1 Introduction

1.1 The Aims of Statistical Mechanics

Statistical mechanics (SM) is the third pillar of modern physics, next to quantum theory and relativity theory. Its aim is to account for the behaviour of macroscopic systems in terms of the dynamical laws that govern their microscopic constituents and probabilistic assumptions about them. The use of probabilities is motivated by the fact that systems studied by SM have a large number of microscopic constituents. Paradigmatic examples of systems studied in SM are gases, liquids, crystals, and magnets, which all have a number of microscopic constituents that is of the order of Avogadro's number (6.022×10^{23}).[1]

The focal point of SM is a particular aspect of the behaviour of macro-systems, namely equilibrium. To introduce equilibrium, and to boost intuitions, let us consider a standard example. A gas is confined to the left half of a container with a dividing wall, as illustrated in Figure 1a. The gas is in *equilibrium* in the sense that there is no manifest change in any of its macro properties like pressure, temperature, and volume, and the gas will have these macro properties so long as the container remains unchanged. Now consider such a change: remove the dividing wall in the middle, as illustrated in Figure 1b. The gas is now no longer in equilibrium because it does not fill the container evenly. As a result, the gas starts spreading through the entire available volume, as illustrated in Figure 1c. The spreading of the gas comes to an end when the entire container is filled evenly and no further change takes place, as illustrated in Figure 1d. At this point, the gas has reached a new equilibrium. Since the process of spreading culminates in a new equilibrium, this process is an *approach to equilibrium*. A key characteristic of the approach to equilibrium is that it seems to be *irreversible*: systems move from non-equilibrium to equilibrium, but not vice versa: gases spread to fill the container evenly, but they do not spontaneously concentrate in the left half of the container. Since an irreversible approach to equilibrium is often associated with thermodynamics, this is referred to as *thermodynamic behaviour*.[2]

[1] From now on, we will use 'micro' and 'macro' as synonyms for 'microscopic' and 'macroscopic', and we will speak of macro systems and their micro constituents, as well as of macro or micro properties, macro or micro characterisations, and so on. We will also use 'micro-states' and 'macro-states', where we add the hyphen to indicate that they are unbreakable technical terms. We use 'SM systems' as a shorthand for 'the systems studied in SM'.

[2] This is commonly seen as a result of the second law of thermodynamics. However, as Brown and Uffink (2001) note, the irreversibility of the kind illustrated in our example is not part of the second law and has to be added as an independent principle to the theory, which they call the *minus-first law*.

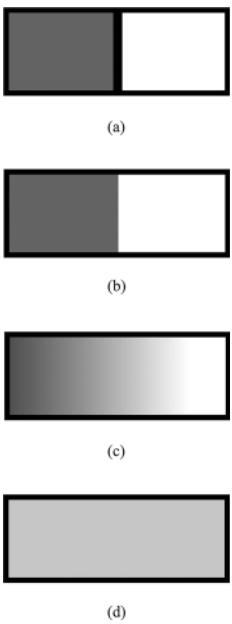

Figure 1 The spreading of gas when removing a shutter.

Characterising the state of equilibrium and accounting for why, and how, a system approaches equilibrium are the core tasks for SM. Sometimes these two problems are assigned to separate parts of SM, namely *equilibrium SM* and *non-equilibrium SM*.

What does characterising the state of equilibrium involve? We said that the gas is in equilibrium when there is no manifest change in any of its macro properties. This is a valid macro characterisation. Yet if the aim of SM is to account for the macro properties of a system in terms of the behaviour of its micro constituents, then we need a characterisation of equilibrium in micro-physical terms: what condition does the motion of the molecules in a gas have to satisfy for the gas to be in equilibrium? And how do the values of macroscopic properties like local pressure and local temperature depend on the state of motion of gas molecules? Equilibrium SM provides answers to these and related questions.

Turning to non-equilibrium, the core question is how the tendency of systems to move to equilibrium when prepared in a non-equilibrium state is grounded in the dynamics of the micro constituents of the system: what is it about the motions of molecules that leads them to spread so that the gas assumes a new equilibrium state when the shutter is removed? And, crucially, what accounts for the fact that the reverse process does not happen?

This Element focuses on these questions. This, however, is not to say that SM deals solely with equilibrium. While equilibrium occupies centre stage, SM also addresses other issues such as phase transitions, the entropy costs of computation, and the process of mixing substances; and in philosophical contexts SM has been employed to shed light on issues like the direction of time and the possibility of knowledge about the past. Space constraints prevent us from delving into these issues, and we will concentrate on equilibrium.

1.2 The Theoretical Landscape of SM

Foundational debates in many other areas of physics can take as their point of departure a generally accepted formalism and a clear understanding of what the theory is. A discussion of the nature of space and time, for instance, can base its considerations on the general theory of relativity, and a discussion of the foundations of quantum mechanics has the formalism of the theory as reference point. The situation in SM is different because, unlike quantum mechanics and relativity theory, SM has not yet found a generally accepted theoretical framework, let alone a canonical formulation. Those delving into SM find a multitude of different approaches and schools of thought, each with its own conceptual apparatus and formal structure.[3] A discussion of the foundations of SM can therefore not simply begin with a concise statement of the formalism of SM and its basic principles. Indeed, the choice, articulation, and justification of a theoretical framework for SM is an integral part of the foundational endeavour!

It has become customary in the foundations of SM to organise most (although not all) theoretical approaches in SM under one of two broad theoretical umbrellas. These umbrellas are known as Boltzmannian SM (BSM) and Gibbsian SM (GSM) because their core principles are attributed to Ludwig Boltzmann (1844–1906) and Josiah Willard Gibbs (1839–1903) respectively. Accordingly, approaches are then classified as 'Boltzmannian' or 'Gibbsian'.[4] In this Element, we follow this classificatory convention and use it to structure our discussion. We note, however, that from a historical point of view, this labelling is not entirely felicitous. While Boltzmann did indeed champion the approach now known as BSM, his work was in no way restricted to it.

[3] For reviews of these approaches see Frigg (2008b), Penrose (1970), Shenker (2017a, 2017b), Sklar (1993), and Uffink (2007).

[4] There is an interesting historical question about the origin and subsequent evolution of this schism. We cannot discuss this question here, but we speculate that it goes back to Ehrenfest and Ehrenfest-Afanassjewa's seminal review (1912/1959). While they do not use the labels BSM and GSM, they clearly separate a discussion of Boltzmann's ideas, which they discuss under the heading of 'The Modern Formulation of Statistico-Mechanical Investigations', and Gibbs' contributions, which they discuss in a section entitled 'W. Gibbs's *Elementary Principles in Statistical Mechanics*'.

Boltzmann in fact explored many different theoretical avenues, among them also ensemble methods that are now classified as Gibbsian. So 'BSM' is a neologism that should not be taken to be an accurate reflection of the scope of Boltzmann's own work.[5]

1.3 Outline

In this Element we discuss how BSM and GSM deal with the questions introduced in Section 1.1. Both BSM and GSM are formulated against the background of dynamical systems theory and probability theory. In Section 2 we introduce the basic concepts of both theories and state results that are important in the context of SM. In Section 3 we turn to BSM. We start by introducing the core concepts of BSM and then discuss different ways of developing them. Section 4 is dedicated to GSM. After introducing the formalism, we discuss its interpretation and different developments, and we end by considering the relation between BSM and GSM.

It goes without saying that omissions are inevitable. Some of them we have already announced in Section 1.1: we focus on equilibrium and the approach to it, and we set aside topics like phase transitions and the entropy costs of computation, and we will only briefly touch on questions concerning reductionism.[6] In addition to these, we set aside approaches to SM that do not clearly fall under the umbrella of either BSM or GSM, which implies that we do not discuss the Boltzmann equation.[7]

2 Mechanics and Probability

As the name 'statistical mechanics' indicates, SM aims to give an account of the behaviour of systems in terms of mechanics and statistics. The use of the word 'statistics' in this context is, however, out of sync with its modern use, where the term usually means something like 'the technology of extracting meaning from data' (Hand 2008, 1). As a branch of theoretical physics, SM is not concerned with the collection and interpretation of data (although data are of course important in testing the theory). In the second half of the nineteenth century, when the foundations of the discipline were laid, one of the main uses of the term 'statistics' was also to designate a 'description of the properties or behaviour of a collection of many atoms, molecules, and so on, based on the application of probability theory' (*OED* 'statistics'). So a 'statistical' treatment of

[5] For an overview of Boltzmann's contributions to SM see Uffink (2022); for detailed discussions see, for instance, Brush (1976), Cercignani (1998), Darrigol (2018), and Uffink (2007).
[6] For recent discussions of reduction with special focus on statistical mechanics, see Batterman (2002), Butterfield (2011a, 2011b), Lavis, Kühn, and Frigg (2021), and Palacios (2022).
[7] For a discussion of the Boltzmann equation, see Uffink (2007).

a problem is simply a treatment in terms of probability theory. This use of 'statistics' has become less common, and hence 'probabilistic mechanics' might better describe the discipline to a modern reader. But, for better or worse, historical labels stick. What this short excursion into the etymology of 'statistical mechanics' brings home is that the theory builds on two other disciplines, namely mechanics and probability, which provide its background theories.

The aim of this section is to introduce these background theories. The mechanical background theory against which SM is formulated can be either classical mechanics or quantum mechanics, resulting in either classical SM or quantum SM. Foundational debates are by and large conducted in the context of classical SM. We follow this practice in this Element, and for this reason the current section focuses on classical mechanics.[8]

The section is structured as follows. We begin by introducing dynamical systems at a general level, along with some basic mechanical notions like *trajectory*, *measure*, and *determinism* (Section 2.1). We then have a closer look at a specific class of dynamical systems, Hamiltonian systems, which are important in this context because they provide the fundamental structure of SM systems (Section 2.2). Hamiltonian systems have two dynamical properties that feature prominently in discussions about SM, namely time-reversal invariance (Section 2.3) and Poincaré recurrence (Section 2.4). Being ergodic is a property that certain time evolutions possess. This property plays an important role in SM because different approaches appeal to it to justify the equilibrium behaviour of SM systems (Section 2.5). This brings our discussion of mechanics to a conclusion, and we turn to probability. We introduce the basic formalism of probability theory and discuss the three main philosophical interpretations of probability (Section 2.6).

2.1 Dynamical Systems

At the level of their micro constituents, SM systems have the structure of a so-called *dynamical system*, a triple (X, ϕ_t, μ).[9] This is illustrated in Figure 2. Here, X is the *state space* of the system: it contains all states that, the system could possibly assume. In classical mechanics, the state of motion of an object (understood as a point particle) is completely specified by saying what its position and momentum are. If a system consists of several objects, the state of motion of the system is specified by saying what the positions and the momenta of all its objects are. Molecules move in three-dimensional physical space, and so a molecule has

[8] For a discussion of foundational issues in quantum SM see, for instance, Emch (2007).
[9] Our characterisation of a dynamical system is intuitive. For mathematical discussions see, for instance, Arnold and Avez (1968) and Katok and Hasselblatt (1995).

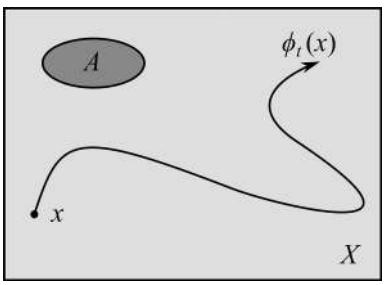

Figure 2 A dynamical system and a subset $A \subseteq X$.

three degrees of freedom. This means that a specification of the position of the molecule requires three parameters, one for each spatial direction. Since the molecule has a momentum in each direction, a further three parameters are needed to specify its momentum. So we need six parameters to specify the state of motion of a molecule. Accordingly, the specification of the state of the entire gas with n molecules requires $6n$ parameters. Hence, X is a $6n$-dimensional space containing the positions and momenta of all molecules.

The second element of a dynamical system, ϕ_t, is the *time evolution function*. This function specifies how the state of a system changes over time, and we write $\phi_t(x)$ to denote the state into which an initial state x evolves after time t. If the time evolution is specified by equations of motion like Newton's, Lagrange's, or Hamilton's, then $\phi_t(x)$ is the solution of that equation (we discuss Hamilton's equations in more detail in the next section). As time passes, $\phi_t(x)$ draws a 'line' through X that represents the time evolution of a system that was initially in state x; this 'line' is called a *trajectory*. This is illustrated in Figure 1.

The third element of a dynamical system, μ, is a *measure* on X. At a general level, a measure is a device to attribute a size to an object. Familiar examples are the attribution of a length to a segment of a line, a surface to a part of a plane, and a volume to a portion of space. From a mathematical point of view X is a set, and the measure μ attributes a 'size' – or 'measure' – to subsets of X in much the same way in which a ruler attributes a length to, say, a pencil. If the measure μ is able to attribute a size to a particular subset $A \subseteq X$, then A is said to be *measurable*. In what follows we assume that all sets are measurable. If a measure is such that $\mu(X) = 1$, then the measure is *normalised*.

Things can be measured in different ways. One measure is of particular importance in the current context: the so-called *uniform Lebesgue measure*. This measure is a mathematically precise rendering of the measure we use when attributing lengths and surfaces to objects. For instance, the interval [2, 5] has uniform Lebesgue measure (or length) 3, and a circle with radius r has uniform Lebesgue measure (or surface) of πr^2.

The time evolution ϕ_t is deterministic. It is common to define determinism in terms of possible worlds (Earman 1986, 13). Let W be the class of all physically possible worlds. The world $w \in W$ is deterministic if and only if for *any* world $w' \in W$ it is the case that: if w and w' are in the same state at some time t_0, then they are in the same state at *all* times t. This definition can be restricted to an isolated subsystem s of w. Consider the subset world $W_s \subseteq W$ of all possible worlds which contain an isolated counterpart of s, and let s' be the isolated counterpart of s in w'. Then s is deterministic if and only if for *any* world $w' \in W_s$ it is the case that if s and s' are in the same state at some time t_0, then they are in the same state at *all* times t. The system s can be a dynamical system of the kind we have just introduced. Determinism then implies that every state x has exactly one past and exactly one future, or, in geometrical terms, trajectories cannot intersect (neither themselves nor other trajectories).

2.2 Hamiltonian Mechanics

The notion of a dynamical system introduced in the previous section is extremely general, and more structure must be added to make it useful for the treatment of SM systems. The structure that is usually added is that of Hamiltonian mechanics.[10]

In Hamiltonian mechanics, the state of an object is described by its position q and its momentum p. In fact, there is a q-and-p pair for every degree of freedom. Hence, for a gas with n molecules moving in the three-dimensional physical space there are $3n$ q-and-p pairs. The $3n$ pairs constitute the state space X of the system, which in this context is often referred to as *phase space* (and since having $3n$ pairs means having $6n$ variables, the phase space has $6n$ dimensions as noted previously).

The Hamiltonian equations of motion are

$$\dot{q}_k = \frac{\partial H}{\partial p_k} \text{ and } \dot{p}_k = -\frac{\partial H}{\partial q_k},$$

where the dot indicates a derivative with respect to time and the index k ranges over all degrees of freedom (so, for a gas we have $k = 1, \ldots, 3n$). This defines a system of k differential equations, and the solution to this equation is the time evolution function ϕ_t of the system. Solving these equations for SM systems is usually a practical impossibility, and so we will not be able to write down ϕ_t explicitly. This does not render Hamiltonian mechanics useless. In fact, the equations ensure that ϕ_t has a number of general features, and these can be

[10] For a brief introduction see, for instance, Argyris, Faust, and Haase (1994, ch. 4).

established even if ϕ_t is not explicitly known. In the remainder of this section we review some of the features that are central to SM.

The Hamiltonian equations contain the function H, which is called the *Hamiltonian* of the system. The Hamiltonian H is the energy function of the system, which, in general, depends on all coordinates q_k and p_k. So, the Hamiltonian equations of motion specify how a system evolves in time based on the energy function of the system. If the energy of the system does not explicitly depend on time (which is the case, for instance, if a system is not driven by outside influences like pushes or kicks), then the system is *autonomous*. In an autonomous system, H is a conserved quantity (or a 'constant of motion'), meaning that it does not change over the course of time. This has the immediate consequence that any function $f(H)$ is a conserved quantity too.

The gas that we used as our introductory example is an autonomous system: it is isolated from the environment through the box and once the shutter has been removed, it is not subject to any outside disturbances. The gas is no exception, and typical systems in SM are autonomous Hamiltonian systems. The fact that the energy is conserved has an important consequence. Once the energy of the system is fixed to have a certain value E, the conservation of energy means that $H = E$ must hold all the time. H is a function of the coordinates q_k and p_k, which means that it defines a hypersurface X_E in X. This surface has one dimension fewer than X itself. In the case of the gas, it is therefore $6n - 1$ dimensional. This surface is known as the *energy hypersurface*.[11] Since $H = E$ holds at all times, it follows that the motion of the system is confined to the energy hypersurface: trajectories that start in an initial condition x in the energy hypersurface will never leave the hypersurface.

The total energy of the system may not be the only conserved quantity. Depending on the nature of the system, other quantities may be conserved too. From a formal point of view, each conserved quantity is a function Q of the coordinates q_k and p_k for which $Q = C$ holds for all times, where C is the value that Q assumes. In geometric terms, each conserved quantity defines a hypersurface to which trajectories remain confined. Such surfaces are also called *invariant hypersurfaces*. The crucial aspect (and this will be important later on) is that these hypersurfaces divide the phase space into regions that are 'disconnected' in the sense that trajectories cannot penetrate the surface to get to the other side. This is schematically illustrated in Figure 3, where we see an invariant surface and three trajectories. While Trajectory 1 and Trajectory 2 are

[11] It is a *hyper*surface because 'surface' has the connotation of being two-dimensional, like a tabletop, which the surface of constant energy is not.