Physicæ

Understanding the ergodic hypothesis via analogies

Andre R. Cunha*
Laboratório de Meios Porosos e Propriedades Termofísicas
Universidade Federal de Santa Catarina,
Campus Universitário - Trindade
Florianópolis, Santa Catarina,
Brasil - CEP 88040-900 - P.O. Box 476

The Ergodic Hypothesis is a hypothesis in Statistical Mechanics that relates the microscopic motion of particles with the macroscopic average, i. e., the observed property. Despite its importance, didactically its understanding is not easy due to technical issues. Therefore, in this article we propose analogies in order to clarify some important features of the referred hypothesis. Our starting point is the perception that the same macroscopic property, i. e., the average of the movement, can be calculated by different procedures. After that, we build the same average in a more convenient way. We do not have as objectives to contemplate advanced implications of the referred hypothesis. On such cases, some papers will be referenced.

I. INTRODUCTION

In many theoretical contexts, technical sophistication often blurs the meaning and intuition behind a theory. Besides, the difficulties that enclosure fundamental questions usually have an inhibitor effect on curious eyes. Accepting such truth, textbooks' authors follow the tendencies of technical tutorials, banning from their pages crucial explanatory steps, which may possibly impair the proper discernment of a hypothesis. With that concern in mind, we propose some analogies to help explaining an important topic in Statistical Mechanics: the Ergodic Hypothesis (EH).

A. Statistical Mechanics

In the last decades of 19th century, great advances were made, concerning the comprehension of matter structure. Despite the resistance of the *energetics' group* [1], the works by Maxwell, Boltzmann, Gibbs and others were successful on interpreting matter as composed by the minors entities called *atoms*. Under this new perspective, many obstacles needed to be overcome, and EH appears with singular voice to solve some problems.

Briefly, the EH states that the temporal average of the movement of molecules and atoms is equal to the spatial average (in phase space, see eqs. (1) and (9)). In other words, this means that a macroscopic property can be interpreted as an average taken over different data.

Despite the fundamental nature of the hypothesis, its presentation in many textbooks is underestimated, giving an impression of being either a mere step in mathematical calculations or even a self-evident conclusion. We mention some examples: ref. [3] just mentions it. Refs. [4, 5] make some comments, specially ref. [6] whose dense book about fundamentals of Statistical Mechanics has discussions about EH's implications. However, no one emphasizes its intuitive meaning. Other texts about Kinetic Theory of Gases [7, 8] make use of that meaning without justifying or even commenting about it. Ref. [9] is an exception that introduces EH to connect micro- and macroscopic scales. In short, even on those texts that recognizes EH's fundamental role, its hard to comprehend of the reasons that take us to adopt it.

We mention refs. [10–12] as good sources of information, despite the fact they are high technical level readings. We may still mention ref. [13] as an example that shows the actual relevance of this subject, which deals with the Quantum Ergodic Theorem according to John von Neumann.

II. DESCRIPTIONS OF THE MICROSCOPIC STATE

The description of a physical system, such as an ideal gas with N particles, is usually carried out in terms of $3N$ spatial coordinates $q_1,\ldots,q_{3N} \equiv \mathbf{q}$; and the respective $3N$ conjugated momenta $p_1,\ldots,p_{3N} \equiv \mathbf{p}$. We are then considering the Hamiltonian formalism. We still define a differential volume of phase space by $d\omega \equiv dq_1 \ldots dq_{3N} dp_1 \ldots dp_{3N}$.

Imagine that we intend to calculate an (intensive) macroscopic property $\mathscr{G}_{obs}$, where *obs* means observable, taking the atomics hypothesis as true. An usual way of trying to extract a simple and unique value from an infinite number of irregularly moving particles is to *properly* conceive an average of the motion (position or velocity) of the particles.

With the aim of testing the validity of the previous procedure, we will perform an experiment. Due to the fact that any experimental measurement occurs in a finite time τ, the required average must be evaluated during that interval[2] [9].

*cunha@lmpt.ufsc.br

[1] The energeticists were those who disagreed that the matter would be subdivided into smaller entities, the *atoms*. As representatives of this current of thought include Helm and Ostwald [1, 2].

[2] Implicitly, we are considering that the system reaches equilibrium during the finite time interval τ, otherwise there would be no use in computing this average. So we focus our attention on studying systems in equilibrium [14].

Mathematically, this is translated as,

$$\mathscr{G}_{obs} = \bar{G}(\tau) = \frac{1}{\tau}\int_0^\tau G(\mathbf{q},\mathbf{p};t)dt \quad . \tag{1}$$

Even if the function G is known, it is impracticable to compute all $6N$ equations of movement together with their precise initial conditions. In this way, the temporal average is useless at getting some information.

How could we obtain $\mathscr{G}_{obs}$? Is eq. (1) adequate? Is that the only possible way to calculate the macroscopic measurement?

III. THE CONSUMER AND ITS EXPENDITURES

Imagine that we wish to know the *mean daily consumption* (macroscopic property $\mathscr{Q}_{obs}$) of one consumer; we call him ℓ^*. In this case, we propose to track him during a period of a month (30 days), after which, we calculate the average. In mathematical notation,

$$\mathscr{Q}_{obs} = \bar{Q}(d) = \frac{1}{30}\sum_{d=1}^{30} Q(d) \quad . \tag{2}$$

We can now suppose that we do not have an entire month to observe ℓ^*; in fact, we have just one day. How could we know that average? Would it be possible to know $\mathscr{Q}_{obs}$ so far?

If we believe – *here is the thesis* – that there are many consumers L with similar profiles of ℓ^*, i.e., a socioeconomic group $\mathfrak{L}$ to which ℓ^* belongs, we could observe many of them for *just one day*, taking notes of their expenditures and calculating the desired average:

$$\langle Q(\ell)\rangle = \frac{1}{L}\sum_{\ell=1}^{L} Q(\ell) \quad . \tag{3}$$

The family of L consumers sent to shopping is commonly called "*ensemble*". If we take ℓ^* as one of the many possibilities of the ensemble $\mathfrak{L}$, we say that ℓ^* is a *realization of ensemble* [10, 15, 16].

It is obvious that in a certain day the consumer ℓ^* can perform and probably he will, different expenditures from that performed by the consumer i, which will be different from those of j, and so on.

In fact, what we are supposing is that the unique consumer ℓ^* could perform, in any of the 30 days of the month, the several expenditures of his pairs, in such way that the average of the socioeconomic class $\mathfrak{L}$ taken during one simple day equals the average of the unique consumer ℓ^* taken during a whole month.

Hypothesis :

$$\left(\begin{array}{c} L \text{ consumers} \\ \text{in just one day} \end{array}\right) \equiv \left(\begin{array}{c} \text{an unique consumer } \ell^* \\ \text{during a whole month} \end{array}\right) \quad .$$

In this way, we change a temporal average $\bar{Q}(d)$ for another one built on socioeconomic class $\mathfrak{L}$, $\langle Q(\ell)\rangle$. The important fact is that we consider ourselves successful in satisfying our curiosity about the *mean daily consumption* of the specific consumer ℓ^*. Mathematically,

$$\mathscr{Q}_{obs} = \bar{Q}(d) = \langle Q(\ell)\rangle \quad . \tag{4}$$

IV. EXPANDING THE ANALOGY

The idea above shows our goal: exchanging the average based on the time evolution of a single agent, for another based on a set of similar agents.

With this motivation we choose not to hold ourselves in the class $\mathfrak{L}$. Instead of sending the consumers of $\mathfrak{L}$ to shopping in just one day, we can – in an abstract way – give life to each possible expenditures instantaneously, i.e., in a time interval $\Delta t \to 0$. This means we are changing the focus from the class of consumers $\mathfrak{L}$ to the *expenditures space* $\mathfrak{M}$ and eliminating the time dependence of the average.

We know that there are an infinity of possibilities of expenditures (from one simple candy from the neighborhood grocery store to a luxurious car)[3], then the sum must be replaced by an integral. Besides, it is natural to believe that buying a simple candy is much more common than buying a sport car. So we need to assign a *probability density* $f(m)$ to the occurrence of each expenditure m of the set $\mathfrak{M}$.

In mathematical terms, we now express the *mean daily consumption*,

$$\langle Q(m)\rangle = \int_{\mathfrak{M}} f(m)Q(m)dm \quad , \tag{5}$$

and,

$$\int_{\mathfrak{M}} f(m)dm = 1 \quad , \tag{6}$$

since f is a probability density.

This procedure allows us to know the same property $\mathscr{Q}_{obs}$ in three different ways,

$$\mathscr{Q}_{obs} = \bar{Q}(d)\big|_t = \langle Q(\ell)\rangle_{\mathfrak{L}} = \langle Q(m)\rangle_{\mathfrak{M}} \quad . \tag{7}$$

V. THE ERGODIC HYPOTHESIS

The EH consists in changing the temporal average of eq. (1) for a *spatial* average. More precisely, for an average on the *phase space* of Hamiltonian formalism.

In this formalism, we express the microscopic state of a system with just one point in a space of $6N$ dimensions. Then instead of tracking a long path of this single point during the time interval τ and computing the integral in eq. (1), we will imagine an *ensemble* of similar points[4] and assign to each portion of that ensemble a probability density $f_N(\mathbf{q},\mathbf{p};t)$ that quantifies the movements in all directions of phase space. Afterwards, we integrate the possibilities in all directions to obtain the macroscopic tendency.

$$\mathscr{G}_{obs} = \langle G(t)\rangle = \int_{\mathbf{\Omega}} f_N(\mathbf{q},\mathbf{p};\Delta t\to 0)\, G(\mathbf{q},\mathbf{p};\Delta t\to 0)d\omega \quad , \tag{8}$$

---

[3] We should have in mind that the only possible expenditures are those that characterize the socioeconomic group $\mathfrak{L}$.

[4] The similarity lies in the fact that all the replicas of the point must correspond to the same macroscopic state.

where $d\omega$ is the volume element of a phase space region $\mathbf{\Omega}$. We may even drop the explicit time dependence and write it down,

$$\mathscr{G}_{obs} = \langle G(t) \rangle = \int_{\mathbf{\Omega}} f_N(\mathbf{q},\mathbf{p}) G(\mathbf{q},\mathbf{p}) d\omega \quad . \tag{9}$$

A. Why choose an average over the phase space?

Think of a landscape with mountains, valleys, cliffs, plains, rivers, sea, etc. Imagine now that you are with a group of friends in some point of that landscape, on a valley, for example. You are lost and tired, and a question arise: where should I go? Two main alternatives appears:

i. if everybody walk together as an unique point during many days in certain direction, after that long period the group will ponder the displacement, i.e., compute an average.

ii. or, individually, each one chooses a direction to go, and after some hours, return with some information. Taking this informations as basis, the group will decide which direction to follow.

In the first alternative, the group goes as unity. During the displacement, it ponders about the mountains, rivers, cliffs (in other words, the topography and geometry) of the landscape. Depending on amount of food and fatigue (energy), the group will choose an specific route. After some days, it may decide the trip path.

Look at the second alternative. Imagine that one person returns after two hours and says that there is a high climb towards the south. Due to the fact that the group is very tired, you choose to give less importance (low probability density) to that information. The same occurs with another information: there is a cliff towards the east, which means that direction is impracticable (null probability density). In short, the group is planning the steps considering the physical descriptions (topography and geometry), as well as the limitation of people (energy). And that is done before computing the collective displacement. *Ergodic Hypothesis affirms that both alternatives get the same result.*

When we decide to describe a physical system by means of Hamiltonian formalism, we are taking into account information about geometry and energy of the system. The richness of the underlying mathematical structure [17–20] is the motivation to conceive an average on the phase space.

VI. CONCLUSION

In this article we use analogies to emphasize the phenomenological side of the Ergodic Hypothesis. The analogy is based on the fact that a specific macroscopic property of a system can be calculated from different sets of data.

EH is a product of intense research and debate, and gives life to a proper branch of Physics and Mathematics. For example, an interesting result due to Birkhoff [11, 21] establishes that the validity of EH ensures the result of *equal a priori probabilities* (many times taken as a postulate). If we translate to the context of the proposed analogy, this means that every product has the same possibility of being purchased by the consumer, i.e., $f(m)$ of eqs. (5) and (6) assumes the same value for any m in $\mathfrak{M}$, either a bicycle or a luxury car. From this perspective, we can realize that this result due to Birkhoff is somewhat counterintuitive, which can lead us to ponder on the validity of a result so restrictive. On the other hand, ref. [22] affirms that ergodic theorems and the mentioned postulate are both the main approaches of foundations of Statistical Mechanics, and one is not necessary for the other; in addition, it discusses cases where the hypothesis cannot be applicable. Ref. [23] goes further arguing that *"there is no complete justification for the postulate, even if the ergodic theorems are applicable"*. As a final judgment, we note that the refs. [9] and [12] says that experimental confirmation is the dominant criteria for either accepting or refuting EH's validity.

Finally, refs. [11] and [12] may serve as a next step in addition to this paper, and may be used for discussing non-ergodic systems and modern advances of Ergodic Theory.

Acknowledgments

We would like to thank the anonymous reviewers for their careful reading and comments. We also thank G. G. Gomes and E. Muller dos Santos for their valuable comments. Finally, we thank CAPES and Material Science and Engineering Post-Graduate Program of Federal University of Santa Catarina.

[1] Sílvio R. Dahmen. A obra de Boltzmann em Física. *Revista Brasileira de Ensino de Física*, 28:281–295, 2006.

[2] Antonio Augusto P. Videira. Atomismo, energetismo e pluralismo teórico no pensamento epistemológico de Ludwig Boltzmann. *Química Nova*, 17:461–464, 1994.

[3] R. K. Pathria. *Statistical Mechanics*. Butterworth-Heinemann, 1996.

[4] W. Greiner, L. Neise, and H. Stöcker. *Thermodynamics and statistical mechanics*. Springer Verlag, 1995.

[5] Gregory H. Wannier. *Statistical Physics*. Dover Publications, Inc, 1966.

[6] R. Tolman. *The Principles of Statistical Mechanics*. Dover Publications, Inc, 1938.

[7] Sydney Chapman and T. G. Cowling. *The Mathematical Theory of Non-Uniform Gases*. Cambrigde Univeristy Press, 3 edition, 1970.

[8] Richard L. Liboff. *Kinetic theory: classical, quantum, and relativistic descriptions*. Springer Verlag, 2003.

[9] Stewart Harris. *An introduction to the theory of Boltzmann Equation*. Dover Publications, Inc, 1971.

[10] N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier, 3 edition, 2007.

[11] L. E. Reichl. *A Modern Course in Statistical Physics*. John Wiley & Sons, Inc, 2 edition, 1998.

[12] César R. de Oliveira and Thiago Werlang. Ergodic hypothesis in classical statistical mechanics. *Revista Brasileira de Ensino de Física*, 29(2):189–201, 2007.

[13] S. Goldstein, J.L. Lebowitz, R. Tumulka, and N. Zanghì. Long-time behavior of macroscopic quantum systems. *The European Physical Journal H*, 2010.

[14] Herbert B. Callen. *Thermodynamics and an Introduction to Thermostatics*. John Wiley & Sons, Inc, 1985.

[15] Steven Kay. *Intuitive Probability and Random Processes*. Springer, 2006.

[16] Athanasios Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, Inc., 1991.

[17] V. I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer-Verlag, 1997.

[18] Nivaldo A. Lemos. *Mecânica Analítica*. Livraria da Física, 2004.

[19] Walter F. Wreszinski. *Mecânica Clássica Moderna*. EDUSP, 1997.

[20] H. Goldstein. *Classical Mechanics*. Addison-Wesley, 1950.

[21] Adrian Patrascioiu. The ergodic-hypothesis: a complicated problem in Mathematics and physics. *Los Alamos Science*, Special Issue:263–279, 1987.

[22] Oliver Penrose. *Foundations of statistical mechanics: a deductive treatment*, page 39. Pergamon Press, 1969.

[23] Colin J. Thompson. *Mathematical statistical physics*, page 57. Macmillan, 1972.