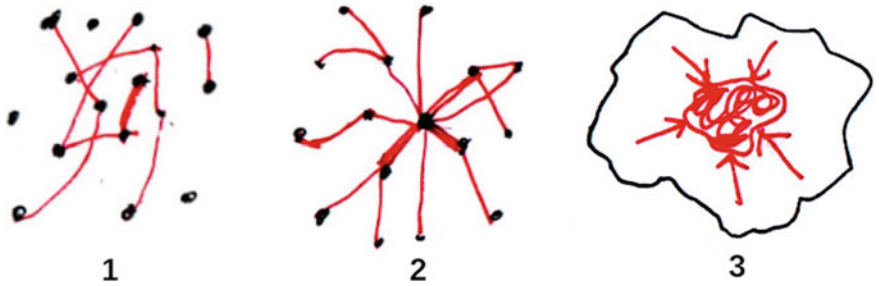


Fig. 3 1: Random communication. 2: Hierarchically structured communication. Few with many and many with few. 3: Pattern center-area that allows many to interact with a few

On the other hand, the city provides one or more gateways that easily connect the population with the outside world. In their simplest form, these gateways are the city gates open along the walls. They can take a more complex structure when moving from one transport system to another as in the case of a port. Large cities usually have a port, because they are located on the sea or are connected to the sea by canals, rivers and lakes. Railway stations and airports are the modern evolutions of the port. The gateways end up concentrating the flows that enter and leave the city in a few privileged points that become urban centers. The public space and the central areas thus become the two main ideas of the city's response to the need for communication.

11 From Communication to the Basic Center-Area Pattern

Communication leads to flows between individuals which can be messages or, in the case of the city, physical interactions that move people and goods. While there may be an interaction between each pair of inhabitants, this is unlikely. Some individuals become the reference for all others: few individuals have many relationships while many have few. An efficient spatial organization of this distribution of relations consists of placing the attractor of flows at the center (Fig. 3). That minimizes the total number of movements. The problem is similar to that of selling goods. In that case, the solution is to put the market at the center and the customers around.

Therefore the center-area pattern is considered the basic pattern for the organization of the city. It is a pattern in Alexander's sense: that is, it contains both the problem, that is, of communicating effectively and the solution, the center-area relationship. The center-area relationship is similar to the point-area relationship described by Aldo Rossi [40]: a monument is located in one point while the surrounding area is filled with ordinary housing. This design produces the primordial geometric figure, that is, the circle that was also the base of the first huts, with the hearth in the center.

This pattern is widespread everywhere. Consider eukaryotic cells with its nucleus and the mitochondria around it, an egg or an avocado. There are also circular cities, and arguably the most famous is the ancient city of Baghdad, with the mosque and

the Caliph's palace in the center. Other geographical concepts go back to the idea of the cell: the core-periphery model of the New Economic Geography and the market area where the market is the fulcrum in the center of the zone where people live.

The center-area pattern is similar to the market area: activities are in the central part and residences in the periphery, without a clear boundary between the two. We can therefore consider the center-area relationship as a workplace-residence connection. People live in this system usually commuting between these two places not necessarily belonging to the same pattern. Commuting flows can occur both within a pattern and between patterns.

A workplace can be, for example, a shopping mall or a public service such as schools. In this sense, the center-area pattern is similar to the neighborhood unit used as a building block in forming the city structure. The idea was introduced by Clarence Perry [38], in the 1920s and applied later to the design of English New Towns [27]. The difference with the notion of the center-area pattern is this. The neighborhood unit is a modular concept. Each unit includes a community center with an elementary school and has a similar size because the minimum population should be able to support the school. In addition, the maximum distance from the center should be a distance that a child can walk, one-quarter mile. In turn, as later shown, the size of center-area patterns differ from each other, mimicking the distribution of the cities size in a regional context, the Zipf's law.

12 The Channeling of Movement and the Establishment of a Center

Movement on earth usually occurs through a canal, i.e. a road or something similar. Only the sea allows movement in any direction without an established channel. That is why many large cities are on the seaside or communicate with the sea through rivers or canals. When land communications were harder than now, the sea was the mean of communication and the small sea allowed for easy communication and thus the development of trade and political control. That was the case of the Mediterranean Sea around which the Roman Empire developed.

The Romans were extraordinary road builders when they had to expand the empire towards the inland areas. But the roads also develop through self-construction processes, when repeated movements lead individuals to follow the same path (Fig. 4). In the periods technology was less developed, the privileged routes were those on the ridges of the mountains since the crossing of rivers was avoided. Bridges were rare, and it is not by chance that their existence was often the origin of a city born nearby.

The bridge is the quintessence of the road since only at that point the flows of people and freights can cross the flow of water. The main function of a transport network is to channel the flows, establishing the areas through which the flows can pass through, and the areas that cannot be crossed by the flows [22]. On the one hand,

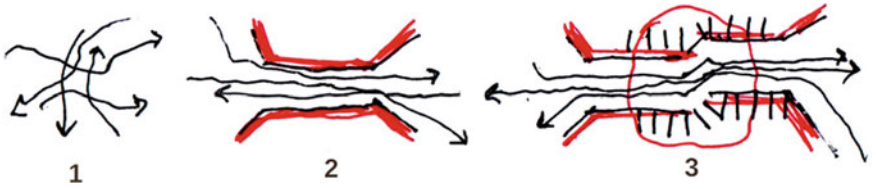


Fig. 4 1: Random movements. 2: Channeling of movements: street. 3: Breakpoint of the canal and beginning of the establishment of the center

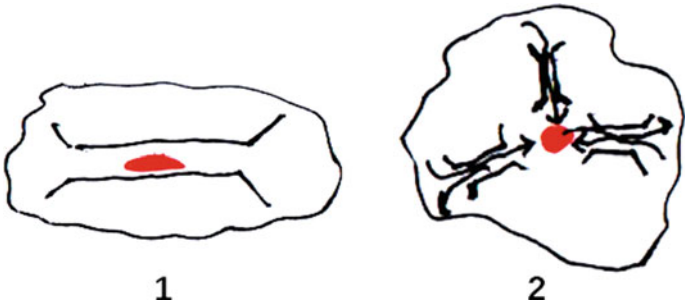


Fig. 5 1: An embryonic center is formed along the way. 2: The streets channel the flows between the center and the area

interaction decreases in randomness, and on the other, it becomes more efficient as roads and other communication networks accelerate the movement.

A center grows when there is a breakpoint (Fig. 4) in the channel, or it forks in two or more directions. It favors a slowing down of movement and therefore the meeting between moving subjects. A first central embryo begins inside the canal, as an elementary center. These embryos are brought into the central area pattern to give rise to the more elaborate form of the pattern which includes a center, an area and the channels (roads) between the two (Fig. 5). This structure is similar to that of the Roman military camp, apart from its rigidity. In addition, it resembles the organization of the eukaryotic cell with energy-carrying mitochondria. Anyway, the channeling of flows generates two basic patterns which are different instances of the center-area pattern. We study them in the next section.

13 The Two Basic Patterns

People on the move, as pilgrims in the Middle Ages, require services which are located along the circulation channels at a regular distance from each other. Thus, the primordial form of organized communication occurs at the edge of the channel by intercepting individuals on the move. That is why the village along the road, with

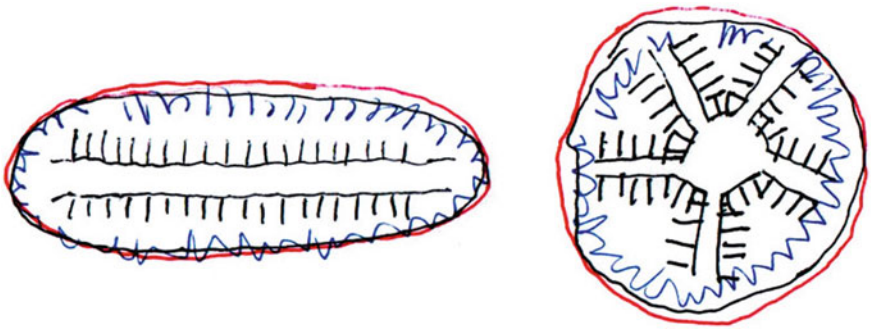


Fig. 6 The two basic patterns: the strip and the center

the buildings on both sides, is the primeval model of the settlement. It generates an axis, which is like an elongated center, and a periphery consisting of residential streets orthogonal to the main road.

The real center arises from the intersection of at least two main roads. When they are only two, as in the Roman camp, it is possible to keep the grid orthogonal. When more than two roads intersect, the radial model is adopted, with different road angles to which the shape of the blocks adapts. At the intersection of the streets, a square is formed which serves both traffic and parking. The building takes on a relevant architectonic aspect. Often there are monuments, that is particular buildings that communicate the meaning of the activities often connected to religion or power that take place in them in a lasting way. This pattern is represented with a circle, a special case of an ellipse with equal axes (Fig. 6). Both patterns coexist in the city.

14 The Self-similarity of the System of Patterns

Understanding what a city is like is not easy. Its functioning arises from the synergy between the central activities, the transport network and the movement of people. All of these elements are not visible on a geographic map. On the other hand, it is easier to understand a vast region with urban centers of various sizes placed where the roads intersect. In a regional context, the urban centers, that develop linearly along the roads, are separated by vast agricultural areas often eroded by urban expansion.

According to the principle of self-similarity, shapes of urban areas are scale-invariant: a smaller part of the shape is similar to the whole and vice-versa. So, we can consider a system of cities, after a scale reduction, as the image of the central areas of a city where activities are located (Fig. 7). The usefulness of the exercise consists in applying the macroscopic laws observed for the system of centers to the internal structure of a city.

On the other hand, studies have confirmed the validity of the Central places theory applied to the internal structure of cities and metropolitan areas [9, 52]. In other

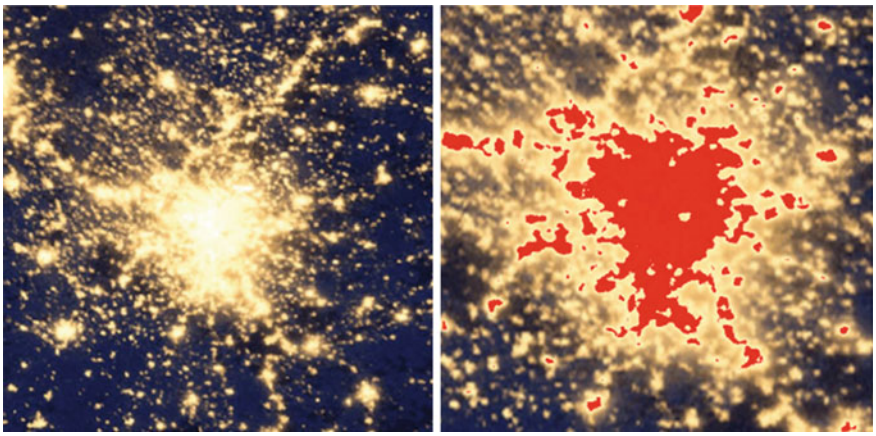


Fig. 7 Two images of Paris at night. The first on a large scale, the second on a smaller scale. In the second, the central areas with greater brightness are red colored. A statistical similarity of these central areas with the built-up areas of the first image can be detected which confirms the self-similarity of the structure of the settlement

words, the shopping and service centers inside the city are organized according to market areas and follow a semi-hierarchical structure.

The previous assumptions allow advancing in the theory of the urban internal structure. The city is composed of interconnected center-area patterns, to which the macroscopic rules found in the analysis of urban centers apply. Therefore the dimensional distribution (i.e. the surface) of the patterns will be similar to that described by Zipf's law. Moreover, we establish the dependency relationships between patterns with this method: a pattern depends on one, or two, larger patterns chosen at random. The application of the method produces a system of semi-hierarchical relations (Fig. 8) [44] similar to the semi-lattice proposed by Alexander. Furthermore, this type of structure, which is also self-similar, has the advantage of wholeness [25].

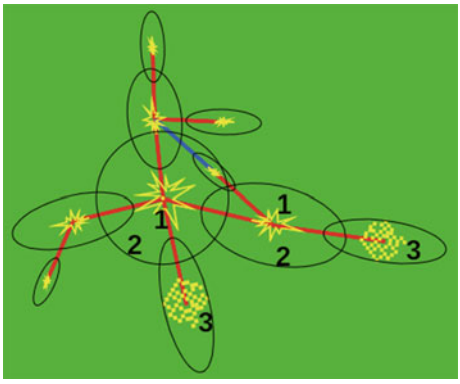


Fig. 8 Semi-hierarchical spatial organization of a simple system of patterns. 1: Centers. 2: Areas. 3: Special patterns. Red: Primary dependencies. Blue: Secondary dependencies

15 The Spatial Organization of the System of Patterns

Building on these theoretical bases, we show an example of a city generation, obtained using a program developed on the Netlogo platform.

The patterns are with a star in the middle that represents the center. A circular or elliptical edge delimits the area or spatial extent of the pattern. Some of the center-area patterns, different in size and variety of activities, are for specialized activities such as industrial areas or centers serving the entire population, e.g. hospitals, waste disposal, and commercial centers (Fig. 8). These centers have no surrounding suburbs or subordinate smaller centers. They are located on the outskirts near the main roads, making them easily accessible both from outside and from within the city. In these cases, the centers are represented in an enlarged and more scattered form, because they do not irradiate their effects in the immediate neighborhood.

The patterns are distributed more or less evenly throughout the space. The distance between each center and its nearest neighbor is similar on average, regardless of the size of the patterns. They may also overlap so that a large pattern may include smaller ones, as in the Central place theory. Environmental features such as lakes, sea, rivers and mountains can be added, although urban development should take place in a flat area.

To arrange the patterns in space, we establish the largest pattern in the center. The secondary patterns dependent on the largest one, are located at a distance proportional to their size. Patterns dependent on secondary patterns are located according to the same principle. The process is repeated for the subordinate patterns (Fig. 9). The shape of the pattern, circular or elliptical, depends on the ratio between the longest and shortest axis. We define this ratio as inversely proportional to the degree (number) of dependency relationships. So that more dependency relationships result in a circular

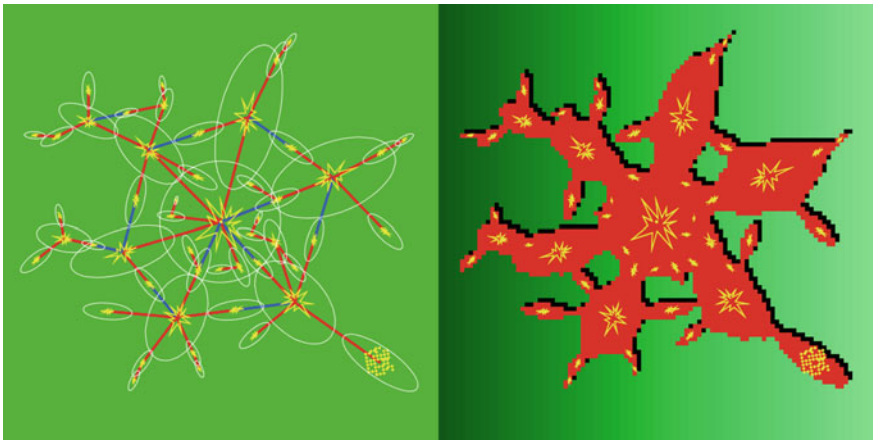


Fig. 9 The center-area patterns, with semi-hierarchical dependency relationship, and the corresponding built-up area

shape, fewer in an elliptical shape. The user can change the position of the patterns, to suit his/her needs or the morphology of the land. The patterns area determine the built-up urban area.

16 The Tracing of the Roads Network

After having established built areas and centers, we need to set up the crucial division between urban public space that allows internal communication and private living space. This division is achieved through the design of the road network that defines the urban morphology consisting of squares and urban axes. As a result, the shape of the blocks is also settled.

There are two principles for the organization of roads: (1) responding to the need of connecting points, typical of the regional scale; (2) ensuring widespread accessibility, together with the possibility of building or using the land for agriculture. The first principle yields triangulation which is the most efficient way to connect a pattern of points on a surface so that any pair of points are connected in the fastest way; the second principle generates the grid which determines rectangular blocks or fields.

The squared mesh is a surface that can be covered with parallel lines all of the same lengths. There are practical advantages in using these surfaces: if you plow a field, you make many parallel passes with the tractor equal in length. If you build a slab, put several parallel beams equal in length. Consequently, if you want to divide the surface in two, you get two equal parts, while the triangle does not give the same result.

The radial shape, often opposite to the grid, derives precisely from the use of the fabric of connections between points, on an urban scale, such as the union of triangles having a common vertex (the center). When triangles are equilateral we get a hexagon that has been proposed for city planning [7]. But usually, the urban network is a combination of the two types of networks that must simultaneously satisfy opposite needs: building and accessibility, and connection between points.

The resulting road network does not have a hierarchical structure like the roots of a tree or the bronchi of the lungs. In those hierarchical networks, it is not necessary to connect two points at the base of the hierarchy, because the flows are tightly organized between the center and the periphery. In the city, however, communication, even if centralized, is made possible between each pair of points. Suppose a hierarchical network and two people in the outermost branch that want to communicate. Instead of going through the center, they will find a direct link, that is, a shortcut that will allow them to communicate easily. This shortcut mechanism produces urban loops, ring roads and blocks (Fig. 10). Overall, the road network is no longer strictly hierarchical.

To trace the road network, two types of roads are distinguished: urban and regional. The first connects the centers within the city, ensures accessibility to all surrounding points and is therefore permeable. The second connects the points, passing near the edges of the pattern area and away from the centers.

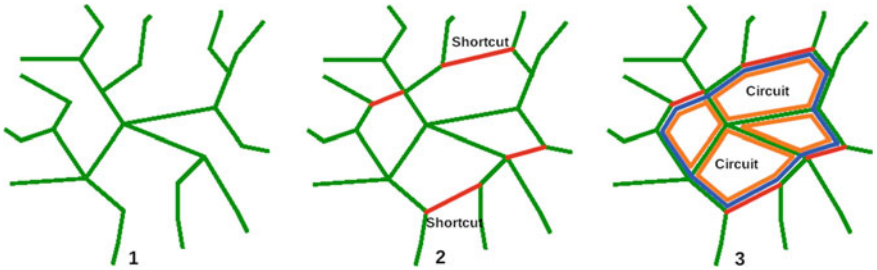


Fig. 10 1: The hierarchical network. 2: Shortcuts are added. 3: Circuits are created

To trace the first type of roads, it helps to simulate random movements between the centers, so that they occur without changes of direction. That is a crucial principle resulting in the urban axes which, along with squares, form the structure of the urban space. To perform the simulation, agents are left free to travel from a randomly chosen center to the most important and nearest one without major changes of direction. In this way, the most relevant possible urban axes are identified, which will be the central elements of the organization of the urban space.

Starting from the identification of the preferred paths, one can manually draw the roads, making sure that they connect the main centers (Fig. 11). Once the main roads have been traced, it is possible to automatically trace the minor roads that allow both access to all points and the definition of the blocks, using the method of fracturing a surface according to the axis of least resistance (Fig. 12). The process stops when the block surface is less than a threshold.

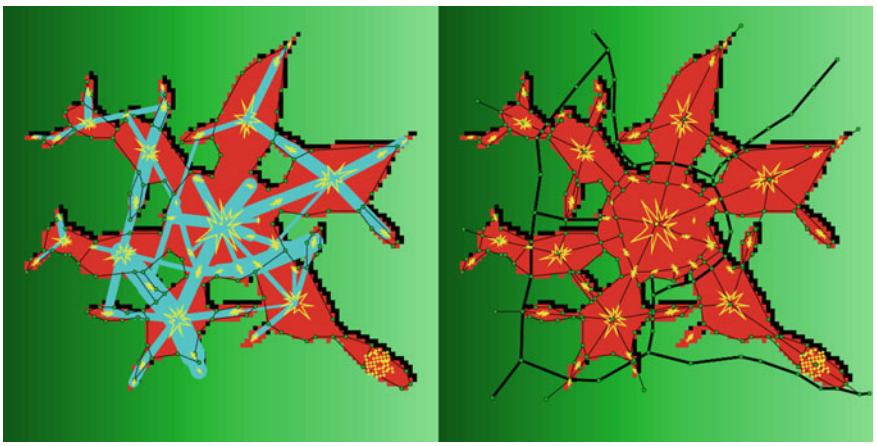


Fig. 11 The main flows, as calculated with agent simulation and the road network as manually designed. The thickest lines are motorways, the others are common roads

17 The Topological Aspects of the Roads Network, and the Production of Urban Space

The network of streets defines the space within the city. The designed urban viability is analyzed with the methods for the analysis of planar graphs: the configurational analysis (the indirect distance of each road axis from the others, [21]), betweenness (the most traveled road section [18]) and distance of each road section from the others [6]. The weighted value of these three measures represents the importance of the road for the urban space. We then calculate the centrality of the urban fabric in each point considering two factors: the importance of the roads nearby and the accessibility to the centers of the patterns.

This analysis makes it possible to establish the width of the streets based on their importance. Squares are located in the most relevant intersections, or close to the centers of the patterns. So, we can define the urban space in its main components: squares and urban axes. The size of squares is proportional to number, if greater than 4, of streets intersecting in it or to the size of the connected pattern center. Streets and square are distributed according to Zipf's law: few are important and many are small but diffused all over the city. That results in a pleasant urban space consisting of a large variety of places [13].

The centrality of the urban fabric allows each block to be occupied by the most suitable use: that is, exclusively by business activities , where the centrality is higher, by residence mixed with activities, and by residence alone, where the centrality is lower. Furthermore, the high centrality combined with the adjacency to a square identifies the blocks for monumental buildings. Some blocks can also be cleared to use the land for gardens and urban parks (Fig. 12).

Centrality represents the value that the land would have in case of realization of the town. The established use depends on the offer per square meter of the various

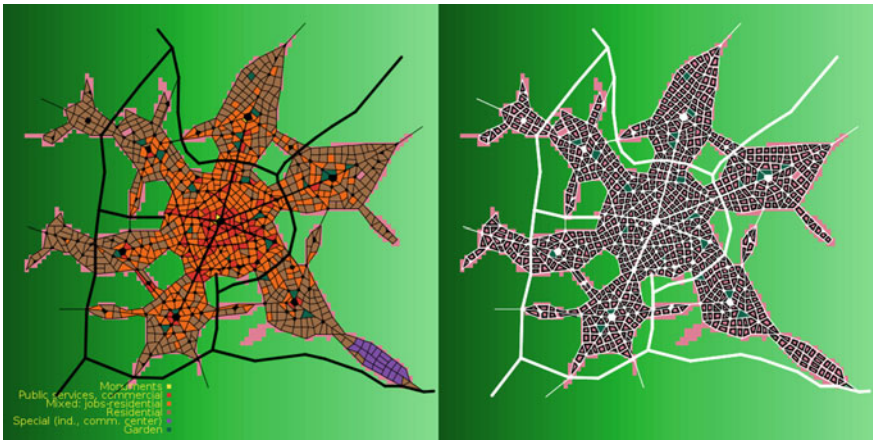


Fig. 12 The urban spaces with squares and urban axes, including land use, and blocks

Fig. 13 A 3d view of the city

activities. As can be seen from Fig. 12, the distribution of activities and residences recalls that resulting from the von Thunen-Alonso bid rent theory. When the planned use of the land is similar to that obtained applying the economic equilibrium, the project can easily resist the economic dynamics that will follow its realization.

The block design is the final act of the process. Blocks are the basic elements of the urban fabric. They are like tiles in a tessellation and similar to citadels within the town. Hence they are isolated, in opposition to the continuity of public space. The block is defined by the streets and establishes the shape of the buildings [36] whose front face the street while the rear-facing the inner courtyard. In so doing, blocks contribute to the definition of the road or square they face (Fig. 12). The resulting 3d, shown in Fig. 13, is just a sketch that needs to be detailed by architectonic projects.

18 Final Remarks

At the end of the process, we get a city that might look like existing cities where each part collaborates in developing the whole. Wholeness depends on the elements, from the few large to the many small, none alike, that make up a city. This distribution favors the emergence of variety and therefore of beauty [26].

What has been shown is just an example. The method leaves room for the human user who intends to design a city while maintaining the general principles based on the use of patterns and their interrelation. See for instance Fig. 14 which shows a 3d view of a city on the seaside, crossed by a river. However, many aspects have been treated incompletely or too broadly. Here is a list.

First, the urban design assumes that the land is flat, as in most large cities. But that's not always the case. It is, therefore, necessary to face the problem of building a road network on hilly terrain. The road network should be adapted to the slopes of the ground, following the contour lines [10, 15].

Moreover, relevant transport infrastructures such as railway stations and the railway network are not included in the method. They can have a double effect. On

Fig. 14 A 3d view of a city on the seaside, crossed by a river

the one hand, they strengthen the centrality of the place, on the other, the network infrastructure may be a barrier for the organization of the city. In general, the barriers that constitute an element of the urban organization have not been considered.

We have sketched a simplified version of the block, adapted to the shape it takes in the traditional western city. The block can have different variants depending on the density, the cultural tradition, the economic model, as it is explained by the Transect Theory [16]. In the peripheral areas of the city, blocks consist mainly of single single-family houses, while in denser urban areas they are built internally to use all the land parcels, due to the high land rent. In the special patterns, blocks are treated with the usual method, while their shape varies, namely in industrial areas where the buildings get light from the roof and require large continuous surfaces. Only an architectural project can solve these problems specifying in detail the shape of the buildings, having as references the urban morphology established by the road system and the use of the land.

Last but the hardest is the intervention in an existing city. What I have shown is a method for designing a new entire city. But most often we add new parts to a town or demolish and rebuild parts to introduce new functions and regenerate neighborhoods. Therefore, to make the method more useful, it must be able to deal with the transformation of a city. It involves studying the organization of the existing city with the identification of the patterns and their dependencies. Then we should develop a method of reorganization by adding new patterns organically linked to the existing ones, and connecting the new roads network.

19 Conclusions

Starting from complex dynamics, a theory of city architecture, including the basic elements and how to compose them, has been shown. Is a theory so relevant for designing a city, even if it risks being a mere theoretical exercise? I think so because a city encompasses all aspects of architectural design: unique buildings such as

monuments, but also ordinary artifacts such as residences, infrastructure and public spaces. We consider these aspects in their interactions, as in a self-organized city, not in a fragmentary way, as in architectural practice. To be able designing an entire city is not a simple school exercise. It is a way to understand the role of each new part of the city and each new building with all existing ones.

Anyway, the reality remains the best book from which to learn. That is why I chose the method of imitation. It may appear a choice anchored more to past experiences than proposing solutions for the future. But it is not so. Imitation is not a way to become a simple mirror of reality, but a way to learn from it. Learning from reality means learning from the collective intelligence that has probably already solved the problem we are facing.

Based on imitation, the theory proposes a top-down method. It starts from the definition of social facts to then arrive at the architectural ones. The organization of centre-area patterns determine the spatial organization. The design of the urban space takes place later. The theory is based on the hypothesis that the normal state of an individual is movement and that communication is an elementary need. For these reasons, the road network is the central element in the definition of urban space. Blocks are designed according to the viability, as in spontaneous growth. Hence each part integrates with the whole.

Wholeness is the key word and the goal of the proposed method. We need a comprehensive approach to the conception of the city. Architecture can face this challenge if it renounces the claim to operate with abstract principles extraneous to people's daily life, recognizing that it does not belong exclusively to an elite because all individuals are potentially involved, as the urban fabric results from a collective work.

Acknowledgements I am particularly grateful to Sylvie Occelli, Sergio Albeverio and Valerio Cutini for discussing and reviewing the text. Disclaimers apply as usual.

References

1. Alexander, C.: The Timeless Way of Building. Oxford University Press, New York (1979)
2. Alexander, C., Ishikawa, S., Silverstein, M.: A Pattern Language: Towns, Buildings, Construction. Oxford University Press, New York (1977)
3. Alonso, W.: Location and Land Use: Toward a General Theory of Land Rent. Harvard University Press, Cambridge, MA (1964)
4. Barthelemy, M., Bordin, P., Berestycki, H., Gribaudi, M.: Self-organization versus top-down planning in the evolution of a city. Sci. Rep. **3**, 11 (2013)
5. Batty, M.: Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals. The MIT Press, Cambridge, USA (2005)
6. Bavelas, A.: A mathematical model for group structures. Hum. Organ. **7**, 16–30 (1948)
7. Ben-Joseph, E., Gordon, D.: Hexagonal planning in theory and practice. J. Urban Des. **5**, 237–265 (2000)
8. Bonino, M., Governa, F., Repellino, M.P.: The City After Chinese New Towns: Spaces and Imaginaries from Contemporary Urban China. Birkhäuser, Basel (2019)

9. Carol, H.: The hierarchy of central functions within the city. Ann. Assoc. Am. Geogr. **50**, 419–438 (1960)
10. Castillo, M.: Urban patterns and disaster risk: the informal city on the hills. In: Kim, Y.O., Park, H.T., Seo, K.W. (eds.) Proceedings of the 2013 International Space Syntax Symposium, pp. 114:1–114:13. Sejong University Press, Seoul (2013)
11. Cheddadi, M.A., Hotta, K., Ikeda, Y.: Reinterpreting self-organizing urban tissues by designing a generative model. In: Agrawal, A., Gupta, R. (eds.) Revisiting the Role of Architecture for Surviving Development. 53rd International Conference of the Architectural Science Association, ANZAScA 2019 - Proceedings, pp. 175–184. Architectural Science Association (2019)
12. Christaller, W.: Die zentralen Orte in Süddeutschland. Fisher Verlag, Jena (1933)
13. Crompton, A.: The fractal nature of the everyday environment. Environ. Plann. B. Plann. Des. **28**, 243–254 (2001)
14. Cuthbert, A.R.: Urban design: requiem for an era - review and critique of the last 50 years. Urban Des. Int. **12**, 177–223 (2007)
15. Cutini, V.: Axial lines and contour lines: climbing up the centre. In: Kubat, S., Ertekin, O., Güney, Y.I., Eyüboğlou, E. (eds.) Proceeding 6th International Space Syntax Symposium, pp. 094.1–094.14. ITU Faculty of Architecture, Istanbul (2007)
16. Duany, A., Talen, E.: Transect planning. J. Am. Plann. Assoc. **68**, 245–266 (2002)
17. Frankhauser, P.: La Fractalité des aires urbaines. Anthropos-Economica, Paris (1994)
18. Freeman, L.C.: A set of measures of centrality based upon betweeness. Sociometry **40**, 35–41 (1977)
19. Gandelsonas, M.: The city as the object of architecture. Assemblage **37**, 128–144 (1998)
20. Garnier, T.: Une Cité industrielle: Étude pour la construction des Villes. Ch. Massin & Cie, Paris (1932)
21. Hillier, B., Hanson, J.: The Social Logic of Space. Cambridge University Press, Cambridge, UK (1984)
22. Holubec, P.: Conceptualizing the urban system as a system of flows. In: Walloth, C., Gebetsroither-Geringer, E., Atun, F., Werner, L.C. (eds.) Understanding Complex Systems, pp. 79–93. Springer International Publishing, Switzerland (2016)
23. Howard, E.: Garden Cities of Tomorrow. Swan Sonnenschein & Co., London (1902)
24. Hursh, J.: Slums could inspire the cities of the future. Here's how. https://www.weforum.org/agenda/2018/01/this-startup-is-turning-slums-into-microcities/ (2018)
25. Jiang, B.: Wholeness as a hierarchical graph to capture the nature of space. Int. J. Geogr. Inf. Sci. **29**, 1632–1648 (2015)
26. Jiang, B.: Living structure down to earth and up to heaven: Christopher Alexander. Urban Sci. **3**, 1–20 (2019)
27. Keeble, L.: Principles and Practice of Town and Country Planning. Estates Gazette, London (1952)
28. Kelso, J.A.S., Stolk, E., Portugali, J.: Self-organization and design as a complementary pair. In: Springer Proceedings in Complexity, pp. 43–53 (2016)
29. Krier, L.: The Architecture of Community. Island Press, Washington, DC (2009)
30. Krugman, P.: First nature, second nature, and metropolitan location. J. Reg. Sci. **33**, 129–144 (1993)
31. Marshall, S.: Cities, Design and Evolution. Routledge, London (2009)
32. Meier, R.: A Communications Theory of Urban Growth. MIT Press, Cambridge, MA (1962)
33. Molotch, H.: The city as a growth machine: toward a political economy of place. Am. J. Sociol. **82**, 309–332 (1976)
34. Netto, V.M., Meirelles, J., Ribeiro, F.L.: Social interaction and the city: the effect of space on the reduction of entropy. Complexity (2017)
35. Oliver, P.: Built to Meet Needs: Cultural Issues in Vernacular Architecture. Routledge, London (2006)
36. Panerai, P., Castex, J., Depaule, J.-C.: Formes urbaines: de l'îlot á la barre. Editions Parenthèses, Marseille (1997)

37. Pedersen Zari, M.: Regenerative Urban Design and Ecosystem Biomimicry. Routledge, Oxon, UK (2018)
38. Perry, C.A.: The Neighborhood Unit, a Scheme of Arrangement for the Family-Life Community. Committee on Regional Plan of New York and Its Environs, New York (1929)
39. Portugali, J.: Self-organization and the City. Springer, Berlin (2000)
40. Rossi, A.: The Architecture of the City. The MIT Press, Cambridge, MA (1982)
41. Rudofsky, B.: Architecture Without Architects: A Short Introduction to Non-Pedigreed Architecture. The Museum of Modern Art, New York (1965)
42. Salingaros, N.A.: Principles of Urban Structure. Techne, Delft (2005)
43. Schelling, T.C.: Models of segregation. Am. Econ. Rev. **59**(2), 488–493 (1969)
44. Semboloni, F.: Hierarchy, cities size distribution and Zipf's law. Eur. Phys. J. B **63**, 295–301 (2008)
45. Semboloni, F.: Cittá e sistemi urbani. Analisi e progetto. In: Albeverio, S., Giordano, P., Vancheri, A. (eds.) Metodi e Modelli Matematici per le Dinamiche Urbane, pp. 215–247. Springer, Italia (2021)
46. Simon, H.: The Sciences of the Artificial. The MIT Press, Cambridge, MA (1969)
47. Sinha, A.: Design of settlements in the Vaastu Shastras. J. Cult. Geogr. **17**, 27–41 (1998)
48. Trummer, P.: The city as an object: thoughts on the form of the city. Log **27**, 51–57 (2013)
49. von Glasersfeld, E.: An introduction to radical constructivism. In: Watzlawick, P. (ed.) The Invented Reality, pp. 17–40. Norton, New York (1984)
50. Wang, L.: China's new town movements since 1949: a state/space perspective. In: Progress in Planning, pp. 1–38 (2020)
51. Watson, B., Müller, P., Veryovka, O., Fuller, A., Wonka, P., Sexton, C.: Procedural urban modeling in practice. IEEE Comput. Graph. Appl. **28**(3), 18–26 (2008)
52. West, D.S., Hohenbalken, B.V., Kroner, K.: Tests of intraurban central place theories. Econ. J. **95**, 101–117 (1985)
53. Wu, F., Webster, C.J.: Simulation of land development through the integration of cellular automata and multicriteria evaluation. Environ. Plann. B. Plann. Des. **25**, 103–126 (1998)
54. Zipf, G.: Human Behavior and the Principle of Least Effort. Addison Wesley, Cambridge, MA (1949)

Randomness, Emergence and Causation: A Historical Perspective of Simulation in the Social Sciences

Raffaello Seri, Davide Secchi, and Mario Martinoli

Abstract This chapter is a review of a selection of simulation models, with special reference to the social sciences. Three critical aspects are identified—i.e. randomness, emergence and causation—that may help understand the evolution and the main characteristics of these simulation models. Several examples illustrate the concepts through a historical perspective.

Keywords Simulation models · Randomness · Emergence · Causation · Agent-based models

1 Introduction

The following pages present a selection of computational models and techniques that have been used in the last 70 years and provide an overview of how the field has evolved. In an era of cheap and fast computation, it is particularly important to look back at the history to understand the specific reasons that make current advanced techniques so remarkably relevant to social scientists.

We do not pretend to present a comprehensive overview of simulation modeling techniques and systems, but have selected those that we believe have contributed the most to build current simulation approaches. In so doing, we identify a thread, concerning *randomness*, *emergence* and *causation*, that specifies some of the most relevant characteristics of current techniques. In an attempt to make these comments

R. Seri (✉)
DiECO, Università degli Studi dell'Insubria, Varese 21100, Italy
e-mail: raffaello.seri@uninsubria.it

R. Seri · D. Secchi
Department of Language and Communication, Centre for Computational & Organisational Cognition (CORG), University of Southern Denmark, Slagelse 4200, Denmark

M. Martinoli
Istituto di Economia and EMbeDS, Scuola Universitaria Superiore Sant'Anna, Pisa 56127, Italy

S. Albeverio et al. (eds.), *Complexity and Emergence*, Springer Proceedings in Mathematics & Statistics 383, https://doi.org/10.1007/978-3-030-95703-2_7

as much visible as possible, they are found under the header "Intermezzo", as they interrupt the flow of the presentation and weave together the different methods.

Before moving forward, a few words need to be spent on these three aspects and why they are so important as to appear as the 'fil rouge' of the chapter. Let us start from *randomness*. Any simulation of a social system has to be able to reproduce elements that appear unpredictably and without any apparent connection to the phenomenon under analysis (or to an outcome variable). The reason is connected to Laplace's demon (see [31, p. 2]): an omniscient intellect having complete knowledge of all the forces and positions of the items composing the natural world as well as unlimited computing possibilities would be able to predict without error all future events. Randomness is a way of accounting for our ignorance of these initial conditions and for our computational limits. That is why a model not taking into account the possibility of unpredictable and/or external events would fall short of capturing the inherent complexity of most phenomena. The second aspect, *emergence*, is tied to the assumption that social systems are complex [29, 50]. When this assumption holds for the computational simulation that models a social system, then uncertainty, ambiguity, and unpredictability are key features of that research effort. Some have argued that a successful simulation is one that presents the modeler with counter-intuitive and surprising results (see, e.g., [54]). We do not subscribe to this view, because it is too radical and only fits certain types of simulations. However, we can certainly support the idea that simulating a social system means to allow for a true intellectual enquiry, where results are not entirely discernible by simply looking at the code. On this respect, "a model is *not* a model" (opposite to what some argue in a recent editorial, see [78]). The third aspect is that of *causation*. One of the defining aspects of a computational simulation refers to the mechanisms that specify how its component parts behave. The interactions of these parts are reflected in the values taken by the aggregate variables describing the system that, on their turn, impact the single components. The components and the aggregates are thus linked by up- and downward relations. In a social system, both causal directions need to be present to explain most phenomena. The history of computational simulation in the social sciences has always bounced back and forth between these two levels, and settled on recent techniques that could account for both.

The methods we are going to present are very heterogeneous. Some of them are specified at the level of the individual, others are aggregate. Some of them require interactions between agents, others don't. Some of them are deterministic, others contain random elements. However, all of them share the same two characteristics: (a) objects (individuals or quantities) are reduced to a finite number of idealized types (that may vary in quality); (b) objects are specified by relations partially or fully connecting them together. The historical review that we are going to present will show how the two mechanisms that have governed the evolution of these methods are indeed the identification of basic units of analysis, at whatever level they are defined, and the determination of the mechanisms that connect them. The solutions that have been proposed to these two questions have led to the development of several simulation methods.

Now we come to the structure of the chapter. In Sect. 2, we review some computational experiments involving early computers. In Sects. 3 and 4 we respectively review System Dynamics and Discrete-Event Simulation, two methods of inquiry considering the aggregate behavior of a system. Then, we review Microsimulation techniques in Economics and Political Science in Sect. 5. Section 6 covers Cellular Automata while Sect. 7 introduces Agent-Based Models. Section 8 wraps up the main conclusions.

2 Experiments with Early Computers

2.1 ENIAC

One of the first electronic computers—the ENIAC, Electronic Numerical Integrator and Computer—was built at the beginning of 1945 at the University of Pennsylvania in Philadelphia [105, p. 125]. In the spring of 1946 at Los Alamos, Stan Ulam suggested that ENIAC could be used to resuscitate some statistical sampling techniques that "had fallen into desuetude because of the length and tediousness of the calculations" [105, p. 126]. He discussed the idea with John von Neumann, who sent, on March 11, 1947, a letter to the leader of the Theoretical Division of the Los Alamos National Laboratory, Robert Richtmyer, with "a detailed outline of a possible statistical approach to solving the problem of neutron diffusion in fissionable material" [105, p. 127]:

> The idea then was to trace out the history of a given neutron, using random digits to select the outcomes of the various interactions along the way. […] von Neumann suggested that […] "each neutron is represented by [an 80-entry punched computer] card …which carries its characteristics," that is, such things as the zone of material the neutron was in, its radial position, whether it was moving inward or outward, its velocity, and the time. The card also carried "the necessary random values" that were used to determine at the next step in the history such things as path length and direction, type of collision, velocity after scattering–up to seven variables in all. A "new" neutron was started (by assigning values to a new card) whenever the neutron under consideration was scattered or whenever it passed into another shell; cards were started for several neutrons if the original neutron initiated a fission [41, p. 133].

This led Nicholas Constantine Metropolis and Stanislaw Ulam to introduce, in 1949, the name of *Monte Carlo method* [106] for a statistical sampling method:

> I (Metropolis) suggested an obvious name for the statistical method—a suggestion not unrelated to the fact that Stan had an uncle who would borrow money from relatives because he "just had to go to Monte Carlo" [105, p. 127].

The name of Monte Carlo method is nowadays generally used to denote a rather heterogeneous array of techniques for solving mathematical problems by:

- reducing their solution to the computation of an expectation with respect to a random variable and

- approximating this expectation with the empirical average based on a sample of realizations of the random variable.

The simplest example is the integration of a function on a bounded domain.

Example 1 (*Monte Carlo integration*) The aim is to compute the integral of a function f defined on a bounded domain that we identify, without loss of generality, with the unit interval [0, 1]. The integral is $\int_0^1 f(x)\,\mathrm{d}x$. A solution is to remark that a random variable X uniformly distributed over the interval [0, 1] has probability density function 1, so that $\int_0^1 f(x)\,\mathrm{d}x = \mathbb{E}f(X)$. This means that, if we have a sample $\{x_1, \ldots, x_n\}$ of realizations from X, we can approximate $\int_0^1 f(x)\,\mathrm{d}x = \mathbb{E}f(X)$ through $\frac{1}{n}\sum_{i=1}^n f(x_i)$.

Modern statistical sampling methods largely predate the Monte Carlo method.

Example 2 (*Buffon's needle*) An often-misquoted antecedent is Buffon's needle, a mathematical problem requiring to compute the probability that a needle of length ℓ, randomly cast on a floor with equally-spaced parallel lines at a distance d, lands on a line. The problem was presented by Buffon in 1733 at the Académie Royale des Sciences of Paris (see [16], where the problem is presented but not solved). A solution was described in [17, pp. 100–105] (see also [30, pp. 359–360]). The original aim of Buffon was to look for an explicit solution. The probability itself is $2\ell/\pi d$ (when $\ell \leq d$) and this explains why several authors have performed the experiment repeatedly to provide, through an empirical approximation to the probability, an approximation to π. Despite the problem lends itself to a sampling solution, there is no evidence that Buffon ever tried to do that. Notwithstanding this, many authors (among which, e.g., [95, p. 120]) have written that Buffon proposed a sampling solution: what probably has misled them is the fact that [33] and [34, pp. 170–171] presented the needle problem together with another problem that Buffon studied with 2048 trials (this is confirmed by the fact that [95, p. 120] states that Buffon tried 2048 tosses in the needle problem). Other authors (see [142, p. 117]) have attributed the sampling solution to [30, p. 360]: despite Laplace is indeed talking about the limit for a large number of draws, he is probably making a reference to a frequentist argument rather than to a real sampling solution to the problem. However, the discussion in [33] (reprinted in [34, pp. 170–171] and [83]) suggests that the sampling approximation of π through Buffon's needle was in use as early as 1855. A search of the literature leads to an even older example, namely [159] (see also [127]).

2.2 *Intermezzo*

As no computer-based method for building random numbers was (and still is) known, John von Neumann went on to study algorithms to generate *pseudo-random numbers*, i.e. numbers with characteristics similar to those of random numbers. A famous but somewhat trite quotation is:

> Any one who considers arithmetical methods of producing random digits is, of course, in a state of sin. For, as has been pointed out several times, there is no such thing as a random number—there are only methods to produce random numbers, and a strict arithmetic procedure of course is not such a method [154, p. 36].

The interpretation that is generally given to this sentence is inaccurate. Von Neumann was not being skeptical, as often interpreted, of the usefulness of pseudo-random numbers. He was just suggesting that "'cooking recipes' for making digits […] probably […] can not be justified, but should merely be judged by their results" [154, p. 36]. This is the main way in which *random-number generators* (*RNG*, though a better name would be *pseudo-random-number generators*, *PRNG*) are evaluated today, through batteries of statistical tests in which their behavior is compared with the theoretical behavior of true random numbers (e.g., [15, 100]). As computer simulation was, at that time, very difficult, random digits were collected in publications among which the famous *A Million Random Digits with 100,000 Normal Deviates* [27], published in 1955.

However, even before and around the construction of ENIAC several computational experiments were being performed using analog computers, and they often used random numbers too.

2.3 FERMIAC

In the 1930s, when he was still in Rome, Enrico Fermi was studying neutron transport: when a neutron (a sub-atomic particle) passes through matter, it can interact with other particles, or it cannot. However, the aggregate behavior of neutrons seemed out of reach. Fermi assumed that each neutron was like an agent whose behavior was dictated by the sampling of some random numbers. He then computed the aggregate results for a large numbers of neutrons on a mechanical calculator:

> Fermi had invented, but of course not named, the present Monte Carlo method when he was studying the moderation of neutrons in Rome [138, p. 221].

According to his student Emilio Segrè (see [105, p. 128]), Fermi kept the technique secret and used it to solve several problems. His colleagues were often astonished by the precision of his computations. These back-of-the-envelope calculations contributed to create the myth of so-called *Fermi estimates*.

In the late 1947 the ENIAC (see Sect. 2.1) was moved at the Ballistics Research Laboratory in Maryland. During the inactivity of the ENIAC, Fermi built an analog simulator of neutron transport, called the FERMIAC (a pun on ENIAC; an image can be found in [105, p. 129]). In the FERMIAC, neutrons were modelled as agents in a planar region, whose behavior was affected when a material boundary was crossed.

2.4 MONIAC

In 1949, at the London School of Economics (LSE), Bill Phillips (Alban William Housego "A. W." "Bill" Phillips, who later introduced the Phillips curve) built an hydraulic machine called *Monetary National Income Analog Computer* or *MONIAC* (see [116]). The name was a pun on "money" and "ENIAC". It was a system of tanks and valves through which water flowed, in a simulation of money circulation in the UK economy. The example of the MONIAC looks very distant from the ones we will see below, but it contains the features outlined above: here the tanks are the objects; the valves are the relations connecting the objects.

Both the FERMIAC and the MONIAC were examples of *analog* or *analogue computers*, i.e. machines using physical (electrical, mechanical, or hydraulic) phenomena expressed in terms of variables measured on a continuous scale to model a phenomenon. Analog computers were very common when no other computing method was available. The difference with respect to digital computers is that the latter use information stored in discrete form: the earliest digital computers were *program-controlled*, i.e. they were programmed by modifying the physical structure (plugs, wires, etc.) of the machine; modern computers are *stored-program*, as the program is stored in memory, without hardware modifications.

3 System Dynamics

System dynamics (*SD*) is an approach to the dynamical study of systems composed of objects in interaction. The idea of SD is to model the change over time of some quantities through feedback loops, accumulation of flows into stocks, and identification of inflows and outflows. The system is generally represented first graphically as a diagram and then mathematically as a system of differential (or difference) equations, that are then solved numerically by a computer program. Its central insight is the fact that the structure connecting the components is sometimes more important than the components themselves in determining the behavior of the system. It was founded, as a branch of *systems theory* [149], by Jay Wright Forrester in the 1950s (see, e.g., [53] or the historical accounts in [55, 81]). At the beginning, SD was developed to analyze complex business problems, in connection to the author's position at the MIT Sloan School of Management. It has been applied to several problems ever since. SD goes through a series of steps to transform a verbal description of the phenomenon under scrutiny into a mathematical model (see, for example, [144] for a worked-out example on new product adoption).

Example 3 (*Lotka-Volterra Model*) An early example of a system of differential equations in which the elements appearing in the system can be interpreted as feedback loops is the Lotka–Volterra Model (LVM), developed by Lotka [90, 91, pp. 92–94] and Volterra [151, 152] in their seminal work. The LVM is a predator-prey model describing the dynamics of two species—i.e. predators and preys—interacting

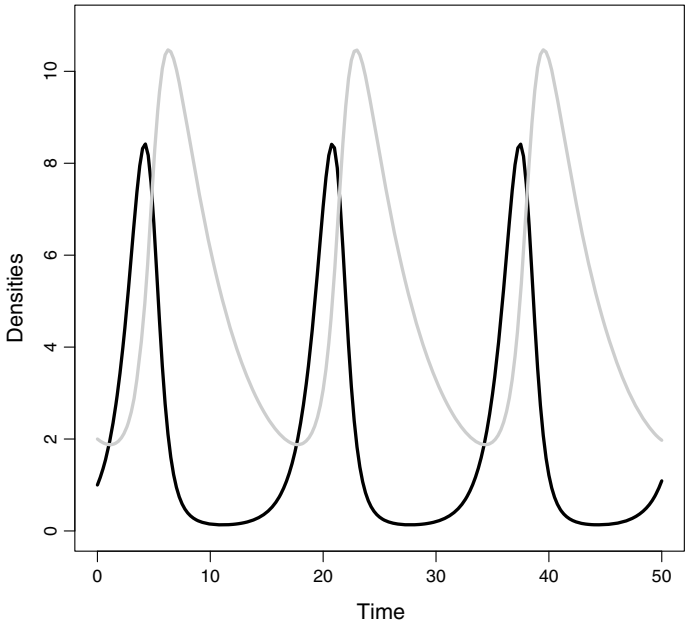


Fig. 1 Population densities of the two species in the Lotka–Volterra model (predator in grey, prey in black)

in an ecological environment. The presence of feedback loops is made clear both by Lotka (see the graphical representation in [90, pp. 411–412]) and by Volterra:[1] "conviene […] di schematizzare il fenomeno isolando le azioni che si vogliono esaminare e supponendole funzionare da sole, trascurando le altre" [152, p. 31]. A SD approach is in [36], while the final behavior of the system is illustrated in Fig. 1.

Example 4 (*The Limits to Growth*) The Club of Rome is a think tank founded in 1968 in Rome as "an informal association of independent leading personalities from politics, business and science, men and women who are long-term thinkers interested in contributing in a systemic interdisciplinary and holistic manner to a better world." Their 1972 book *The Limits to Growth* [104] used SD to study the world economy and population, and raised considerable interest and concern about their sustainability.

Example 5 (*The Lorenz system*) In 1963, Edward Norton Lorenz studied atmospheric convection through differential equations (see [89]). He realized that a small change in the initial conditions could have long-term effects on the behavior of the system. The idea is to take two starting points on two nearby trajectories: moving along them, they will eventually diverge. Similar insights had already been advanced

[1] In English, "it is more effective […] to schematize the phenomenon by isolating the actions that one wants to examine and assuming they behave independently, irrespectively of the others" (our translation).

by Henri Poincaré in 1890 while studying the three-body problem and by Jacques Hadamard while studying motion on surfaces of negative curvature, but had little impact on the literature. Lorenz's discovery, instead, sparked a small revolution. It led to the identification of so-called *deterministic chaos, chaos theory* or, simply, *chaos*, i.e. sensitivity to initial conditions in deterministic systems (often called *dynamical systems*). Lorenz coined the term *butterfly effect* for this phenomenon. An oft-quoted sentence is taken from the title of Lorenz's talk at the 139th meeting of the American Association for the Advancement of Science in 1972: "Does the flap of a butterfly's wings in Brazil set off a tornado in Texas?"

Instances of chaos in models from several domains, and SD among them, were described. As an example, in 1986, Erik Mosekilde and Javier Aracil received the Jay W. Forrester Award for their work on chaos in SD.

Example 6 (*A Sound of Thunder*) In the June 28, 1952, issue of Collier's magazine, a science fiction short story by Ray Bradbury was published under the title *A Sound of Thunder* [13]. It described a time travel into the past whose impact on the future goes awry because of a butterfly (no spoilers). This story is sometimes miscredited with the origin of the name butterfly effect but, despite being a wonderful example of the very concept, it had no bearing on its development.

3.1 Intermezzo

While chaos is extremely important from a theoretical point of view, its relevance in real examples is difficult to work out:

> An essential point made by Poincaré is that *chance* and *determinism* are reconciled by long-term unpredictability. Here it is, in one crisp sentence: *A very small cause, which escapes us, determines a considerable effect which we cannot ignore, and we then say that this effect is due to chance* [130, p. 48; emphasis in the original].

The problem with chaos is that the dependence on the initial conditions makes difficult to forecast the future of the system, as initial conditions are always observed with a small error. This is why chaotic dynamical systems may be modelled as *stochastic processes*:[2]

> En dernière analyse, le hasard réside donc […] dans l'œil de l'observateur. [43, p. 14]

For us, what matters most is that chaos is a property of the system that is not shared by its components when considered in isolation. Properties like this are called *emergent*:

> The ability to reduce everything to simple fundamental laws does not imply the ability to start from those laws and reconstruct the universe. […] The constructionist hypothesis breaks down when confronted with the twin difficulties of scale and complexity. […A]t each level of complexity entirely new properties appear. […] Psychology is not applied biology, nor is biology applied chemistry. […T]he whole becomes not only more than but very different from the sum of its parts. [3, pp. 393–395]

[2] In English: "So, in the end chance lies […] in the eye of the observer" [44, p. 4].

Emergent properties arise when the system as a whole displays a behavior that is not explicit in its single components. As put forth in [3], a system of interacting quantities/agents is not only more than the sum of its components, it is different from their sum. We will see below some examples of *emergence* (some authors use *supervenience* for a related concept).

For the moment we review some of the history of the concept. As remarked in [79, p. 49], one of the first disciplines to embrace emergence as its central phenomenon was Economics, through the work of Adam Smith:[3]

> [E]very individual […] neither intends to promote the publick interest, nor knows how much he is promoting it. […H]e intends only his own gain, and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention. […] By pursuing his own interest he frequently promotes that of the society more effectually than when he really intends to promote it. [141, p. 35]

The philosopher John Stuart Mill wrote, when dealing with failures of the principle of the Composition of Forces:

> The chemical combination of two substances produces, as is well known, a third substance with properties entirely different from those of either of the two substances separately, or of both of them taken together. Not a trace of the properties of hydrogen or of oxygen is observable in those of their compound, water. […W]e are not, at least in the present state of our knowledge, able to foresee what result will follow from any new combination, until we have tried it by specific experiment. [107, pp. 426–427]

The metaphor of water was a recurrent one in the work of early emergentists [12, p. 37]. In Biology, Thomas Henry Huxley, in the book [71, pp. 16–17], introduced the idea that "there is no sort of parity between the properties of the components and the properties of the resultant": he used the "aquosity" of the oxide of hydrogen (i.e. water) as a comparison for the "vitality" of living systems, and he admonished that those who say that "the properties of water may be properly said to result from the nature and disposition of its component molecules" are "placing [their] feet on the first rung of a ladder which, in most people's estimation, is the reverse of Jacob's, and leads to the antipodes of heaven." Another oft-quoted antecedent involving a different discipline is the recognition by the French sociologist Émile Durkheim that social facts cannot be reduced to the agents that are involved in them:[4]

---

[3] The sentence is often misquoted replacing the obsolete "publick" with the more modern "public".

[4] The article containing this quotation became the Préface of the second edition of *Les Règles de la méthode sociologique* [39], and is generally quoted as such (despite the article is antecedent); the sentence is not in the first, 1895, edition. In English:

> The solidity of bronze lies neither in the copper, nor in the tin, nor in the lead which have been used to form it, which are all soft or malleable bodies. The solidity arises from the mixing of the two. The liquidity of water, its nutritive and other properties, are not in the two gases of which it is composed, but in the complex substance they form by coming together. […Social facts] reside in the society itself that produces them and not in its parts, namely, its members. [40, pp. 39–40]

It is difficult to say whether Durkheim was aware of Huxley's example, but he was surely well acquainted with the work of Huxley's friend, Herbert Spencer (see [45]), on social organisms. By

> La dureté du bronze n'est ni dans le cuivre ni dans l'étain ni dans le plomb qui ont servi à le former et qui sont des corps mous ou flexibles; elle est dans leur mélange. La fluidité de l'eau, ses propriétés alimentaires et autres ne sont pas dans les deux gaz dont elle est composée, mais dans la substance complexe qu'ils forment par leur association. […Les faits sociaux] résident dans la société même qui les produit, et non dans ses parties, c'est-à-dire dans ses membres. [38, p. 9]

Several other historical examples are in [102, pp. 63–64, 863]; outside Biology, the economist Elinor Ostrom quoted this book [120, p. 44] as one of her major sources of inspiration.[5]

4 Discrete-Event Simulation

Discrete-Event Simulation (*DES*) is a kind of simulation developed in the 1950s that:

> utilizes a mathematical/logical model of a physical system that portrays state changes at precise points in simulated time. Both the nature of the state change and the time at which the change occurs mandate precise description. Customers waiting for service, the management of parts inventories, or military combat are typical application domains for discrete event simulation. [115, p. 370] or [114, p. 149]

Strictly speaking, the name DES denotes (almost) any simulation taking place in discrete time, but this has several consequences on how it is performed. The system has finitely many components, with finitely many states. These components interact through events having no duration. In general, the state of the system is described by a *state variable*.

Example 7 (*Queueing Systems*) The most classical example of a DES is a queue. As an example, individuals from a *calling population* arrive at random times in front of one or more *servers*, servicing them in *FIFO* (*first in first out*) order with random serving times; if all servers are busy, a *waiting line* creates and its length is the state variable. Several queueing systems are easily solvable, others are not, and require simulation to be solved. An example may be found in [6, Sect. 2.1]: the random values for the arrival and service times are collected in tables, and the number of customers in the system can be obtained combining these values through the rules of behavior of the queue.

---

the way, the metaphor of water is used in [45, p. 96] to describe the shorthand system developed by William George Spencer, Herbert Spencer's father.

[5] These pages by Mayr contain some mistakes. First, the book of Lloyd Morgan cited by Mayr is probably the one from 1923 [111], not from 1894, as emergence starts appearing in his work from 1912 (see [12, p. 59]). Second, the quotation just after that is not by Morgan but is taken from the book [121, p. 72] where it is used to illustrate the reasoning in [112, p. 59].

5 Microsimulation

5.1 Microsimulation in Economics

Microsimulation was introduced in 1957 by Guy Henderson Orcutt in [119]. Here is a definition adapted from the one provided by the International Microsimulation Association:

> Microsimulation refers to a wide variety of modeling techniques that operate at the level of individual units (such as persons, firms, or vehicles), with rules applied to simulate changes in state or behavior. These rules may be deterministic or stochastic, with the result being an estimate of the outcomes of applying these rules, possibly over many steps involving many interactions. These estimates are also at the micro level, allowing analysis of the distribution of the outcomes and changes to them, as well as the calculation of any relevant aggregate. [48, p. 2142]

Microsimulation bears resemblances with agent-based modelling (see Sect. 7) but they "have remained very distinct fields in the literature with microsimulation methods drawing heavily on micro-data" [48, p. 2142]. The reliance on micro-data for the construction of rules of behavior—that is generally considered a positive, when not the defining feature, of this method—has somewhat limited the scope of application of microsimulation to situations in which these data are available.

5.2 Intermezzo

What is missing from Microsimulation can be illustrated using the so-called *Coleman's boat* reproduced in Figs. 2 and 3.

The simplest form of this diagram (see [24, pp. 8 or 10]) is shown in Fig. 2. It illustrates the causal paths between micro- and macro-level phenomena: a macro-level cause influences agents at a micro-level and this in turn influences the macro-

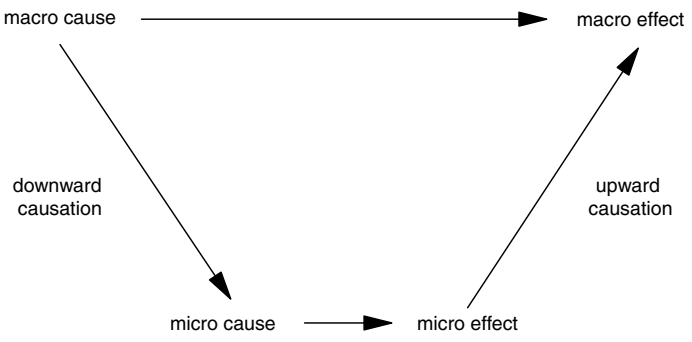


Fig. 2 Coleman's boat as represented in [24, pp. 8 or 10]

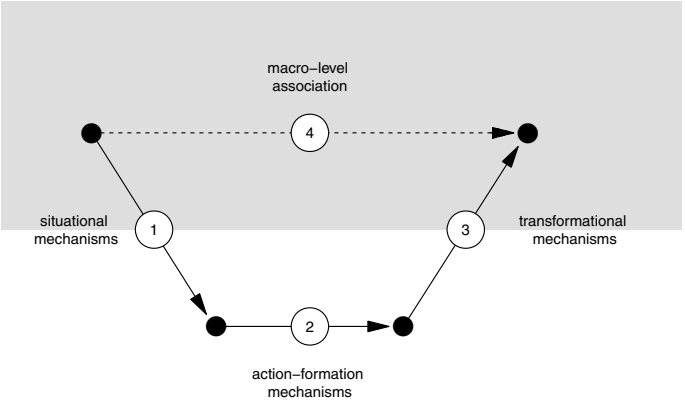


Fig. 3 Coleman's boat as represented in [65, p. 59]

level. Here, "the macro level is an abstraction, nevertheless an important one" [24, p. 12]. A slightly different representation, from [65, p. 59], is in Fig. 3. The grey area represents the macro level, while the white one represents the micro level. Applications to simulation methods are in [146, p. 35], [65, p. 59] and [66].

Before turning to the explanation of these graphs, we remark that some authors (see [18, p. 454] and [98, p. 42]) prefer to refer to this representation as a *Boudon-Coleman diagram* (see [126] for a study of the antecedents of Coleman's boat) while others present modifications of the boat without an upper macro-macro path [23, p. 1322].

The mechanisms "by which social structures constrain individuals' action and cultural environments shape their desires and beliefs" [65, p. 59] represented by path 1 are called *situational*. They represent *downward causation*. *Action-formation* mechanisms (path 2) "[link] individuals' desires, beliefs, etc., to their actions" [65, p. 59]. The mechanisms "by which individuals, through their actions and interactions, generate various intended and unintended social outcomes" [65, p. 59] are called *transformational* (path 3). They represent *upward causation*. Path 4 does not represent causality as "explanations that simply relate macro properties to each other […] are unsatisfactory" [65, p. 59]. One can iterate the "boat" over time: at each step of the simulation, there is an upward causation path from the agents towards the macro-level, and a downward causation path from the macro-level to the behavior of the agents.

The link of emergence with Coleman's boat is that emergent phenomena can generally be identified with macro behaviors induced by the mechanisms taking place in the bilges of the boat. Transformational mechanisms are especially relevant as they are the final step through which upward causation generates emergent phenomena. It is no surprise, therefore, that simulation methods built from the characteristics of the single agents may be better at modeling emergence (see [140, p. 230]).

Now, returning to Microsimulation, Economics has at least two mechanisms of individual market coordination that are coherent with both upward and downward causation: *general equilibrium* and *partial equilibrium*. The former takes place when equilibrium between demand and supply is achieved on all markets inside an economy at the same time, and changes in one market affect all other markets. The latter happens when one market is considered in isolation and is supposed not to affect the other markets. Both mechanisms predict that individuals, without coordination but only through *tâtonnement*, select prices achieving a macro-level equilibrium characterized by *market clearing*—i.e. full allocation of goods in any market of the economy. As individuals face these prices, this constitutes a source of downward causation, from the macro level to the micro one. But the application of partial and general equilibria in simulated models has two problems. First, it is not credible that these concepts of equilibrium hold exactly true, as perfect market clearing seems to be the exception rather than the norm. Second, in models representing a proper subset of the economy, it is difficult to imagine quantitative mechanisms of downward causation.

Coleman's boat can also be useful to classify simulation models. Indeed, some authors [58, 92] identify three categories of models:

- Macrosimulations (e.g., System Dynamics, see Sect. 3, Discrete-Event Simulation, see Sect. 4) focus on an aggregate level and operate at the level of the deck of Coleman's boat;
- Microsimulations (e.g., Microsimulation, see Sect. 5.1, Simulation of Voting Behavior, see Sect. 5.3) focus at the individual level and take place in the bilges of Coleman's boat;
- the third category is composed of models in which there is an iteration between the two levels. These are identified with so-called Agent-Based Models (see Sect. 7).

5.3 Early Simulation of Voting Behavior

Some models similar to economic Microsimulations can be found in the early literature on simulation of voting behavior. We include these models in this review because of their accent on agents' heterogeneity.

Pool and Abelson [122] presented a model they had developed for the Democratic Party in the 1960 US presidential campaign, the so-called Simulmatics project. The original model used the positions of 480 types of voters (that were consolidated to 15 in the published paper) on 52 issues. The data for the model were based on over 100,000 interviews in polls collected over 10 years by polling firms. The researchers advised Kennedy that he would benefit from taking a strong stance in favor of civil rights and from openly dealing with his Catholic religious beliefs. The paper was so influential that in 1964 Eugene Leonard Burdick, a political scientist and novelist, wrote a novel, called *The 480* (see [19]), criticizing the fact that the use of computer models made easy to choose strategies to maximize votes and manipulate electors.

In 1965, Ithiel de Sola Pool, Robert P. Abelson and Samuel L. Popkin published *Candidates, Issues, and Strategies: A computer simulation of the 1960 and 1964 presidential elections*, a book describing in detail their model (see [123]). [1] considered a simulation model of voting in the fluoridation referendum (i.e. whether tap water should be compulsorily fluoridated or not). The model had 500 agents behaving according to 51 rules (22 about information processing, 27 on information exchange, 2 for voting behavior).

6 Cellular Automata

Cellular automata (CA, sing. *automaton*) are systems composed of individuals taking on one of a discrete number of states, arranged in fixed cells (hence the name *cellular*) on a grid, interacting according to deterministic rules depending on the neighboring agents' state (hence the name *automata*). A more formal definition is the one in the Stanford Encyclopedia of Philosophy:

> CA are (typically) spatially and temporally discrete: they are composed of a finite or denumerable set of homogeneous, simple units, the atoms or cells. At each time unit, the cells instantiate one of a finite set of states. They evolve in parallel at discrete time steps, following state update functions or dynamical transition rules: the update of a cell state obtains by taking into account the states of cells in its local neighborhood (there are, therefore, no actions at a distance). [10]

They have been used both as specific examples of real-world phenomena and as abstract examples of how complex behavior can arise from simple rules. Note that the rules of behavior of cellular automata are deterministic and fixed. They were first formalized by Stanislaw Ulam and John von Neumann in the 40s, while the former was working on the growth of crystals and the latter on self-replicating systems. The work of von Neumann culminated in the classic [153] (note that the symposium for which the paper was written was held in 1948), while the work of Ulam was published in [148]. However, it was only in the 1970s that CA rose to prominence with the following example.

Example 8 (*Game of Life*) In 1970, in [56], Martin Gardner popularized "a fantastic solitaire pastime" invented by John Horton Conway. This is indeed a cellular automaton with very simple rules:

- each cell can be either occupied by a living creature or empty;
- the creature in the cell dies if it has 1 or 4+ neighbors (resp. of loneliness and overcrowding);
- an empty cell comes to life if it has 3 living neighbors.

The neighbors of a cell are the ones in the *Moore neighborhood*, i.e. the set of 8 cells in contact through a side or a corner with the cell (a *von Neumann neighborhood*, instead, is the set of 4 cells in contact through a side with the cell). The game is generally applied starting from a configuration of activated cells. At the beginning,

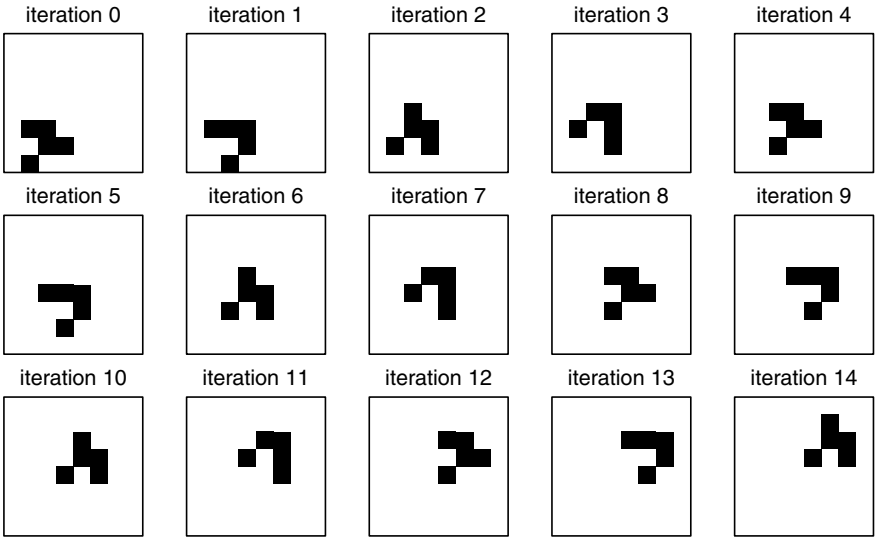


Fig. 4 Game of Life: evolution of a glider on a 8×8 checkerboard

Conway thought that such a system could not create a universe in constant expansion, but he was soon proved to be wrong (see the *glider* in [9, p. 931] or Fig. 4 and the *glider gun* in [9, p. 935]). The number of configurations that have been explored is incredibly large (see [9, Chap. 25]).

Around the same years, Thomas Crombie Schelling introduced a model dealing with *segregation*, i.e. the enforced separation of different ethnic groups in a community. By taking inspiration from James Sakoda, who created a set of so-called checkboard models (see [67] for the detailed story), *Schelling's Segregation Model* [131, 132] showed that a personal slight preference towards a less diverse neighborhood could create in the long run a segregated community.

Example 9 (*Schelling Segregation Model*) Schelling [131, 132] proposed a model in which two types of individuals, say, A and B, are located on a one-dimensional or two-dimensional grid. Some of the cells may be empty. At each step, each individual counts how many in their Moore neighborhood are like them: if the proportion is smaller than a threshold value x, they move to a new position. This position is chosen deterministically: it is the nearest empty position satisfying their threshold. Schelling did not quote explicitly cellular automata in his paper, but to keep the paper inside the framework was compelled to introduce awkward deterministic rules (as an example, individuals choose to move according to a certain order in the grid). However, this is not the version of Schelling's model that is generally used: in the latter, some randomness is generally introduced in the relocation of moving individuals. This small step gets this cellular automaton close to agent-based models (see Sect. 7).

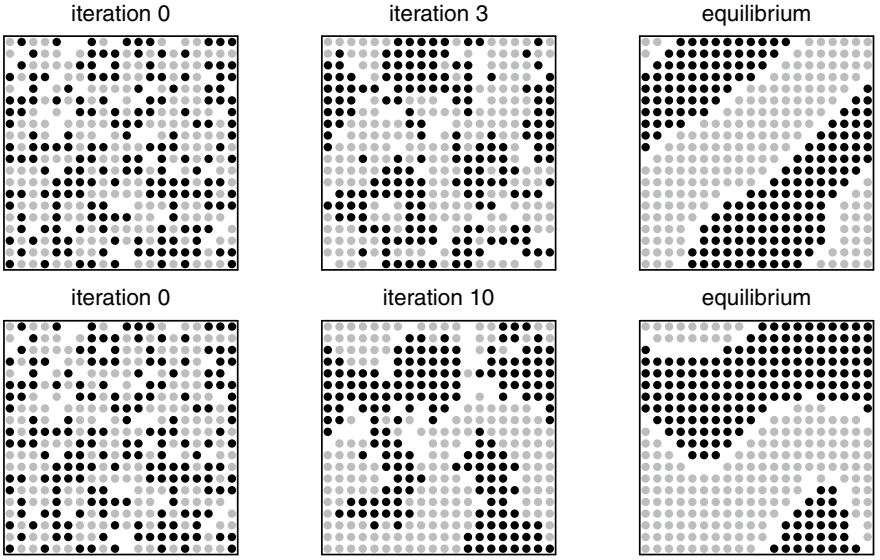


Fig. 5 Schelling Segregation Model for different values of x: the row above shows the case $x = 0.5$, the row below the case $x = 0.9$; the left column displays a random initial configuration that is equal for both values of x; the second column shows the model after 3 (above) and 10 (below) iterations; the last column shows what happens after an arbitrarily large number of iterations

This instance of the model is displayed in Fig. 5. The evolution of the system for different values of x has been characterized:

> Initially the system quickly develops small clusters, but then a slow evolution toward larger clusters follows. […T]he system evolves toward one big cluster or very few clusters. In the case of $x = 1/2$ the cluster surface tends to form flat surfaces […] In $x \neq 1/2$ cases the surface is bumpy and irregular [150, p. 19263].

6.1 Intermezzo

In this model, segregation is an example of *emergence*. We discuss in the following the implications of emergence and its importance for the model and for the development of simulation in the social sciences.

Example 10 (*Schelling Segregation Model*) In Schelling's model, segregation is an emergent property, as nobody necessarily wants it to take place. Schelling stated this opposition in the famous title of one of his book, *Micromotives and Macrobehavior* [133], whose blurb summarizes the idea as follows: "small and seemingly meaningless decisions and actions by individuals often lead to significant unintended consequences for a large group." The description in the blurb replicates the definition of the upward causation path in Coleman's boat.

Note the striking difference with respect to two economic frameworks that gained traction in the last decades.

On the one hand, in modern Macroeconomics, several models are based on a *representative agent*, i.e. an agent that represents the whole economy (see, e.g., [64]). This became a central element of *Real Business Cycle* (*RBC*) *theory*, first, and, after that, of *Dynamic Stochastic General Equilibrium Models* (*DSGE*) introduced in the two seminal papers [80, 88]. DSGE are, in general, macroeconomic models featuring an economy in general equilibrium; moreover, the models are microfounded, i.e. they are not formulated in terms of aggregate quantities but they derive their behavior by aggregating microeconomic individual models. In the case of DSGE, their behavior can be reduced to that of a representative agent maximizing expected utility. In [62], the authors provide an interesting point of view focused on the causal structure of these dynamic models:

> These types of models are nowadays the most widely used to draw and to evaluate policy claims because they bear the advantage of simultaneously addressing two critical issues about causal structures. On the one hand, under the acceptance of the rational expectation hypothesis, the structure modeled by the RBC/DSGE approach remains invariant under policy intervention because it takes into account the forward-looking behavior of the economic agents. On the other hand, the theoretical structure has an empirical counterpart in which the distinction between endogenous and exogenous variables is eschewed. [62, p. 126]

The emphasis on the representative agent implies that any characteristic of the economy is a characteristic of the agent, and no emergence seems possible (see [79, p. 51]).

On the other hand, in Microeconomics, as well as in other Natural and Social Sciences, some market and non-market interactions, in which tactical and strategic factors are preeminent, are studied through the lens of *game theory*, a branch of economics/mathematics introduced by Oskar Morgenstern and John von Neumann [155] in which agents interact taking into account other agents' reactions. Here emergence is possible as a consequence of strategic interaction between the agents.

These two situations describe a whole spectrum of models, from one in which no emergence is possible to one in which emergence is a consequence of strategic interaction. But Schelling's model is different as emergence is a consequence of the myopic behavior of individuals in a dynamic context. There is no planning at all. (One could even show that, if x is very high, no stable equilibrium is possible: if individuals have strong preferences against diversity, they do not get what they want!)

7 Agent-Based Models

An *Agent-Based Model* (henceforth *ABM*) is a computational model whose unit is the *agent*, an autonomous individual behaving in a given *environment* according to established *rules* [134]. The agent is the unit of analysis and it can be anything the

modeler is interested in, from a neutron to a country. Its general features can be characterized by:

- *autonomy*: each agent is modeled independently from the others and it can develop in ways that are not predictable solely by looking at the initial conditions set;
- *interaction*: exchanges with other agents may modify the characteristics of the agent and the way in which it perceives the environment and the self;
- *complexity*: some characteristics "emerge" during the interactions.[6]

Agents interact in a limited space that is usually referred to as *environment*, such that their position can represent either their physical location (see Schelling's Segregation Model in Sect. 6) or their psychological state of mind (see the Garbage Can Model below).

The *rules* are the norms that regulate what happens in the model and are sometimes identified as mechanisms. They can be:

- *behavioral*: they define what each agent should be doing in general and/or as a function of their characteristics;
- *interactional*: they define what happens to an agent and/or to the environment when they interact;
- *time-dependent*: rules may modify agents' characteristics, other rules, or the shape of the environment as time—however defined in the simulation—goes by;
- *developmental*: rules set the conditions for agents (and/or the environment) to change, evolve or die.

The next step we deem appropriate at this point in the chapter is to try and explain the difference between ABM and other simulation frameworks, especially because ABM is the latest and most advanced of all known techniques so far. A classification of simulation models can be based on the following dichotomies:

- the backings of the model can be based on equations or on properties of the objects;
- the approach can be either at the macro- or at the micro-level;
- agents can be homogeneous or heterogeneous;
- rules can be homogeneous or heterogeneous;
- the environment can be either static or dynamic.

While the other concepts have already been explained, it can be interesting to spend a word on backings. Most simulation models involve behaviors that are dictated by equations that connect the different elements of the model. This is clearly true for simulation models working at the aggregate level, but also some microsimulation models are based on actions described through the application of equations on individual-level variables. Some models, however, start from the specification of the characteristics of the objects—be they agents and/or rules—and let them interact more or less freely. A first consideration is that this further increases the distance between the observed behavior of the components of the model and the elements

[6] It is a bit odd to attach this aspect to agents, but we want to highlight that agents can be characterized as complex; see below and [42].

Table 1 A comparison of computational simulation models

	Backings		Approach		Agents		Rules		Environments	
	Equation	Object	Macro	Micro	Homogeneous	Heterogeneous	Homogeneous	Heterogeneous	Static	Dynamic
Early comp.	X		X	X	X		X		X	
SD	X		X			X	X		X	
DES	X		X		X		X		X	
Microsim.	X			X	X		X			X
Voting models	X			X	X		X		X	
CA	X	X		X	X		X		X	
ABM		X	X	X		X		X		X

governing it. Indeed, while the result of an equation is often rather predictable, it is not the case for the interactions of objects possessing their own characteristics and behaving on their basis. For this reason, models whose backings are based on objects often offer more opportunities for the development of those features that are not possessed by their own components but that are born out of the interactions—i.e. exactly emergent properties. This is not to say that equation-backed models cannot exhibit emergence—as the examples above show, they can and they often do—but only that emergence is usually more unpredictable in object-backed models. A second, related, point is that object-backed models are often built using a *bottom-up* approach—i.e. starting from the properties of the objects—and they recover aggregate features only as a result of the interactions. From this point of view, they start from the bilges of Coleman's boat (see Figs. 2 and 3) but they also involve its deck.[7]

From Table 1, it is probably more apparent to understand why ABM is considered the most advanced computational simulation approach as of yet. In fact, by comparing core components of the simulation approaches reviewed so far, it becomes clear how ABM stands at odds with most of them. Of course, the agent-based approach has taken from past simulation techniques, but its comprehensive reach makes it stand as a jump ahead. Omitted from the table and not explicitly mentioned (only cursorily in the introduction) in this chapter so far are the surrounding conditions that make ABM a viable option. We are referring to the surge of computational power and to the possibility that even home computers are capable of performing complicated operations, unthinkable twenty years ago. This means that, for example, having heterogeneous agents in a simulation came at very high costs before the middle of the 1990s, while it is relatively (computationally) cheap to allow them today. This technical hardware and software innovation opened up for the possibility of a different approach to simulation modeling.

In the following pages, we introduce an example concerning the ABM version of the celebrated Garbage Can Model.

Example 11 (*Garbage Can Model*) The *Garbage Can Model* (GCM) of [21] is a model of decision making in an *organized anarchy*, i.e. an organization characterized by the three properties of problematic preferences, unclear technology, and fluid participation.

> [A]n organization is a collection of choices looking for problems, issues and feelings looking for decision situations in which they might be aired, solutions looking for issues to which they might be the answer, and decision makers looking for work. [21, p. 2]

The GCM has been extremely influential in organizational behavior. The model was implemented using one of the earliest computer languages developed by IBM, FORTRAN, and it was not an ABM. Yet, by using an ABM jargon, four types of agents can be identified: (a) problems, (b) opportunities, (c) solutions, and (d)

[7] It is worth noting that ABM can also be backed by equations, better, by a mix of equations and object-based modeling. Actually, we are not aware of ABM that do not have any equation embedded in their coding. The difference of this approach is in the ability to mix and mash both object- and equation-based techniques.

decision makers. The overall goal of the model is to determine whether a formal (hierarchic) organizational structure provides an institutional backbone for problem solving that is better than an informal (anarchic) organizational structure, or not. In the first case, the four types of agent interact following a specified sequence while in the other they interact at random. There are two ways in which participants make decisions in the organization. One is *by resolution*: it happens when problems are solved once participants match opportunities to the right solutions; i.e. when the right combination of the four agents are on the same patch at the same time. The other is *by oversight*: it is when solutions and opportunities are available to participants but no problems are actually solved.

Not all problems are solved automatically, just by having opportunities, decision makers, and solutions available. In fact, all problems have difficulty levels, participants have abilities, and solutions have a certain degree of efficiency. The problem is solved if the match of the participant with an opportunity and a solution overcomes the difficulty of the problem.

The findings of the original model are the following: resolution is not the most common style of decision making; hierarchies reduce the number of unresolved problems but increase problem latency; important and early problems are more likely to be solved.

The model has later been implemented by [51, 52] in ABM form. In the agent-based version of the model, there are three types of structures:

1. *Anarchy*. There is no hierarchy so that abilities, efficiencies, and difficulties are randomly distributed among agents.
2. *Hierarchy-competence*. Abilities, efficiencies, and difficulties increase as one moves up the hierarchical ladder.
3. *Hierarchy-incompetence*. Abilities, efficiencies, and difficulties decrease as one moves up the hierarchical ladder.

Finally, the model implements two modes of (not) dealing with problems (i.e. *flights*):

1. *buck passing*: when one participant has the alternative of passing the decision on a problem to another participant;
2. *postpone*: when problems are kept on hold by participants and eventually solved at an unspecified future time.

The results of the model are summarized as follows by the authors:

> The […] properties point to very interesting features of organizational decision-making. […]
>
> 1. Decisions by oversight are very common, much more common than decisions made in order to solve problems. This result suggests that the rational mode of decision-making is a very rare case. Most decisions are socially induced acts, made with the purpose of obtaining legitimacy by conforming to required rituals.
> 2. If there is a hierarchy, then top executives are busy with gaining legitimacy for their organization by means of decisions by oversight, whereas the bottom line cares about solving problems.

3. Organizations make themselves busy with a few problems that present themselves again and again. So participants have the impression of always facing the same problems. [52, p. 123]

7.1 A Menagerie of Names

ABM have been used in several disciplines: Biology/Ecology [59, 60, 125], Computer Science [37, 117], Sociology [4, 92, 143, 147], Management [2, 110] and Organizational Behavior/Organization Science [7, 49, 136], Political Science [32], Psychology [25, 26] and Cognitive Sciences [93], Population Studies [61], Economics [5, 145] and Finance [157], Transportation Research [94].

As expected, their very general nature implies that ABM can be adjusted in several ways according to the discipline. The different definitions of ABM are exposed in the following parts of the chapter. Since Social Sciences include many different disciplines, we decided to restrict our attention only to a part of these and to compare them with a selection of other disciplines, from the area of Natural Sciences and Techniques (e.g., Physics, Engineering, and Biology). In particular, we dedicate special attention to Economics and Management, highlighting the peculiarities of agents as heterogeneous individuals, with bounded rationality, interacting with each other.

7.1.1 Biology/Ecology

In Biology, an ABM is sometimes called an *Individual-Based Model* (*IBM*), but there is considerable confusion as to the definition. Some authors consider individual-based and agent-based as synonyms (see, e.g., [125, p. 3]). Others reserve the term individual-based when individuals are simpler and rules are formulated probabilistically at the individual level (see, e.g., [11, p. 338]).

Example 12 (*Conservation Biology*) In this branch of biology, and especially in *population viability analysis* (*PVA*)—i.e. the quantitative assessment of risk of extinction concerning a population or a species—most classical models fail to take into account the spatial nature of habitats: spatially dispersed animals, like the giraffe (see [28, 72]), may be at higher risk than commonly thought. *Metapopulation models* were the first to consider several populations interacting in separated locations, but they generally constrain the locations to be discrete. ABM offer a continuous improvement.

7.1.2 Computer Science

In Computer Science, one finds the concept of *Multi-Agent System* (*MAS*, see [37] for a survey and [117] for the relation between MAS and ABM), used to denote a

computer system of intelligent agents in interaction inside an environment. The aim is not to study the behavior of agents, but to solve a problem or compute a quantity. An agent is seen here only as a computing entity (and this is why, at odds with most ABM, it must be intelligent). As agents share knowledge or computing power, they approach the solution of the problem. From this point of view, MAS can be seen as a subfield of *Distributed Artificial Intelligence* (*DAI*):

> DAI is the study, construction and application of multiagent systems, that is, systems in which several interacting intelligent agents pursue some set of goals or perform some set of tasks. [158, p. 1]

Another related, but slightly more general concept, is that of *Artificial Adaptive Agents* (*AAA*). This name is often connected with so-called *Complex Adaptive Systems* (*CAS*, e.g., [108]):

> Such a system is *complex* in a special sense: (i) It consists of a network of interacting agents (processes, elements); (ii) it exhibits a dynamic, aggregate behavior that emerges from the individual activities of the agents; and (iii) its aggregate behavior can be described without a detailed knowledge of the behavior of the individual agents. An agent in such a system is *adaptive* if it satisfies an additional pair of criteria: the actions of the agent in its environment can be assigned a value (performance, utility, payoff, fitness, or the like); and the agent behaves so as to increase this value over time. A complex adaptive system, then, is a complex system containing adaptive agents, networked so that the environment of each adaptive agent includes other agents in the system. [68, p. 365; emphasis in the original]

What differentiates CAS from ABM is the emphasis, that is generally lacking in the latter, on adaptivity, but the two approaches are not mutually exclusive (see [118] for an example).

7.1.3 Physics

Physics does not generally require the introduction of agents as sentient and autonomous entities. Optimization provides an example of the difference between Social Sciences and Physics:

> In physics and the natural sciences, maximization typically occurs without a deliberate "maximizer." […M]aximizing behavior differs from nonvolitional maximization because of the fundamental relevance of the choice act. Fermat's "principle of least time" in optics was a fine minimization exercise (and correspondingly, one of maximization). It was not, however, a case of maximizing behavior, since no volitional choice is involved (we presume) in the use of the minimal-time path by light. [139, p. 745]

A notable exception is the area of complex systems, that is at the border of Physics. Outside this area, the generic name of Monte Carlo is used for models that would elsewhere give rise to ABM.

Example 13 (*Why the Brazil Nuts are on Top*) According to physical intuition, when we shake a box of nuts, the largest nuts should go on the bottom while the smallest should float; the evidence suggests that the contrary is true. In [128, 129] the authors build a Monte Carlo model to show how and when this happens.

7.1.4 Economics

In Economics, the most appreciated features of ABM are that they allow for heterogeneity of the agents, both in their types and their characteristics, for interactions taking place in non-trivial, often dynamic, networks [57], and for a wide range of individual behaviors, from perfect foresight to bounded rationality [124].

The name *Agent-based Computational Economics* (*ACE*) is sometimes used to denote the computational study of economic processes modeled as dynamic systems of interacting agents. According to [85, p. 246], an agent refers, in general, to "an encapsulated collection of data and methods representing an entity residing in a computationally constructed world."

ACE may overcome some critiques moved to DSGE models in macroeconomics (see [20, 75, 76]):[8]

> The advantage of the ACE approach for macroeconomics in particular is that it removes the tractability limitations that so limit analytic macroeconomics. ACE modeling allows researchers to choose a form of microeconomics appropriate for the issues at hand, including breadth of agent types, number of agents of each type, and nested hierarchical arrangements of agents. It also allows researchers to consider the interactions among agents simultaneously with agent decisions, and to study the dynamic macro interplay among agents. Researchers can relatively easily develop ACE models with large numbers of heterogeneous agents, and no equilibrium conditions have to be imposed. Multiple equilibria can be considered, since equilibrium is a potential outcome rather than an imposed requirement. Stability and robustness analysis can be done simultaneously with analysis of solutions. [22]

A second difference between ACE and DSGE concerns the agent's expectations, that in DSGE are generally *rational*. This is a mechanism of expectation formation introduced in [113], according to which ex-ante expectations concerning the future value of a variable differ from its real value by a zero-mean random term. This can be justified supposing that agents know the model representing the economy, from which the alternative name of *model-consistent expectations*, i.e. individuals and researchers share the same model of the economy: "Muth's notion was that the professors, even if correct in their model of man, could do no better in predicting than could the hog farmer or steelmaker or insurance company. […] The professors declare themselves willing to attribute to economic actors at least as much common sense as is embodied in professional theories" [103, p. 53]. Now, this kind of *model-consistent rationality*, according to which agents are able to analyze the economy as economists, cannot be generally assumed in ACE (and more generally in ABM) because of the very way in which models are built.

ACE are widely used in Finance, although, as pointed out by [35, 69, 70, 84], in Financial Economics several features of ABM are not used (interactions taking place over networks, coexistence of several kinds of agents, etc.) and the attention is more focused on the heterogeneity and bounded rationality of consumers.

Another class of models is formed by *history-friendly models* (*HFM*, see [96, 97]):

[8] We refer to [46, 47, 63, 77, 86] for a detailed discussion on the differences between ACE and DSGE.

> "[H]istory-friendly" models (HFMs) of industrial evolution […] are variants of ABMs which aim at capturing in stylized form qualitative theories about mechanisms and factors affecting industrial evolution, technological advances and institutional changes. HFMs consist of three steps: appreciative theories of the history of a specific industry, history-replicating simulations, and history-divergent simulations. In HFMs, model building and calibration are conducted with the guidance of the history. [160, p. 45]

As explained above, an important step in HFM is *calibration*, namely the search for values of the parameters of the model producing an output that is approximately similar to a set of real data.

7.1.5 Management and Organization Research

The status of ABM among the management disciplines is still controversial. In fact, and in spite of the few attempts made so far [49, 109, 134, 137], there is still no clear definition of what ABM are for scholars in the area of Management and Organizational Research (MOR). If one excludes the more engineering-related area of MOR, that is Operations and Supply Chain Management, there are very few examples of ABM (a recent publication guides the reader on how to develop ABM in this field of research; [135]).

In a review of the literature, Wall [156] divides the contributions in two groups, those related to exploration/exploitation and those dealing with differentiation and integration. Models pertaining to the former originate from James G. March's famous categorization of possible opposite decisions an organization usually faces [99] between—to make a very long story short—putting existing resources to work (exploitation) or seeking additional resources (exploration). The latter dichotomy is engrained into very old discourses within the MOR literature, and relates to the basic decision to "make or buy" [82]. One of the most interesting findings of the review—although not discussed openly—is that almost all models are of a special kind: they are NK models (see below). According to another recent study [8], it seems there is a trend in MOR where scholars engage in one of the simplest kinds of ABM while (a) not calling them as such, and (b) de facto establishing a parallel literature. Using NK models to address MOR topics is probably a sign that the field is struggling with concepts such as complexity, emergence, and randomness.

But, outside of this quite remarkable trend, there still are areas of MOR where ABM have appeared. This is the case, for example, of those ABM dealing with routines (e.g., [14, 110]), work team dynamics (e.g., [7, 101]), and, more broadly, with organizational behavior [136].

Example 14 (*Adaptation on a Rugged Landscape*) Given their prominence in MOR, it makes sense to introduce NK models in an example. Before shortly describing the example, it is worth dedicating a few words on this typology of models. They were introduced by evolutionary biologist Stuart A. Kauffman in the late 1980s [73, 74] to study fitness and adaptation. Using Kauffman and Weinberger's words:

> The distribution of the fitness values over the space of genotypes constitutes the fitness landscape. […] The space consists of all 20 N proteins, length N, arranged such that each

> protein is a vertex next to all 19 N single mutant variants obtained by replacing one amino acid at one position by one of the 19 remaining possible coded amino acids. Each protein in the space is assigned some "fitness" with respect to a specific property, such as binding a specific ligand, where "fitness" can be defined as the affinity of binding. An adaptive walk can be conceived as a process which begins at a single protein in the space and passes via ever fitter 1-mutant variants. […] N is the number of "sites" in the model genotype or protein, while K is the number of sites whose alternative states, "alleles" or amino acids, bear on the fitness contribution of each site. Thus K measures the richness of epistatic interactions among sites. [74, pp. 211–212]

Outside of Biology, a modeler could attribute a diversity of characteristics to the two main parameters N and K, preserving their relations, and adapting to the study of different types of "fitness". This was the intuition of Daniel A. Levinthal [87] who was probably the first to introduce the MOR community to NK models. He studied how organizations adapt to different forms (organizational design). In the model, there are N organizational attributes and K other attributes that affect an organization's fitness—i.e. the extent to which interaction affects adaptation. Dependence on initial conditions was one of the main findings of this simulation model.

8 Conclusions

The aim of the chapter was to show the central role played by randomness, emergence and causation for the development of different groups of simulation models. Following the literature, we have outlined a short and necessarily partial history of simulation models, with special attention to the Social Sciences. The models that we have covered are some early works with analog and digital computers, System Dynamics, Discrete-Event Simulation, Microsimulation in Economics and Political Science, Cellular Automata and Agent-Based Models.

To conclude, ABM can be considered the most advanced computational simulation approach so far. Indeed, although this approach has taken from past simulation techniques, its comprehensive reach makes ABM stand as a jump ahead.

Acknowledgements The second author gratefully acknowledges financial and logistic support from the DiECO, Università degli Studi dell'Insubria, during a visiting period. The third author acknowledges the PRIN Grant 2017 "How Good Is Your Model? Empirical Evaluation and Validation of Quantitative Models in Economics." We thank Eugenio Caverzasi, Ernesto Carrella, Alan Kirman, Alessio Moneta, and Massimo Rusconi for useful feedback on the paper.

References

1. Abelson, R.P., Bernstein, A.: A computer simulation model of community referendum controversies. Public Opin. Q. **27**(1), 93 (1963)
2. Aggarwal, V.A., Siggelkow, N., Singh, H.: Governing collaborative activity: interdependence and the impact of coordination and exploration. Strateg. Manag. J. **32**(7), 705–730 (2011)

3. Anderson, P.W.: More is different. Science **177**(4047), 393–396 (1972)
4. Axelrod, R.: The dissemination of culture: a model with local convergence and global polarization. J. Conflict Resolut. **41**(2), 203–226 (1997)
5. Axelrod, R., Tesfatsion, L.: Appendix A: a guide for newcomers to agent-based modeling in the social sciences. In: Tesfatsion, L., Judd, K.L. (eds.) Handbook of Computational Economics, vol. 2, pp. 1647–1659. Elsevier (2006)
6. Banks, J., Carson, J.S. II, Nelson, B.L., Nicol, D.M.: Discrete-Event System Simulation, 4th edn. Prentice-Hall International Series in Industrial and Systems Engineering. Pearson Prentice Hall, Upper Saddle River, NJ (2005)
7. Bardone, E., Secchi, D.: Inquisitiveness: distributing rational thinking. Team Perform. Manag. Int. J. **23**(1/2), 66–81 (2017)
8. Baumann, O., Schmidt, J., Stieglitz, N.: Effective search in rugged performance landscapes: a review and outlook. J. Manag. **45**(1), 285–318 (2019)
9. Berlekamp, E.R., Conway, J.H., Guy, R.K.: Winning Ways for Your Mathematical Plays, 2nd edn. A.K. Peters, Natick, MA (2001)
10. Berto, F., Tagliabue, J.: Cellular automata. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, Fall (2017)
11. Black, A.J., McKane, A.J.: Stochastic formulation of ecological models and their applications. Trends Ecol. Evol. **27**(6), 337–345 (2012)
12. Blitz, D.: Emergent Evolution: Qualitative Novelty and the Levels of Reality. Springer, Dordrecht (2010)
13. Bradbury, R.: A sound of thunder. Collier's **28**, 20–21, 60–61 (1952)
14. Breslin, D., Romano, D., Percival, J.: Conceptualizing and modeling multi-level organizational co-evolution. In: Secchi, D., Neumann, M. (eds.) Agent-Based Simulation of Organizational Behavior. New Frontiers of Social Science Research, pp. 137–157. Springer International Publishing, Cham (2016)
15. Brown, R.G.: Dieharder: A Random Number Test Suite (2019)
16. Buffon, G.-L.L.: Geometrie [Résolution des problémes qui regardent le jeu du franc-carreau]. Histoire de l'Académie Royale des Sciences, Année **1733**, 43–45 (1735)
17. Buffon, G.-L.L.: Essais d'Arithmétique morale. In: Histoire Naturelle, Générale et Particulière, Supplément, Tome Quatrième, pp. 46–123. Imprimerie Royale, Paris (1777)
18. Bunge, M.: Mechanism and explanation. Philos. Soc. Sci. **27**(4), 410–465 (1997)
19. Burdick, E.: The 480. McGraw Hill, New York, NY (1964)
20. Caiani, A., Godin, A., Caverzasi, E., Gallegati, M., Kinsella, S., Stiglitz, J.E.: Agent based-stock flow consistent macroeconomics: towards a benchmark model. J. Econ. Dyn. Control **69**, 375–408 (2016)
21. Cohen, M.D., March, J.G., Olsen, J.P.: A garbage can model of organizational choice. Adm. Sci. Q. **17**(1), 1 (1972)
22. Colander, D., Howitt, P., Kirman, A., Leijonhufvud, A., Mehrling, P.: Beyond DSGE models: toward an empirically based macroeconomics. Am. Econ. Rev. **98**(2), 236–240 (2008)
23. Coleman, J.S.: Social theory, social research, and a theory of action. Am. J. Sociol. **91**(6), 1309–1335 (1986)
24. Coleman, J.S.: Foundations of Social Theory. Belknap Press of Harvard University Press, Cambridge, MA (1990)
25. Conte, R.: Agent-based modeling for understanding social intelligence. Proc. Natl. Acad. Sci. **99**(suppl 3), 7189–7190 (2002)
26. Conte, R., Paolucci, M.: On agent-based modeling and computational social science. Front. Psychol. **5** (2014)
27. RAND Corporation (ed.): A Million Random Digits with 100,000 Normal Deviates. Free Press, Glencoe, IL (1955)
28. Courchamp, F., Jaric, I., Albert, C., Meinard, Y., Ripple, W.J., Chapron, G.: The paradoxical extinction of the most charismatic animals. PLOS Biol. **16**(4), e2003997 (2018)
29. Cunningham, B.: The reemergence of 'emergence'. Philos. Sci. **68**, S62–S75 (2001)
30. de Laplace, P.-S.: Théorie analytique des probabilités. Veuve Courcier, Paris (1812)

31. de Laplace, P.-S.: Essai philosophique sur les probabilités. Veuve Courcier, Paris (1814)
32. de Marchi, S., Page, S.E.: Agent-based models. Annu. Rev. Polit. Sci. **17**(1), 1–20 (2014)
33. De Morgan, A.: Supplement to the budget of paradoxes (No. IV). The Athenæum **2017**, 835–836 (1866)
34. De Morgan, A.: A Budget of Paradoxes. Longmans, Green, and Co., London (1872)
35. Dieci, R., He, X.-Z.: Chapter 5: Heterogeneous agent models in finance. In: Hommes, C., LeBaron, B. (eds.) Handbook of Computational Economics, vol. 4, pp. 257–328. Elsevier (2018)
36. Doore, K., Fishwick, P.: Prototyping an analog computing representation of predator prey dynamics. In: Proceedings of the Winter Simulation Conference 2014, pp. 3561–3571a, Savannah, GA, December 2014. IEEE
37. Dorri, A., Kanhere, S.S., Jurdak, R.: Multi-agent systems: a survey. IEEE Access **6**, 28573–28593 (2018)
38. Durkheim, É.: De la Méthode Objective en Sociologie. Revue de Synthèse Historique **II**(1), 3–17 (1901)
39. Durkheim, É.: Les Règles de la méthode sociologique, revue et augmentée d'une préface nouvelle, 2nd edn. Bibliothèque de philosophie contemporaine, Alcan, Paris (1901)
40. Durkheim, É.: The Rules of Sociological Method. The Free Press, New York, NY (1982)
41. Eckhardt, R.: Stan Ulam, John von Neumann, and the Monte Carlo method. Los Alamos Sci. **15**(Special Issue), 131–137 (1987)
42. Edmonds, B., Moss, S.: From KISS to KIDS—an 'Anti-simplistic' modelling approach. In: Davidsson, P., Logan, B., Takadama, K. (eds.) Multi-Agent and Multi-Agent-Based Simulation, vol. 3415, pp. 130–144. Springer-Verlag, Berlin, Heidelberg (2005)
43. Ekeland, I.: Au hasard: la chance, la science et le monde. Seuil, Paris (1991)
44. Ekeland, I.: The Broken Dice, and Other Mathematical Tales of Chance. University of Chicago Press, Chicago, IL (1993)
45. Elwick, J.: Containing multitudes: Herbert Spencer, organisms social and orders of individuality. In: Francis M., Taylor, M.W. (eds.) Herbert Spencer: Legacies, pp. 89–110. Routledge, London (2014)
46. Fagiolo, G., Roventini, A.: Macroeconomic policy in DSGE and agent-based models Redux: new developments and challenges ahead. J. Artif. Soc. Soc. Simul. **20**(1), 1 (2017)
47. Farmer, J.D., Foley, D.: The economy needs agent-based modelling. Nature **460**(7256), 685–686 (2009)
48. Figari, F., Paulus, A., Sutherland, H.: Chapter 24: Microsimulation and policy analysis. In: Atkinson, A.B., Bourguignon, F. (eds.) Handbook of Income Distribution, vol. 2B, pp. 2141–2221. Elsevier (2015)
49. Fioretti, G.: Agent-based simulation models in organization science. Organ. Res. Methods **16**(2), 227–242 (2013)
50. Fioretti, G.: Emergent organizations. In: Secchi, D., Neumann, M. (eds.) Agent-Based Simulation of Organizational Behavior. New Frontiers of Social Science Research, pp. 19–41. Springer International Publishing, Cham (2016)
51. Fioretti, G., Lomi, A.: An agent-based representation of the garbage can model of organizational choice. J. Artif. Soc. Soc. Simul. **11**(1), 1 (2008)
52. Fioretti, G., Lomi, A.: Passing the buck in the garbage can model of organizational choice. Comput. Math. Organ. Theory **16**(2), 113–143 (2010)
53. Forrester, J.W.: Principles of Systems. MIT Press, Cambridge, MA (1968)
54. Forrester, J.W.: Counterintuitive behavior of social systems. Technol. Forecast. Soc. Chang. **3**, 1–22 (1971)
55. Forrester, J.W.: The beginning of system dynamics. McKinsey Q. **1995**(4), 4–16 (1995)
56. Gardner, M.: The fantastic combinations of John Conway's new solitaire game "life". Sci. Am. **223**(4), 120–123 (1970)
57. Gilbert, N., Pyka, A., Ahrweiler, P.: Innovation networks - a simulation approach. J. Artif. Soc. Soc. Simul. **4**(3), 8 (2001)

58. Gilbert, N., Troitzsch, K.G.: Simulation for the Social Scientist, 2nd edn. Open University Press, Maidenhead, New York, NY (2005)
59. Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S.K., Huse, G., Huth, A., Jepsen, J.U., Jørgensen, C., Mooij, W.M., Müller, B., Pe'er, G., Piou, C., Railsback, S.F., Robbins, A.M., Robbins, M.M., Rossmanith, E., Rüger, N., Strand, E., Souissi, S., Stillman, R.A., Visser, R.V.U., DeAngelis, D.L.: A standard protocol for describing individual-based and agent-based models. Ecol. Model. **198**(1–2), 115–126 (2006)
60. Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W.M., Railsback, S.F., Thulke, H.-H., Weiner, J., Wiegand, T., DeAngelis, D.L.: Pattern-oriented modeling of agent-based complex systems: lessons from ecology. Science **310**(5750), 987–991 (2005)
61. Grow, A., Van Bavel, J. (eds.): Agent-Based Modelling in Population Studies. The Springer Series on Demographic Methods and Population Analysis, vol. 41. Springer International Publishing, Cham (2017)
62. Guerini, M., Moneta, A.: A method for agent-based models validation. J. Econ. Dyn. Control **82**, 125–141 (2017)
63. Guerini, M., Napoletano, M., Roventini, A.: No man is an Island: the impact of heterogeneity and local interactions on macroeconomic dynamics. Econ. Model. **68**, 82–95 (2018)
64. Hands, D.W.: Conundrums of the representative agent. Camb. J. Econ. **41**(6), 1685–1704 (2017)
65. Hedström, P., Ylikoski, P.: Causal mechanisms in the social sciences. Ann. Rev. Sociol. **36**(1), 49–67 (2010)
66. Hedström, P., Ylikoski, P.: Analytical sociology and social mechanisms. In: Kaldis, B. (ed.) Encyclopedia of Philosophy and the Social Sciences, pp. 26–29. SAGE Publications, Cham (2013)
67. Hegselmann, R., Schelling, T.C., Sakoda, J.M.: The intellectual, technical, and social history of a model. J. Artif. Soc. Soc. Simul. **20**(3) (2017)
68. Holland, J.H., Miller, J.H.: Artificial adaptive agents in economic theory. Am. Econ. Rev. **81**(2), 365–370 (1991)
69. Hommes, C.H.: Chapter 23: Heterogeneous agent models in economics and finance. In: Tesfatsion, L., Judd, K.L. (eds.) Handbook of Computational Economics, vol. 2, pp. 1109–1186. Elsevier (2006)
70. Hommes, C.H., Wagener, F.: Chapter 4: Complex evolutionary systems in behavioral finance. In: Hens, T., Schenk-Hoppe, K. (eds.) Handbook of Financial Markets: Dynamics and Evolution, pp. 217–276. Elsevier (2009)
71. Huxley, T.H.: On the Physical Basis of Life. The College Courant, New Haven, CT (1869)
72. IUCN: Giraffa camelopardalis (amended version of 2016 assessment). In: The IUCN Red List of Threatened Species 2018: E.T9194A136266699. International Union for Conservation of Nature (2018)
73. Kauffman, S.A.: Cambrian explosion and Permian quiescence: implications of rugged fitness landscapes. Evol. Ecol. **3**(3), 274–281 (1989)
74. Kauffman, S.A., Weinberger, E.D.: The NK model of rugged fitness landscapes and its application to maturation of the immune response. J. Theor. Biol. **141**(2), 211–245 (1989)
75. Kirman, A.P.: The intrinsic limits of modern economic theory: the emperor has no clothes. Econ. J. **99**(395), 126 (1989)
76. Kirman, A.P.: Whom or what does the representative individual represent? J. Econ. Perspect. **6**(2), 117–136 (1992)
77. Kirman, A.P.: Ants and nonoptimal self-organization: lessons for macroeconomics. Macroecon. Dyn. **20**(2), 601–621 (2016)
78. Knudsen, T., Levinthal, D.A., Puranam, P.: Editorial: a model is a model. Strat. Sci. **4**(1) (2019)
79. Kuperberg, M.: The two faces of emergence in economics. Sound. Interdiscip. J. **90**(1/2), 49–63 (2007)

80. Kydland, F.E., Prescott, E.C.: Time to build and aggregate fluctuations. Econometrica **50**(6), 1345 (1982)
81. Lane, D.C.: The power of the bond between cause and effect: Jay Wright Forrester and the field of system dynamics. Syst. Dyn. Rev. **23**(2–3), 95–118 (2007)
82. Lawrence, P.R., Lorsch, J.W.: Differentiation and integration in complex organizations. Adm. Sci. Q. **12**(1), 1–47 (1967)
83. Lazzarini M.: Un'applicazione del calcolo della probabilità alla ricerca sperimentale di un valore approssimato di π. Periodico di Matematica per l'insegnamento secondario **IV**(II), 140–143 (1901)
84. LeBaron, B.: Chapter 24: Agent-based computational finance. In: Tesfatsion, L., Judd, K.L. (eds.) Handbook of Computational Economics, vol. 2, pp. 1187–1233. Elsevier (2006)
85. LeBaron, B., Tesfatsion, L.: Modeling macroeconomies as open-ended dynamic systems of interacting agents. Am. Econ. Rev. **98**(2), 246–250 (2008)
86. Lengnick, M., Wohltmann, H.-W.: Agent-based financial markets and New Keynesian macroeconomics: a synthesis. J. Econ. Interac. Coord. **8**(1), 1–32 (2013)
87. Levinthal, D.A.: Adaptation on rugged landscapes. Manag. Sci. **43**, 934–950 (1997)
88. Long, J.B., Plosser, C.I.: Real business cycles. J. Polit. Econ. **91**(1), 39–69 (1983)
89. Lorenz, E.N.: Deterministic nonperiodic flow. J. Atmos. Sci. **20**(2), 130–141 (1963)
90. Lotka, A.J.: Analytical note on certain rhythmic relations in organic systems. Proc. Natl. Acad. Sci. **6**(7), 410–415 (1920)
91. Lotka, A.J.: Elements of Physical Biology. Williams & Wilkins Company, Baltimore, MD (1925)
92. Macy, M.W., Willer, R.: From factors to actors: computational sociology and agent-based modeling. Ann. Rev. Sociol. **28**(1), 143–166 (2002)
93. Madsen, J.K., Bailey, R., Carrella, E., Koralus, P.: Analytic versus computational cognitive models: agent-based modeling as a tool in cognitive sciences. Current Dir. Psychol. Sci. **28**(3), 299–305 (2019)
94. Maggi, E., Vallino, E.: Understanding urban mobility and the impact of public policies: the role of the agent-based models. Res. Transp. Econ. **55**, 50–59 (2016)
95. Maĭstrov, L.E.: Probability Theory: A Historical Sketch. Academic Press, New York, NY (1974)
96. Malerba, F., Nelson, R.R., Orsenigo, L., Winter, S.G.: History-friendly models of industry evolution: the computer industry. Ind. Corp. Change **8**(1), 3–40 (1999)
97. Malerba, F., Nelson, R.R., Orsenigo, L., Winter, S.G.: History-friendly models: an overview of the case of the computer industry. J. Artif. Soc. Soc. Simul. **4**(3) (2001)
98. Manzo, G.: Variables, mechanisms, and simulations: can the three methods be synthesized? A critical analysis of the literature. Rev. Fr. Sociol. **48**(5), 35 (2007)
99. March, J.G.: Exploration and exploitation in organizational learning. Organ. Sci. **2**(1), 71–87 (1991)
100. Marsaglia, G.: The Marsaglia random number CDROM including the DieHard battery of tests of randomness (1995)
101. Mäs, M., Flache, A., Takács, K., Jehn, K.A.: In the short term we divide, in the long term we unite: demographic crisscrossing and the effects of faultlines on subgroup polarization. Org. Sci. **24**(3), 716–736 (2013)
102. Mayr, E.: The Growth of Biological Thought: Diversity, Evolution, and Inheritance. Harvard University Press, Cambridge, MA (1982)
103. McCloskey, D.N.: The Rhetoric of Economics. Rhetoric of the Human Sciences, 2nd edn. University of Wisconsin Press, Madison, WI (1998)
104. Meadows, D.H., Meadows, D.L., Randers, J., III Behrens, W.W. (eds.) The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind. Universe Books, New York, NY (1972)
105. Metropolis, N.: The beginning of the Monte Carlo method. Los Alamos Sci. **15**(Special Issue), 125–130 (1987)

106. Metropolis, N., Ulam, S.: The Monte Carlo method. J. Amer. Statist. Assoc. **44**, 335–341 (1949)
107. Mill, J.S.: A System of Logic, Ratiocinative and Inductive, vol. 1. John W. Parker, West Strand, London (1843)
108. Miller, J.H., Page, S.E.: Complex Adaptive Systems: An Introduction to Computational Models of Social Life. Princeton University Press, Princeton, NJ (2007)
109. Miller, K.D.: Agent-based modeling and organization studies: a critical realist perspective. Organ. Stud. **36**(2), 175–196 (2015)
110. Miller, K.D., Pentland, B.T., Choi, S.: Dynamics of performing and remembering organizational routines: performing and remembering organizational routines. J. Manage. Stud. **49**(8), 1536–1558 (2012)
111. Morgan, C.L.: Emergent Evolution: The Gifford Lectures, Delivered in the University of St. Andrews in the Year 1922. Williams and Norgate, London (1923)
112. Morgan, C.L.: The Emergence of Novelty. Williams & Norgate, London (1933)
113. Muth, J.F.: Rational expectations and the theory of price movements. Econometrica **29**(3), 315 (1961)
114. Nance, R.E.: A history of discrete event simulation programming languages. In: The Second ACM SIGPLAN Conference on History of Programming Languages, HOPL-II, pp. 149–175, New York, NY. ACM (1993)
115. Nance, R.E.: A history of discrete event simulation programming languages. In: History of Programming Languages—II, pp. 369–427. ACM, New York, NY (1996)
116. Ng, T., Wright, M.: Introducing the MONIAC: an early and innovative economic model. Reserve Bank of New Zealand: Bull. **70**(4), 46–52 (2007)
117. Niazi, M., Hussain, A.: Agent-based computing from multi-agent systems to agent-based models: a visual survey. Scientometrics **89**(2), 479–499 (2011)
118. Nilsson, F., Darley, V.: On complex adaptive systems and agent-based modelling for improving decision-making in manufacturing and logistics settings: experiences from a packaging company. Int. J. Oper. Prod. Manag. **26**(12), 1351–1373 (2006)
119. Orcutt, G.H.: A new type of socio-economic system. Rev. Econ. Stat. **39**(2), 116 (1957)
120. Ostrom, E.: The ten most important books. Tidsskriftet Politik **4**(7), 36–48 (2004)
121. Pantin, C.F.A.: Relations Between Sciences. Cambridge University Press, Cambridge (1968)
122. Pool, I.D.S., Abelson, R.P.: The simulmatics project. Publ. Opin. Quart. **25**(2), 167, 22 (1961)
123. Pool, I.D.S., Abelson, R.P., Popkin, S.L.: Candidates, Issues and Strategies: A Computer Simulation of the 1960 and 1964 Presidential Elections. MIT Press, Cambridge, MA (1965)
124. Pyka, A., Fagiolo, G.: Agent-based modelling: a methodology for neo-schumpetarian economics. In: Elgar Companion to Neo-Schumpeterian Economics, pp. 467–488. Edward Elgar Publishing, Cheltenham (2007)
125. Railsback, S.F., Grimm, V.: Agent-Based and Individual-Based Modeling: A Practical Introduction. Princeton University Press, Princeton, NJ (2012)
126. Raub, W., Voss, T.: Micro-macro models in sociology: antecedents of Coleman's diagram. In: Jann, B., Przepiorka, W. (eds.) Social Dilemmas, Institutions, and the Evolution of Cooperation. De Gruyter, Berlin, Boston, MA (2017)
127. Riedwyl, H.: Rudolf Wolf's contribution to the Buffon needle problem (an early Monte Carlo experiment) and application of least squares. Am. Stat. **44**(2), 138–139 (1990)
128. Rosato, A., Prinz, F., Standburg, K.J., Swendsen, R.H.: Monte Carlo simulation of particulate matter segregation. Powder Technol. **49**(1), 59–69 (1986)
129. Rosato, A., Strandburg, K.J., Prinz, F., Swendsen, R.H.: Why the Brazil nuts are on top: size segregation of particulate matter by shaking. Phys. Rev. Lett. **58**(10), 1038–1040 (1987)
130. Ruelle, D.: Chance and Chaos. Princeton University Press, Princeton, NJ (1993)
131. Schelling, T.C.: Models of segregation. Am. Econ. Rev. **59**(2), 488–493 (1969)
132. Schelling, T.C.: Dynamic models of segregation. J. Math. Sociol. **1**(2), 143–186 (1971)
133. Schelling, T.C.: Micromotives and Macrobehavior. Fels Lectures on Public Policy Analysis. Norton, New York, NY (1978)

134. Secchi, D.: A case for agent-based models in organizational behavior and team research. Team Perform. Manag. Int. J. **21**(1/2), 37–50 (2015)
135. Secchi, D.: How do I Develop an Agent-Based Model? Edward Elgar Publishing, Cheltenham (2022)
136. Secchi, D., Neumann, M. (eds.): Agent-Based Simulation of Organizational Behavior. New Frontiers of Social Science Research. Springer International Publishing, Cham (2016)
137. Secchi, D., Seri, R.: Controlling for false negatives in agent-based models: a review of power analysis in organizational research. Comput. Math. Organ. Theory **23**(1), 94–121 (2017)
138. Segrè, E.: From X-Rays to Quarks: Modern Physicists and Their Discoveries. W. H. Freeman, San Francisco, CA (1980)
139. Sen, A.: Maximization and the act of choice. Econometrica **65**(4), 745 (1997)
140. Seri, R., Secchi, D.: How many times should one run a computational simulation? In: Edmonds, B., Meyer, R. (eds.) Simulating Social Complexity: A Handbook. Understanding Complex Systems, pp. 229–251. Springer International Publishing, Cham (2017)
141. Smith, A.: An Inquiry into the Nature and Causes of the Wealth of Nations, vol. 2. Printed for W. Strahan and T. Cadell, in the Strand, London (1776)
142. Spanier, J.: Monte Carlo Methods. In: Nuclear Computational Science, pp. 117–165. Springer, Dordrecht (2010)
143. Squazzoni, F.: Agent-Based Computational Sociology. John Wiley & Sons Ltd, Chichester (2012)
144. Sterman, J.D.: System dynamics modeling: tools for learning in a complex world. Calif. Manag. Rev. **43**(4), 8–25 (2001)
145. Tesfatsion, L.: Chapter 16: Agent-based computational economics: a constructive approach to economic theory. In: Handbook of Computational Economics, vol. 2, pp. 831–880. Elsevier (2006)
146. Troitzsch, K.G.: Perspectives and challenges of agent-based simulation as a tool for economics and other social sciences. In: Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems, AAMAS '09, vol. 1, pp. 35–42, Richland, SC, 2009. International Foundation for Autonomous Agents and Multiagent Systems
147. Tubaro, P., Casilli, A.A.: An ethnographic seduction: how qualitative research and agent-based models can benefit each other. Bull. Sociol. Methodol./Bull. Méthodol. Sociol. **106**(1), 59–74 (2010)
148. Ulam, S.: On some mathematical problems connected with patterns of growth in figures. In: Bellman, R.E. (ed.) Mathematical Problems in the Biological Sciences. Number 14 in Proceedings of Symposia in Applied Mathematics, pp. 215–224. American Mathematical Society, Providence, RI (1962)
149. van Bertalanffy, L.: General System Theory: Foundations, Development, Applications. Braziller, New York, NY (1968)
150. Vinković, D., Kirman, A.P.: A physical analogue of the Schelling model. Proc. Natl. Acad. Sci. **103**(51), 19261–19265 (2006)
151. Volterra, V.: Fluctuations in the abundance of a species considered mathematically. Nature **118**(2972), 558–560 (1926)
152. Volterra, V.: Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. Atti della R. Accademia nazionale dei Lincei. Memorie della Classe di scienze fisiche, matematiche e naturali **2**(III), 31–113 (1926)
153. von Neumann, J.: The general and logical theory of automata (with discussion). In: Jeffress, L.A. (ed.) Cerebral Mechanisms in Behaviour, pp. 1–41. Wiley, Chapman & Hall, New York, NY; London (1951)
154. von Neumann, J.: Various techniques used in connection with random digits. In: Householder, A.S., Forsythe, G.E., Germond, H.H. (eds.) Monte Carlo Method. National Bureau of Standards Applied Mathematics Series, vol. 12, Chap. 13, pp. 36–38. US Government Printing Office, Washington, DC (1951)
155. von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press, Princeton, NJ (1944)

156. Wall, F.: Agent-based modeling in managerial science: an illustrative survey and study. Rev. Manag. Sci. **10**(1), 135–193 (2016)
157. Wang, L., Ahn, K., Kim, C., Ha, C.: Agent-based models in financial market studies. J. Phys. Conf. Ser. **1039**, 012022 (2018)
158. Weiss, G. (ed.): Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence. MIT Press, Cambridge, MA (1999)
159. Wolf, R.: Versuche zur Vergleichung der Erfahrungswahrscheinlichkeit mit der mathematischen Wahrscheinlichkeit: Vierte Versuchsreihe. Mittheilungen der naturforschenden Gesellschaft in Bern **176**, 85–88 (1850)
160. Yoon, M., Lee, K.: Agent-based and "History-Friendly" models for explaining industrial evolution. Evol. Inst. Econ. Rev. **6**(1), 45–70 (2009)