

of Evolution. XIX. Second Supplement to a Memoir on Skew Variation', *Philosophical Transactions of the Royal Society of London*, Series A, 216 (1916), pp. 429–457.

131. Thomas Kuhn, *The Structure of Scientific Revolutions*, 3rd ed. (Chicago: University of Chicago Press, 1996), Chapter 11, p. 136 ff. Our thanks to Adrian Wilson for informed discussion on this point. Staley points out that Kuhn's views on this topic were considerably refined in his later volume on black-body radiation, which describes how Planck's views on quantum theory evolved slowly over a long period. Staley, *Einstein's Generation* (n. 1), pp. 349, 375.

CHAPTER 25

THE EMERGENCE OF STATISTICAL MECHANICS

OLIVIER DARRIGOL
AND JÜRGEN RENN¹

25.1 MECHANICAL MODELS OF THERMAL PHENOMENA

25.1.1 Heat as a Challenge to Mechanics

Statistical mechanics is the name given by the American physicist Josiah Willard Gibbs to the study of the statistical properties of a large number of copies of the same mechanical system, with varying initial conditions. In this chapter we will outline the history of statistical mechanics in a broad sense, and include any attempt to explain the thermodynamic properties of macroscopic bodies as statistical regularities of systems that encompass a very large number of similar constituents.

Statistical mechanics emerged in the second half of the nineteenth century as a consequence of efforts to account for thermal phenomena on the basis of mechanics, which was then considered to be the most fundamental of the physical sciences. Although phenomena such as light, electricity, magnetism, and heat were apparently not of a mechanical nature, scientists tried to explain them by invisible mechanical entities such as the ether or small particles. Mechanical models of thermal phenomena are part of intuitive physics and have been used since antiquity. For example, the communication of heat from one body to another can be made plausible as the motion of invisible particles, or as the flow of an invisible fluid, or else as the effect of a

Institut for Naturfagenes Didaktik
København Universitet
Øster Voldgade 3
DK-1350 København K

wave. Such mechanical explanations of heat can also help to understand its links with visible mechanical effects, for instance, the thermal expansion of bodies, the pressure exerted by gases, or the possibility to generate heat by motion. Which of these mechanical models, at a given historical moment, appeared to be most suitable as an explanation of heat depended on various factors including the available knowledge of thermal phenomena, the place of the model in the overall architecture of physics, and its state of elaboration. In the following, we briefly review some of these models and then show how several circumstances, among them the establishment of thermodynamics in the mid-nineteenth century, led to a focus on the model of heat as a motion of particles, which lies at the origin of statistical mechanics.

25.1.2 Heat as a Fluid

The late eighteenth century saw the development of the notion of heat as a fluid—the so-called ‘caloric’—in parallel to contemporary ideas on electricity. Heat and electricity share properties that suggest their understanding as ‘imponderable fluids’. Both lack a definite shape. As was demonstrated by calorimetric experiments, heat can be stored, just as electricity, in appropriate ‘containers’ and ‘flow’ from one body to another, both tend to spread out and fill their containers as much as possible, and both are imponderable; that is, they have no appreciable weight. The fluid model also accounted for the role played by heat in chemical reactions; the caloric there acted in a way similar to other substances, being either bound or set free. At the turn of the century the fluid model lent itself to Laplace’s reduction of physics to the play of central forces acting among the particles of ponderable and imponderable substances. Later, in Sadi Carnot’s reflections of 1824, it became the basis for discussing the theoretical limits of thermal machines such as steam engines (Brush, 1976; Fox, 1971, chap. 1, 9).

25.1.3 Heat as a Motion

In the first half of the nineteenth century, several developments contributed to the gradual demise of the fluid model and to the widespread acceptance of the notion of heat as a motion. First, in the 1820s the success of the undulatory theory of light affected the understanding of heat, because of the close relationship between light and heat. Second, in the 1830s Italian physicists strengthened the analogies between heat and light, for instance, by revealing the possibility of heat reflection and refraction. Third, in the 1840s James Joule performed careful experiments showing the convertibility of mechanical work into heat. By 1850 the model of heat as a motion was associated with the principle of energy conservation, whose most influential proponents were William Thomson and Hermann Helmholtz (Brush, 1976, chap. 1–3).

Ideas about what kind of motion actually constituted heat varied considerably. Originally, the most common conception was that of a vibratory motion of molecules

transmitted by the ether, following André Marie Ampère’s suggestion. While this conception accounted both for radiant heat and for heat conduction in matter, it was less relevant to the behaviour of gases. At mid-century major British physicists, among them William Rankine, William Thomson, and James Clerk Maxwell, imagined a vortex motion of gas molecules, in rough analogy with Newton’s old explanation of gas pressure in terms of a repulsion between contiguous molecules (at rest). In this picture, heat corresponded to the rotation of the molecules, and elasticity to their centrifugal force.

25.1.4 The Kinetic Theory

Around that time, James Joule, Rudolph Clausius, and others elaborated another conception of heat, as translational molecular motion. The molecules of a gas are assumed to occupy only a small fraction of its volume and to have a rectilinear, uniform motion, occasionally interrupted by mutual collisions or by collisions with the walls of the container. While this model had roots in ancient atomism, its application to gases dates from the eighteenth century. In 1738 Daniel Bernoulli first gave the corresponding explanation of gas pressure, John Herapath rediscovered it in 1820, and John Waterston did so again in the 1840s. The pressure of a gas is roughly assumed to be proportional to the number of collisions of gas particles with the wall and to their momentum. The collision number is itself proportional to the density of the gas and to the velocity of the molecules. Hence, the pressure is proportional to the density of the gas, in conformity with Boyle’s law. It is also proportional to the squared velocity of a molecule, in conformity with Gay-Lussac’s law if temperature is assumed to depend linearly on squared velocity. Initially, this model had a limited range of applicability. But in contrast to the fluid model, it accounted for the conversion between heat and mechanical work, and in contrast to the vibrational and rotational models, it allowed a quantitative description of the behaviour of gases (Brush, 1976, chap. 1–3; Brush, 1983, sec. 1.5–1.8).

25.2 GAS THEORY AS A BRIDGE BETWEEN MECHANICS AND THERMODYNAMICS

25.2.1 Thermodynamics as a Challenge to Mechanics

The formulation of thermodynamics in the 1840s and 1850s by Rankine, Thomson, and Rudolf Clausius led to a theory of heat which no longer required a mechanical model. The first and second law of this theory are enunciated purely in terms of heat-converting engines, without mention of a caloric substance or of a thermal

motion. Nevertheless, the heat-as-motion model acquired a new significance in the context of thermodynamics. Whereas it had been confined to the modelling of the thermal behaviour of gases, it now served a bridge function between two fundamental columns of physics. It made plausible the conversion processes between mechanical and thermal energy in thermodynamic engines and thus offered a mechanical underpinning for the new thermodynamics. The support gained in this way had repercussions for mechanics itself, then still widely conceived as providing a conceptual framework for all of physics. The kinetic theory of gases translated the advances of thermodynamics into challenges for mechanics that eventually led to the creation of statistical mechanics.

In particular, the study of the kinetic theory allowed scientists to confront one of the principal structural differences that distinguish thermodynamics from mechanics: the irreversibility of its laws under time reversal. When two portions of liquid of different temperatures are mixed, for example, the mixture will attain an intermediate temperature. This process is not reversible, that is, the mixture will never spontaneously separate into a cooler and a hotter component. In contrast, the time reversal of any (purely) mechanical process leads to another possible process. This conflict became one of the challenges in the development of the kinetic theory, along with an improved description of the thermal properties of specific physical systems.

25.2.2 Gases as Particles in Motion

The physicist who effectively revived the kinetic theory of gases was Rudolph Clausius, one of the founders of thermodynamics. In 1857, Clausius gave Bernoulli's relation the exact form

$$PV = \frac{1}{3}Nmu^2,$$

where P is the pressure, N the number of molecules, m their mass, and u^2 the mean value of their squared velocity. From this formula he concluded that under Avogadro's hypothesis (according to which the number of molecules in a unit volume is the same for all gases under normal conditions) the average translational energy of a molecule was the same in any gas at a given temperature. Further, he determined the internal energy of the gas as the total translational energy of its molecules. The resultant specific heat at constant volume is $(3/2)R$, where R is the constant of perfect gases. As this value was well below the observed ones, Clausius concluded that the molecules had a rotational and a vibrational motion besides their translational motion. Mutual collisions acted to maintain constant proportions for the average energies of the three kinds of motion (Clausius, 1857).

Although Clausius was aware of the spread of the velocities of the molecules due to collisions, his computations assumed, for simplicity, the same velocity for all molecules. In the approach of his first memoir, he equated the properties of the system as a whole with individual properties of the particles. For instance, he identified the temperature of the gas with the kinetic energy of its molecules. Statistical

considerations, such as the formation of average values, only played a minor role in his arguments. Clausius simply assumed that the molecules occupied a negligible fraction of the volume of the gas and that they moved in straight lines with the same velocity until they hit a wall, without treating their mutual encounters in any detail and without taking molecular forces into account. However, the direct identification of properties of the system with properties of its constituents soon brought the model into conflict with empirical properties of gases (Brush, 1976, chap. 4).

25.2.3 The Transport Properties of Gases

The issue that triggered the development of the notion of gases as statistical collections was the understanding of transport properties. One of Clausius's readers, the Dutch meteorologist C. H. D. Buys Ballot, objected that the kinetic theory implied a much faster diffusion of gases than observed, because the molecular velocities computed by Clausius were comparable to the speed of sound in normal conditions. In response, Clausius introduced what amounts to the first non-trivial statistical concept of the kinetic theory: the 'mean free path'. This concept neither describes the property of an individual particle nor the property of a gas considered macroscopically. It makes sense only if the gas is conceived as a statistical collection.

The mean free path is defined as the average length travelled by a molecule before it collides with another. The proportion of the molecules travelling in the direction Ox that hit another molecule in a slice dx of the gas, Clausius reasoned, is given by the fraction $\pi r^2 n dx$ of the surface of this slice occulted by the 'spheres of action' of the molecules, where r is the radius of action, and n the number of molecules in a unit volume. Consequently, the probability that the motion of a molecule remains free over the distance x is

$$W = e^{-\alpha x}, \text{ with } \alpha = \pi r^2 n,$$

and the probability that this molecule experiences its first collision between x and $x + dx$ is

$$W(x) - W(x + dx) = \alpha e^{-\alpha x} dx.$$

The mean free path l is the average $1/\alpha$ of the distance x under the latter probability law. Clausius judged this length to be so small that Buys Ballot's over-rapid diffusion never occurred (Clausius, 1858).

Starting from Clausius's work, James Clerk Maxwell developed the statistical aspects of the kinetic theory of gases much further. In particular, Maxwell took into account the velocity spread generated by the mutual encounters of gas molecules. In 1860 he determined the equilibrium distribution $f(\mathbf{v})$ of the molecular velocities by assuming its isotropy as well as the statistical independence of the three Cartesian components of velocity. These two conditions lead to the functional equation

$$f(\mathbf{v}) = \varphi(v^2) = \psi(v_x)\psi(v_y)\psi(v_z),$$

whose solution has the form

$$f = \alpha e^{-\beta v^2},$$

wherein α and β are two constants. Maxwell later judged this argument 'precarious', because the second assumption remained to be justified. Yet he never doubted the velocity distribution which now bears his name. As we will see, he and Boltzmann later gave better justifications (Brush, 1976; Maxwell, 1860, chap. 4, 5; Everitt, 1975).

The main purpose of Maxwell's memoir of 1860 was to develop new physical consequences of the mean-free-path concept by using a model of the gas as a large number of perfectly elastic spheres. He computed gas interdiffusion, and also internal friction (viscosity) and heat conduction regarded as diffusions of momentum and kinetic energy respectively. The molecules moving from a given layer of the gas, he reasoned, carry their mass, momentum, and kinetic energy over a distance of the order of the mean free path l and thus communicate it through collisions to another layer of the gas. The order of magnitude of the net flux of these quantities in the direction Ox is $lu\partial(nq)/\partial x$, where q stands for m , mu_y , and $(1/2)mu^2$ respectively. The corresponding coefficients of diffusion, viscosity, and heat conduction are lu , $nmlu$, and $(P/T)lu$ up to numerical coefficients which Maxwell obtained by performing the implied free-path averages. George Stokes' values for the viscosity of air and Thomas Graham's for the interdiffusion coefficient of two gases in normal conditions yielded compatible estimates of the mean free path, about 10^{-7} m (the thermal conductivity was still unknown). This convergence naturally pleaded in favour of Maxwell's approach.

Maxwell's understanding of viscosity had a less welcome implication. Since the mean free path is inversely proportional to the density, it follows that internal friction does not depend on the density of the gas, a counter-intuitive result which 'startled' Maxwell. After deriving this and other unwanted consequences, he briefly doubted the overall validity of the kinetic theory. In 1866, however, he carefully measured the viscosity of gases with the help of his wife, and thereby confirmed the surprising independence of viscosity from density. This spectacular finding lent much credibility to the kinetic theory in general. At the same time, Maxwell's experiments contradicted the more specific hard-sphere model. The measured viscosity turned out to be proportional to the absolute temperature, whereas the model yielded proportionality to the square root of temperature (Maxwell, 1867).

In reaction to this difficulty, Maxwell switched to a different model, maintaining Clausius's general assumptions but replacing the hard spheres of 1860 with repulsive centres of force. He also gave up the mean-free-path method and turned to a more powerful approach to transport phenomena based on computing the number of collisions of various kinds of molecules and the resulting variations of average molecular properties (see the following subsection). The integration of the relevant equations required the knowledge of the perturbed velocity distribution, except when the collision force varied as the inverse of the fifth power of the distance. Maxwell determined the transport of mass, momentum, and kinetic energy in the gas in this special case. He thus retrieved the Navier-Stokes equation for the motion of viscous gases, and

obtained quantitative relations between viscosity, diffusion rate, and thermal conductivity. Viscosity turned out to be proportional to temperature, in agreement with his experiments. To his pleasure, nature seemed to have chosen the mathematically favourable case of the $1/r^5$ force law. Maxwell knew, however, that this law could apply neither to large intermolecular distances for which the force is attractive, nor to small distances for which the structure of molecules comes into play (Maxwell, 1867).

A few years later, Oscar Meyer's and Joseph Stefan's accurate gas-viscosity measurements yielded a temperature-dependence of viscosity at variance both with the $1/r^5$ and with the hard-sphere model. In the 1880s Maxwell's followers elaborated in vain on his transport coefficients. As we may retrospectively judge, this problem eluded their mathematical techniques and physical models.

25.2.4 The Maxwell-Boltzmann Law

A crucial element in the understanding of gases as statistical collections of particles was the equilibrium distribution of velocities. Maxwell and his most outstanding follower Ludwig Boltzmann therefore attempted to establish this distribution as firmly as possible and to generalize it to degrees of freedom other than translation. In 1866, Maxwell provided a new demonstration based on the aforementioned collisions-number approach. In order to determine this number, he considered the trajectory of one of the colliding molecules in the reference system of the other. Call b the distance (impact parameter) between the first asymptote of this trajectory and the second molecule, ϕ the azimuth of the plane of this trajectory, \mathbf{v}_1 and \mathbf{v}_2 the initial velocities of the two molecules. An encounter for which the impact parameter lies between b and $b + db$ and the azimuth lies between ϕ and $\phi + d\phi$ occurs within the time δt if and only if the first molecule belongs to the cylindrical volume $|\mathbf{v}_1 - \mathbf{v}_2| \delta t b db d\phi$. To the number of collisions of this 'kind' per unit time and in a unit volume of the gas, Maxwell gave the natural expression

$$d\nu = |\mathbf{v}_1 - \mathbf{v}_2| b db d\phi f(\mathbf{v}_1) d^3 v_1 f(\mathbf{v}_2) d^3 v_2,$$

where $f(\mathbf{v}) d^3 v$ is the number of molecules per unit volume in the velocity range $d^3 v$ around \mathbf{v} (Maxwell, 1867).

A sufficient condition for the velocity distribution to be stationary, Maxwell reasoned, is the equality of the collision number $d\nu$ with the number

$$d\nu' = |\mathbf{v}_1' - \mathbf{v}_2'| b db d\phi f(\mathbf{v}_1') d^3 v_1' f(\mathbf{v}_2') d^3 v_2'$$

of the same kind (b , ϕ) for which the final velocities are \mathbf{v}_1' within $d^3 v_1'$ and \mathbf{v}_2' within $d^3 v_2'$. Owing to the conservation of energy and momentum, for a given kind of collision the initial and final velocities are in a one-to-one correspondence that leaves $|\mathbf{v}_1 - \mathbf{v}_2|$ and $d^3 v_1 d^3 v_2$ invariant. Therefore, the equality of $d\nu$ and $d\nu'$ requires that

$$f(\mathbf{v}_1) f(\mathbf{v}_2) = f(\mathbf{v}_1') f(\mathbf{v}_2')$$

for any two velocity pairs such that $v_1^2 + v_2^2 = v_1'^2 + v_2'^2$ and $\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{v}_1' + \mathbf{v}_2'$. The only isotropic solution of this functional equation is Maxwell's law. It remains to be shown that the equality $d\mathbf{v} = d\mathbf{v}'$ is a necessary condition of stationarity. Maxwell's proof of this point was impenetrably concise, even for his most perspicacious readers.

In a Viennese context favourable to atomistics and British methods, Ludwig Boltzmann was naturally fascinated by Maxwell's memoir of 1867 on the dynamical theory of gases. In 1868 and 1871 he generalized Maxwell's collision-number approach to molecules that had internal degrees of freedom and responded to external forces such as gravitation. He thus arrived at the form $\alpha e^{-\beta E}$ of the stationary distribution for the molecular variables on which the energy E of a molecule depends. This law, now called the Maxwell-Boltzmann distribution, has been an essential element of statistical physics to this day. We will return to the means of Boltzmann's generalization (Boltzmann, 1868; 1871a).

25.2.5 The Problem of Specific Heats

The kinetic theory of gases makes it possible to calculate the specific heats of a gas from the energy distribution over the various degrees of freedom. Maxwell and Boltzmann's solution to this problem is the so-called equipartition theorem. In its most general form, this theorem states that to each quadratic term in the expression for the energy of a molecule corresponds an average energy $(1/2)kT$, where k is Boltzmann's constant. In particular, the equipartition implies Clausius's result of 1857 that for a monoatomic gas composed of N point-like particles, the total energy should be $(3/2)NkT$ and the specific heat at constant volume $(3/2)Nk$.

In his hard-sphere model of 1860, assuming the randomness of the impacts between gas molecules, Maxwell showed that collisions between two different sorts of molecules tended to equalize their kinetic energies. This implies the truth of Avogadro's hypothesis (following Clausius's aforementioned reasoning). Maxwell also proved that in the case of non-spherical, hard, elastic molecules, the collisions induce rotations with an average kinetic energy equal to that of translation. To his disappointment, the resultant specific heat was much higher than observed for most gases (Maxwell, 1860).

A few years later, Maxwell became convinced that no consistent mechanical model of the molecules could reproduce the observed specific heats of polyatomic gases. As Boltzmann proved in 1871, the Maxwell-Boltzmann law leads to an average energy $(1/2)kT$, for each quadratic term in the energy of a molecule. As Maxwell argued in 1875, this implies the value $1 + 2/r$ for the ratio γ of the specific heats at constant pressure and at constant volume, wherein r is the number of quadratic terms in the energy function of a molecule. If the degrees of freedom of rotation and vibration of polyatomic molecules are taken into account, the resulting value of γ differs widely from the experimental value. 'Here we are brought face to face with the greatest difficulty which the molecular theory has yet encountered', lamented Maxwell (Maxwell,

1875, p. 433). At the turn of the century, opinions varied on the seriousness of this difficulty. We now know that its solution requires quantum theory.

25.2.6 The Boltzmann Equation and the H-Theorem

By 1871 Boltzmann had proven the stationarity of the Maxwell-Boltzmann law but not its uniqueness. In order to fill this gap, in 1872 he traced the evolution of the velocity distribution of a gas from an arbitrary initial state to equilibrium. If the number of collisions occurring in a spatially homogenous gas is known, he reasoned, the evolution of its velocity distribution can be computed. The variation in a given short time δt of the number of molecules with the velocity \mathbf{v}_1 within $d^3\mathbf{v}_1$ is equal to the number of collisions for which the final velocity of one of the colliding molecules belongs to this velocity range, minus the number of collisions for which the initial velocity of one of the colliding molecules belongs to this velocity range. Using Maxwell's formulas for the collision numbers of direct and inverse collisions of a given kind, and taking into account the invariance of the relative velocity $|\mathbf{v}_1 - \mathbf{v}_2|$ and of the product $d^3\mathbf{v}_1 d^3\mathbf{v}_2$, this balance gives the simplest case of 'the Boltzmann equation',

$$\frac{\partial f(\mathbf{v}_1, t)}{\partial t} = \int [f(\mathbf{v}_1')f(\mathbf{v}_2') - f(\mathbf{v}_1)f(\mathbf{v}_2)] |\mathbf{v}_1 - \mathbf{v}_2| b db d\phi d^3\mathbf{v}_2,$$

where the velocities \mathbf{v}_1' and \mathbf{v}_2' are the final velocities in a collision with the initial velocities \mathbf{v}_1 and \mathbf{v}_2 , the impact parameter b , and the azimuthal angle ϕ (Boltzmann, 1872).

This equation completely determines the evolution of the distribution f from its initial value. It implies the stationarity of Maxwell's distribution, since the vanishing of the square bracket amounts to Maxwell's stationarity condition. Boltzmann then considered the function (originally denoted E)

$$H = \int f \ln f d^3\mathbf{v}.$$

As a consequence of the Boltzmann equation, this function is a strictly decreasing function of time, unless the distribution is Maxwell's. Hence Maxwell's distribution is the only stationary one, and any other distribution tends toward Maxwell's. This is the so-called H-theorem. Boltzmann further noted that the value of $-H$ corresponding to Maxwell's distribution was identical to Clausius's entropy. For other distributions, he proposed to regard this function as an extension of the entropy concept to states out of equilibrium, since it was an ever increasing function of time.

Boltzmann then generalized his equation to more general distributions and systems. The Boltzmann equation has become the central tool for deriving transport phenomena in statistical physics. In Boltzmann's times, however, it could only be solved for Maxwell's $1/r^5$ forces despite Boltzmann's brave efforts in the hard-sphere case. Only at the beginning of the twentieth century did efficient perturbative methods become available for solving the equation in more realistic cases.

25.2.7 Challenges to the Second Law

The bridge established by the kinetic theory between mechanics and thermodynamics not only had repercussions on mechanical thinking by introducing statistical notions into the mechanical description of molecular motion but also affected the understanding of thermodynamics in a similar way. Maxwell used the kinetic molecular theory to 'pick a hole' in the second law of thermodynamics and point to its statistical character. In a letter to Tait of December 1867 and in his *Theory of Heat* of 1871, he argued that a 'finite being' who could 'see the individual molecules' would be able to create a heat flow from a cold to a warm body without expense of work (Maxwell, 1995, pp. 331–332). The being—soon named 'Maxwell's demon' by William Thomson—could indeed control a diaphragm on the wall between warm and cold gas, and let only the swifter molecules of the cold gas pass into the warm gas. In discussions with William Thomson and William Strutt (Lord Rayleigh), Maxwell related this exception to the second law with another obtained by mentally reversing all molecular velocities at a given instant. 'The 2nd law of thermodynamics', he wrote to Strutt in 1870, 'has the same degree of truth as the statement that if you throw a tumblerful of water into the sea you cannot get the same tumblerful of water out again' (Maxwell, 1995, pp. 582–583). Later, Maxwell spoke of 'a statistical certainty' of the second law (Knott, 1911, pp. 214–215). In 1878 he remarked that the dissipation of work during the interdiffusion of two gases depended on our ability to separate them physically or chemically, and concluded: 'The dissipation of energy depends on the extent of our knowledge . . . It is only to a being in the intermediate stage, who can lay hold of some forms of energy while others elude his grasp that energy appears to be passing inevitably from the available to the dissipated state' (Maxwell, 1878, p. 646) (Klein, 1970b).

In 1871, the year before his publication of the H-theorem, Boltzmann emphasized that a mechanical interpretation of the second law of thermodynamics required probability considerations. He also noted that in the mechanical picture the energy of a system in contact with a thermostat fluctuated in time. Even earlier, in 1868, he had noted that special initial states of the gas, for instance one in which all molecules originally lay on the same plane, failed to reach equilibrium. In the introduction of his memoir on the Boltzmann equation, he insisted that 'the problems of the mechanical theory of heat [were] problems of probability calculus', and that the observed regularity of the average properties of a gas depended on the exceedingly large value of the number of molecules (Boltzmann, 1872, p. 317). Yet he stated the H-theorem in absolute terms: the function H 'must necessarily decrease' (Boltzmann, 1872, p. 344). In 1876 Boltzmann's Viennese colleague Joseph Loschmidt remarked that not every initial state of the gas satisfied the theorem. He noted, in particular, that the reversibility of the laws of mechanics implied that to every H-decreasing evolution of the gas system corresponded a reverse evolution for which H increased. Boltzmann's proof of the decrease of H thus seemed to contradict the mechanical foundation of the theory (Boltzmann, 1909, vol. 1, pp. 295, 297, 96; Loschmidt, 1876).

To this 'very astutely conceived' paradox (Boltzmann, 1877a, p. 117), Boltzmann replied (in the more intuitive case of the spatial distribution of hard spheres):

One cannot prove that for every possible initial positions and velocities of the spheres, their distribution must become more uniform after a very long time; one can only prove that the number of initial states leading to a uniform state is infinitely larger than that of initial states leading to a non-uniform state after a given long time; in the latter case the distribution would again become uniform after an even longer time. (Boltzmann, 1877a, p. 120)

Boltzmann's intuition, expressed in the modern terminology of micro- and macro-states, was that the number of microstates compatible with a uniform macrostate was enormously larger than that compatible with a non-uniform macrostate. Consequently, an evolution of the gas leading to increased uniformity was immensely more probable.

25.2.8 The Probabilistic Interpretation of Entropy

To this elucidation of Loschmidt's paradox, Boltzmann appended the remark: 'Out of the relative number of the various state-distributions one could even calculate their probability, which perhaps would lead to an interesting method for the computation of the thermal equilibrium' (Boltzmann, 1877a, p. 121). This is precisely what he managed to do a few months later (Boltzmann, 1877b). The probability he had in mind was proportional to the number of microstates corresponding to a given macrostate. Such a number is ill-defined as long as the configuration of the molecules can vary continuously. That Boltzmann could nonetheless conceive it depended on his peculiar understanding of the continuity introduced in calculus. In his view, integrals and differentials were only condensed expressions for sums with many terms and for small differences. Discrete objects and processes were more basic and more rigorously defined than continuous ones. Whenever Boltzmann faced difficult integrations or integro-differential equations, he studied their discrete counterparts to get a better grasp of the solutions. He did so for instance in his derivation of the H-theorem (Boltzmann, 1909, vol. 1, 84–86, 346–361).

Boltzmann started his state-probability considerations with a 'fiction' wherein the energy of a molecule can only be an integral multiple of the finite element ε . A list of N integers giving the number of energy elements for each molecule defines the microstate of the gas, or 'complexion'. The macrostate is the discrete version of the energy distribution: it is defined by giving the number N_i of molecules carrying the energy $i\varepsilon$ for every value of the integer i . The probability of a given macrostate is proportional to its 'permutability':

$$\Pi = N! / N_1! N_2! \dots N_i! \dots$$

For a given value of the total number $\sum N_i$ of molecules and of the total energy $\sum N_i i\varepsilon$, and in the Stirling approximation for factorials, the permutability is a maximum when N_i is proportional to $e^{-\beta i\varepsilon}$ (wherein β is the Lagrange multiplier associated with the constraint of fixed total energy). Boltzmann next replaced the

uniform division of the energy axis with a uniform division of the velocity space, and took the continuous limit of the distribution N_i . This procedure yields Maxwell's velocity distribution. For any distribution N_i the logarithm of the permutability is $-\sum N_i \ln N_i$ in the Stirling approximation (up to a constant), which tends to $-H$ in the continuous limit. Hence the function $-H$, or the entropy of a gas with a given velocity distribution, corresponds to the combinatorial probability of this distribution, as Boltzmann already suspected in his reply to Loschmidt.

In 1878 Boltzmann used the combinatorial approach to explain the existence of a mixing entropy for two chemically indifferent gases. In 1883, after reading Helmholtz's memoirs on the thermodynamics of chemical processes, he showed how his combinatorics, when applied to a reversible chemical reaction, explained the dependence of the equilibrium on the entropy of the reaction. In this context Helmholtz distinguished between 'ordered motion' that could be completely converted into work, and 'disordered motion' that allowed only partial conversion. Accordingly, Boltzmann identified the permutability with a measure of the disorder of a distribution. The mixing entropy thus became the obvious counterpart of increased disorder (Boltzmann, 1878; 1883).

25.2.9 The H-Curve

Boltzmann's probabilistic interpretation of the H-function failed to silence criticism of the H-theorem. In 1894, British kinetic theorists invited Boltzmann to the annual meeting of the British Association, in part to clarify the meaning of this theorem. One of them, Samuel Burbury, offered a terminological innovation: the 'molecular chaos', defined as the validity condition for Maxwell's collision formula. Burbury and Boltzmann also provided an intuitive understanding of this condition: it corresponds to the exclusion of specially arranged configurations, for instance, those in which the velocities of closest neighboring molecules point toward each other (it should not be confused with Helmholtz's molar notion of disorder). As long as the gas remains molecularly disordered, the H-function evolves according to the Boltzmann equation. Boltzmann did not entirely exclude ordered states. He even indicated that an initially disordered state could occasionally pass through ordered states leading to entropy-decreasing fluctuations. However, he judged such occurrences to be extremely improbable (Boltzmann, 1895a; Boltzmann, 1896; 1898; Burbury, 1894, vol. 1, pp. 20–21; Brush, 1976, pp. 616–626).

To this view, Boltzmann's British interlocutors opposed a refined version of the reversibility paradox: H-decreasing and H-increasing states of an isolated gas should be equally frequent, they reasoned, for they correspond to each other by time reversal. In order to elucidate this point, Boltzmann studied the shape of the real H-curve determined by molecular dynamics and discussed its relation with the variations of H given by the Boltzmann equation. The real curve results from the cumulative effect of the rapid succession of collisions in the gas. It therefore has an extremely irregular shape, and does not admit a well-defined derivative dH/dt . The refined paradox

of reversibility fails, because it implicitly identifies the decrease of H with the negative sign of its derivative. An accurate statement of Boltzmann's interpretation of the decrease of H reads: for an initial macrostate out of equilibrium and for a finite time of evolution, the number of compatible microstates for which H decreases is much higher than the number of compatible microstates for which H increases. This statement is perfectly time-symmetrical. It means that over a very long time H is for the most time very close to zero, and that the frequency of its fluctuations decreases very quickly with their intensity. Hence any significant value of H is most likely to be very close to a summit of the H-curve. From that point H may increase for some time, but this time is likely to be very short and to be followed by a mostly uniform decrease (Boltzmann, 1895b; Ehrenfest and Ehrenfest, 1911; Klein, 1970a).

The following year Max Planck's assistant Ernst Zermelo formulated another objection to the H-theorem based on Henri Poincaré's recurrence theorem. According to this theorem, any mechanical system (governed by Hamilton's equations) evolving in a finite space with a finite number of degrees of freedom returns, after a sufficiently long time, as close to its initial configuration as one wishes (except for some singular motions). The theorem, Zermelo and Planck argued, excluded any derivation of the entropy law from a mechanical, molecular model. With obvious lassitude, Boltzmann replied that his description of the H-curve was perfectly compatible with recurrences and yet agreed with the statistical validity of the second law because the recurrence times were far beyond human accessibility. Through a simple calculation he estimated these times to have some 10^{10} digits for a gas. He compared Zermelo to a dice player who would declare a dice to be false because he has never obtained a thousand zeros in a row (Poincaré, 1889; Boltzmann, 1896; 1898; Brush, 1976; Zermelo, 1896, pp. 627–639).

25.2.10 Reception of the Kinetic Theory

Zermelo's attack was a symptom of the hostility of many physicists to atomistic considerations. Influential leaders of German physics such as Gustav Kirchhoff, Helmholtz, and Planck favoured a purely macroscopic physics based on differential equations ruling observable quantities. Experimental predictions of the kinetic-molecular theory could often be rederived by purely macroscopic methods. The only important exceptions were the convergent determinations of the mean free path from viscosity, diffusion, and heat conduction, and the compatible estimates of Avogadro's number by Loschmidt (in 1865), Thomson (in 1870), and others from the mean free path and from other phenomena (Brush, 1976, pp. 76–77). To these successes, Boltzmann would have added his explanation of the specific heats of monoatomic and diatomic gases in terms of rigid molecules with spherical and cylindrical symmetry. This was, however, a controversial achievement. Maxwell himself believed that the elasticity of Boltzmann's molecules implied energy equipartition over all their degrees of freedom, including internal vibrations. In sum, empirical success

could hardly justify the mathematical and conceptual difficulty of the kinetic theory (Boltzmann, 1876; Brush, 1976; Maxwell, 1877, 356–362).

In Germany, Clausius and Meyer were the only important investigators of the kinetic theory and its experimental consequences. In France, the persistence of Ampère's vibrational theory of heat and the growing distaste for molecular theories prevented the early acceptance of Maxwell's and Boltzmann's ideas. Emile Verdet's lectures long remained the only exception, until at the turn of the century Marcel Brillouin, Poincaré, and Emile Borel entered the scene. The most favourable ground was Britain, owing to the general enthusiasm for dynamical theories. John Tyndall's best-selling *Heat as a Mode of Motion*, published in 1862, popularized kinetic concepts very efficiently. In 1876, Henry Watson's valuable *Treatise on the Kinetic Theory of Gases* appeared. By 1890, British activity was flourishing in this field. In Austria, the traditional interest in molecular theories eased the spread of Boltzmann's theory and prompted Stefan's relevant experimental research. Yet the kinetic theory still lacked a full-fledged account of its conceptual and technical foundations (Brush, 1976; Principe, 2008).

The founders themselves disagreed on the status and achievements of their theory. They did not develop a canonical presentation of its core that could serve as the foundation of a common tradition. Maxwell never wrote a treatise on the kinetic theory. In a letter to Tait he ironically commented on Boltzmann's writings: 'By the study of Boltzmann I have become unable to understand him. He could not understand me on account of my shortness and his length was and is an equal stumbling block to me' (Maxwell, 1995, pp. 915–916). Boltzmann's *Lectures on Gas Theory* came late, in the 1890s, and they covered only some aspects of his and Maxwell's work (Boltzmann, 1896; 1898). As a result, the field remained open for the extensions and reinterpretations that led to modern statistical mechanics.

25.3 FROM KINETIC THEORY TO STATISTICAL MECHANICS

25.3.1 Statistical Mechanics as a New Perspective

Statistical mechanics provides tools for analyzing thermal processes not only in gases but also in general physical systems with microscopic degrees of freedom, whatever their precise constitution may be. At the heart of statistical mechanics is the notion of a virtual ensemble of macroscopic systems, all of which are characterized by the same dynamics but which vary in their initial microscopic configuration. Instead of tracking the statistical behaviour of atomistic constituents of a macroscopic system, statistical mechanics studies the properties of such an ensemble. Different kinds of thermodynamic systems in equilibrium are represented by different statistical

ensembles—an isolated thermodynamic system by a 'microcanonical ensemble' in which all members have the same energy; a system which is in contact with a heat reservoir by a 'canonical ensemble', in which the energies of the members are distributed according to an exponential law. Thermodynamic properties are then derived by taking ensemble averages. Due to its generality, statistical mechanics can be employed in classical and, with appropriate modifications, also in quantum physics. For this reason, it played a key role in the transition from classical to modern quantum physics.

Practically all building blocks of statistical mechanics can be found in the numerous publications of Maxwell and Boltzmann. They there appear, however, under perspectives different from that of statistical mechanics as we understand it today. It was only in Josiah Willard Gibbs' *Elementary Principles in Statistical Mechanics*, published in 1902, that a first coherent and autonomous form of statistical mechanics was presented. The now standard terminology 'microcanonical ensemble' and 'canonical ensemble' is due to Gibbs. In the same year, 1902, Albert Einstein published the first of a series of three articles on statistical physics which established, independently of Gibbs, another form of statistical mechanics. This work provided the basis for Einstein's exploration of quantum systems, and also for his analysis of Brownian motion and other fluctuation phenomena as evidence for the existence of atoms. In the same period, the kinetic theory was taken up and further developed, with the result that at the beginning of the twentieth century a variety of approaches were available for dealing with the most diverse problems of statistical physics.

This development of statistical methods was stimulated by the necessity to integrate the growing knowledge of atomistic and statistical processes such as ionic conduction, electronic conduction, and heat radiation. These new contexts of application shifted the emphasis within the kinetic theory of heat from Maxwell's and Boltzmann's questions concerning mechanical foundations to the problem of treating general physical systems in thermal equilibrium. Following this change of perspective, the results of the kinetic theory assumed a new meaning as cornerstones of a more broadly conceived statistical physics. In the following, we sketch the genesis of this new kind of physics by first recapitulating its roots in 'global approaches' to the kinetic theory of Maxwell and Boltzmann, by then discussing the extensions of the kinetic theory to new phenomena, and finally by presenting Gibbs' and Einstein's achievements as the result of a reflection on this development.

25.3.2 Global Approaches

In its earlier and simpler form, the kinetic theory of gases rested on molecular statistics. Boltzmann and Maxwell also developed approaches based on the consideration of the probability of the configurations of the whole system. Whereas the former approach is more intuitive and lends itself to the study of irreversible processes, the latter yields powerful methods for deriving the equilibrium properties of very general systems. Unfortunately, there is no unique, straightforward way to define and derive

the global probabilities. This explains why Boltzmann's pioneering considerations on such probabilities were followed by a variety of reinterpretations, including Gibbs' and Einstein's statistical mechanics.

The first example of a global approach is found in Boltzmann's fundamental memoir of 1868. There Boltzmann introduced the distribution $\rho(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N; \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N)$ such that the product $\rho d\sigma$ gives the fraction of time spent by a system of N point atoms around the phase $(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N; \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N)$ within

$$d\sigma = d^3r_1 d^3r_2 \dots d^3r_N d^3v_1 d^3v_2 \dots d^3v_N$$

after a very long time has elapsed. He then considered the element $d\sigma'$ in phase space made of the phases that the system takes after evolving during a constant time τ from any phase of $d\sigma$. The fraction of time spent by the system in these two elements is obviously equal:

$$\rho d\sigma = \rho' d\sigma'.$$

Since by Liouville's theorem (which Boltzmann rediscovered in this circumstance)

$$d\sigma = d\sigma'$$

and since the time τ is arbitrary, the distribution ρ is constant along the trajectory of the system in phase space. Boltzmann further assumed that the trajectory of sufficiently complex systems filled the energy shell, except for special, highly unstable configurations. This is what the Ehrenfests later called the ergodic hypothesis. Consequently, the distribution ρ must be the uniform distribution over the energy shell, which Gibbs later called microcanonical (Boltzmann, 1868).

From this distribution of the global system, one can derive the distribution of any subsystem that is weakly (thermally) coupled to the complementary system by integrating over the variables of the latter system. When the subsystem is relatively small, its distribution function is approximately proportional to $e^{-\beta H}$, where H is the energy of the subsystem as a function of its phase. The complementary system plays the role of a thermostat, whose absolute temperature is $1/\beta$ in proper units. Boltzmann obtained this general result, corresponding to Gibbs' later canonical law, in 1871. It of course contains the Maxwell-Boltzmann law as the particular case for which the subsystem is reduced to a single molecule (Boltzmann, 1871b, pp. 284–287).

In the same year, 1871, Boltzmann offered an 'analytical proof of the second principle of the mechanical theory of heat' based on the canonical distribution of a system in contact with a thermostat. The idea was to identify the internal energy of the system with the time-average $\langle H \rangle$ of its energy-function H , and the work provided to the system during an infinitesimal change of the external conditions with the time-average of the corresponding change of its potential energy. Then, the heat exchanged with the thermostat is

$$\delta Q = d\langle H \rangle - \langle dH \rangle$$

when measured in energy units. The averages are based on the distribution

$$\rho = e^{-\beta H}/Z,$$

wherein

$$Z = \int e^{-\beta H} d\sigma.$$

Then the product $\beta \delta Q$ is easily seen to be the differential of $\beta \langle H \rangle + \ln Z$. In other words, there exists an entropy function, which Boltzmann rewrote the following year as

$$S = - \int \rho \ln \rho d\sigma.$$

In this circumstance, Boltzmann emphasized the necessity of probability considerations to derive an expression of the entropy on a mechanical basis (Boltzmann, 1871c; Darrigol, 2003).

Despite these interesting results, Boltzmann quickly came to doubt the validity of the underlying hypothesis of ergodicity. He returned to the kinetic-molecular approach, which he believed to be better founded though less general. In 1879, Maxwell revived the global approach and gave it a new twist. Whereas Boltzmann had reasoned in terms of temporal probabilities for a single system, Maxwell adopted the 'statistical specification' of a system, in which the equilibrium properties of a thermodynamic system are to be compared not with those of a single mechanical system but with those of a stationary ensemble of such systems. He proved the stationarity of the microcanonical ensemble for any Hamiltonian dynamics and derived the Maxwell-Boltzmann distribution and energy equipartition from this ensemble. But he did not explain why stationary ensembles represented the thermal properties of macroscopic bodies. He regarded this property as a plausible assumption, to be tested by experiment and perhaps to be justified someday by ergodicity (Maxwell, 1879).

Boltzmann, who welcomed Maxwell's contribution, tried to fill this gap with a heuristic argument. Since, he pointed out, the observed, time-averaged behaviour of a thermodynamic system does not depend on its initial microscopic configuration, almost every system of an ensemble must lead to the same time-averages. Therefore, the time-average of a single system can be replaced with the ensemble average of the time average, which is also the time average of the ensemble average. Hence, for a stationary ensemble, the ensemble average should be equal to the time average. In a non-rigorous manner, this argument justified the use of statistical ensembles as mechanical models of thermodynamics (Boltzmann, 1881).

Yet Boltzmann remained open to other possibilities for constructing mechanical models of thermodynamics. For example, in the 1880s Boltzmann took up Helmholtz's analogy between the so-called 'monocyclic mechanical systems' and thermodynamic systems. Helmholtz had not attempted a realistic description of the mechanical processes responsible for thermodynamic observations. He rather considered simple mechanical systems, with but a few degrees of freedom, for which a function with the properties of the entropy function could be identified. Boltzmann's interest in this approach makes it evident that his principal aim was to identify mechanical analogies apt to clarify the relation between mechanics and thermodynamics. In a memoir of 1885 he introduced and discussed, under the name 'holode',

what later came to be called the canonical ensemble. But whereas in modern statistical mechanics the canonical ensemble serves to explore physical properties of quite general systems in thermal equilibrium at constant temperature, Boltzmann merely conceived this ensemble as a mechanical analogue on the same footing as Helmholtz's monocyclic systems. This illustrates a more general fact: although most of the elements of statistical mechanics were anticipated in the work of Boltzmann and Maxwell, they were there embedded in contexts different from that of modern statistical mechanics (Boltzmann, 1885; Klein, 1972a; Renn, 1997).

25.3.3 Beyond Gases

An essential motor of the transition toward modern statistical mechanics was the application of the kinetic theory to systems other than gases. In the limited context of gases, kinetic theory served mostly as a mechanical model for thermodynamics and transport phenomena. Despite a few empirical successes, it retained a precarious character, as is evident from the controversies in which Boltzmann was involved. Yet toward the end of the century it was successfully applied to newer domains of physics involving ions and electrons. In this process the foundations of the kinetic theory were gradually secured and widened.

A first example is the theory of the Dutch physicist Johannes Diderik van der Waals, which represented the first successful attempt at a quantitative understanding of the liquid-gas transition. In 1857, Clausius had already given a qualitative explanation of the three states of matter and sketched a kinetic picture of the transitions between them. Clausius's and Maxwell's kinetic theory, however, was essentially confined to ideal gases, for which the average distance between two molecules is much larger than their diameter, and for which the intermolecular forces are negligible between collisions. In 1873, van der Waals relaxed these two conditions, with the ambition of creating a kinetic theory of liquids. He assumed hard, elastic, spherical molecules, with a rapidly decreasing mutual attraction. The effect of the latter force, he reasoned, boils down to an attraction of the molecules on the fluid's free surface toward its interior. On the one hand, this surface correction implies a correction a/V^2 to the pressure P that balances the dynamical effect of collisions on the surface. On the other hand, this dynamical effect departs from Clausius's ideal value $(1/3)Nmu^2/V$, because the finite extension of the molecules alters their flux near the wall in the proportion $(V - b)/V$, where b is four times the total volume of the spheres. The resulting equation of states is

$$(P + a/V^2)(V - b) = RT.$$

Van der Waals successfully applied this equation to a unified description of the liquid and gas states. This simple, non-rigorous model provides a sound intuition of the most evident properties of real fluids. Maxwell and Boltzmann both contributed to its spread (Brush, 1976; van der Waals, 1873, chap. 7, 11).

Another, later, example for the extension of the methods of the kinetic theory beyond gas theory is Planck's theory of heat radiation of 1896–1900. Planck originally belonged to the denigrators of the kinetic theory and believed that the thermal properties of radiation should be explained without statistical methods. This attitude was rooted in his faith in the absolute validity of the laws of thermodynamics. In his eyes, Boltzmann's interpretation of the entropy law could not be true, for it restricted the validity of this law to a statistical one. Instead, Planck believed that the diffusion of electromagnetic waves by electric resonators was an intrinsically irreversible process, from which the increase of entropy could be deduced. Under the pressure of Boltzmann's criticism, however, Planck soon came to admit the necessity of an additional assumption of 'natural radiation' as the counterpart of molecular chaos, but without accepting Boltzmann's statistical understanding of disorder. According to Planck, the intricacies of the invisible dynamics of resonators, or those of the walls of a gas container, permitted a strictly irreversible evolution of the system, in full harmony with his absolute conception of the second law of thermodynamics. Elementary disorder—a generic name for natural radiation and molecular chaos—thus became the central concept of a non-statistical understanding of the relation between micro- and macro-world (Darrigol, 1988).

Planck's reinterpretation of Boltzmann's kinetic-theoretical reasoning did not stop there. In 1900 his further exploration of heat radiation brought him to apply Boltzmann's relation between entropy and probability, albeit only in a formal way that prevented revolutionary conclusions (Planck, 1900). He justified his new formula for the equilibrium spectrum of thermal radiation ('Planck's black-body law') by means of Boltzmann's counting of complexions, reinterpreted as a quantitative estimate of elementary disorder among resonators of the same frequency ν . This procedure allowed finite energy-elements $h\nu$ to enter the calculation without contradicting the continuous dynamics of the resonators nor the continuous nature of radiation. In this context Planck wrote the formula for the probability interpretation of entropy,

$$S = k \ln W,$$

which can now be read on Boltzmann's grave. Yet he did not admit the statistical validity of the entropy law until 1914. Einstein and Paul Ehrenfest were the first to treat black-body radiation as a thermo-statistical system and thus to arrive at its incompatibility with the laws of classical physics.

Toward the close of the century, the interest in Maxwell's and Boltzmann's kinetic theory rose considerably owing to the ever growing significance of atomistic ideas and statistical methods at the forefront of physics. Prominent examples are Hendrik Antoon Lorentz's electron theory (an atomistic version of Maxwell's electromagnetism), Paul Drude's electron theory of metals, the ionic theory of electrolytic conduction, the kinetic theory of solutions, and the use of atomistic models in inorganic and organic chemistry. Novel opportunities offered themselves for experimental research: there were new kinds of radiation, either waves or elementary particles; new interactions between matter and radiation such as the Zeeman and photoelectric effects, which confirmed atomistic models of matter or suggested new

ones; new studies of colloidal suspensions that seemed to reveal molecular motions, and so on. Measures of Avogadro's number could be gained from as many different sources as the study of capillarity, the kinetic theory of gases, experiments with thin layers, and, surprisingly, also from the theory of black-body radiation. The agreement between these measures not only increased confidence in atomistic hypotheses but also made it seem urgent to develop molecular theories bridging the diverse domains of physics (Renn, 1997).

25.3.4 Gibbs' and Einstein's Formulations of Statistical Mechanics

Early in the twentieth century, Gibbs and Einstein proposed two versions of statistical mechanics. They were both convinced that their work did not constitute a fundamentally new contribution but merely a change of perspective on what Maxwell, Boltzmann, and others had achieved. Yet their approaches introduced a new emphasis and also a conceptual organization different from that of the earlier global approaches.

Gibbs' perspective was more mathematical and more abstract than that of his predecessors. As a witness of the late nineteenth-century multiplication of molecular theories, he grew suspicious of special molecular assumptions and instead sought the most general foundation for statistical physics. His strategy was to develop the study of statistical ensembles for their own sake, and then to look for analogies between the properties of such ensembles and the laws of thermodynamics. In most of his elegant *Principles of Statistical Mechanics* of 1902, he described the underlying mechanical system in a formal manner, by generalized coordinates subjected to Hamilton's equations, for he agreed with Maxwell that the violations of energy equipartition made the foundations of molecular dynamics insecure. He introduced and systematically studied the three fundamental ensembles of statistical mechanics: the micro-canonical, the canonical, and the grand-canonical ensemble (in which the number of molecules may vary). He examined the relations between these three ensembles and their analogies with thermodynamic systems, including fluctuation formulas. Unlike Boltzmann, he did not seek an *a priori* justification for the thermodynamic significance of his ensembles. His approach was essentially axiomatic (Gibbs, 1902; Klein, 1972b).

Albert Einstein, who was not aware of Gibbs' *Principles*, approached statistical thermodynamics from still a different point of view, with partially overlapping results but decisively new aims. In his student years, he had been impressed by the rise of molecular physics, despite the skepticism of many German physicists. He especially admired Drude's theory of metals of 1900, which combined Lorentz's electron theory with Boltzmann's kinetic theory. On the basis of his experiences with the new applications of kinetic-theoretical methods, including Planck's work on heat radiation, he perceived Boltzmann's approach as too focused on the mechanical foundation of thermodynamics and too little oriented towards new evidence for the atomistic constitution of matter and radiation. From his perspective, there were gaps

in Boltzmann's methods—more so because he only had a fragmentary knowledge of them, mainly through the *Lectures on Gas Theory* (Boltzmann, 1896–1898). Einstein's articles of 1902–04 aimed at providing a new foundation for statistical mechanics, and it was here that he derived the second law of thermodynamics from the canonical ensemble with slowly varying external forces and justified the canonical ensemble as a small sub-ensemble of the micro-canonical ensemble. He interpreted all probabilities physically, as measures of the fraction of time spent by the system in various configurations (Einstein, 1902; 1903; 1904; Renn, 1997).

What distinguished Einstein from his predecessors was a difference of emphasis which turned out to be most consequential. Boltzmann, for instance, had displayed his ability to retrieve the macroscopic laws of thermodynamics and played down the departures from these laws that the molecular structure of matter implied. Einstein did the reverse. He wanted to use statistical mechanics to prove the molecular structure of matter and to probe the non-classical structure of radiation. For this reason, he focused on the fluctuations around equilibrium that were negligible for Boltzmann and inexistent for Planck. He interpreted the probability in the Boltzmann–Planck relation $S = k \ln W$ as the temporal frequency of the fluctuations of a system, and the constant k as the measure of its 'thermal stability'. In 1905 his analysis of Brownian motion showed how fluctuations could become observable on mesoscopic systems such as smoke particles. At the same time, he inverted Boltzmann's relation in order to explore unknown aspects of the dynamics of microsystems. This is how he arrived at the light-quantum hypothesis. Whereas Maxwell and Boltzmann meant to provide a mechanical foundation of thermodynamics, Einstein used statistical mechanics to question this foundation (Büttner, Renn, and Schemmel, 2003).

25.3.5 The Boltzmann Legacy

In their development of statistical mechanics as a synthetic framework, Gibbs and Einstein selected and reinterpreted some aspects of Boltzmann's and Maxwell's work. They neglected other aspects that were essential to Boltzmann or Maxwell, for instance, the discussions of irreversible processes and transport phenomena. More faithful to Boltzmann was his disciple Paul Ehrenfest, who strove with his wife Tatiana to elucidate the interconnections between Boltzmann's various approaches and to clarify the relevant probabilistic assumptions. Their encyclopedia article of 1911 remains an instructive synthesis of Boltzmann's views. They shared his skepticism about the ergodic hypothesis, so much so that Paul imagined a connection between quantum properties and violations of ergodicity. They dwelt on irreversible processes, on the Boltzmann equation, and on the paradoxes of the H-curve. They reproached Gibbs with neglecting this part of Boltzmann's legacy, and hardly noted Einstein's contributions (Ehrenfest and Ehrenfest, 1911; Klein, 1970a).

Statistical mechanics has now become an essential part of the canon of physics. Yet some of its original diversity has survived. Moreover, the nature and solidity of its foundations remain controversial issues. Physicists and mathematicians still discuss the pertinence of the ergodic hypothesis or the split between reversible and

irreversible processes. Philosophers still reflect on the relevant notions of probability or on the arrow of time. A century later, statistical mechanics is as open and alive as Maxwell and Boltzmann wanted theories to be.

REFERENCES

- BOLTZMANN, LUDWIG (1868), Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten, *Wien. Ber.*, 58: 517–560, repr. in Boltzmann (1909), vol. 1, 49–96.
- (1871a) Über das Wärmegleichgewicht zwischen mehratomigen Gasmolekülen, *Wien. Ber.*, 63: 397–418; repr. in Boltzmann (1909), vol. 1, 237–258.
- (1871b) Einige allgemeine Sätze über Wärmegleichgewicht, *Wien. Ber.*, 63: 679–711; repr. in Boltzmann 1909, vol. 1, 259–287.
- (1871c) Analytischer Beweis des zweiten Hauptsatzes der mechanischen Wärmetheorie aus den Sätzen über das Gleichgewicht der lebendigen Kraft, *Wien. Ber.*, 63: 712–732; repr. in Boltzmann (1909), vol. 1, 288–308.
- (1872) Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen, *Wien. Ber.*, 66: 275–370; repr. in Boltzmann (1909), vol. 1, 316–402.
- (1876) Über die Natur der Gasmoleküle, *Wien. Ber.*, 74: 553–560; repr. in Boltzmann (1909), vol. 2, 103–110.
- (1877a) Bemerkungen über einige Probleme der mechanischen Wärmetheorie, *Wien. Ber.*, 75: 62–100; repr. in Boltzmann (1909), vol. 2, 112–148.
- (1877b) Über die Beziehung zwischen dem zweiten Hauptsatz der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht, *Wien. Ber.*, 76: 373–435; repr. in Boltzmann (1909), vol. 2, 164–223.
- (1878) Über die Beziehung der Diffusionsphänomene zum zweiten Hauptsatz der mechanischen Wärmetheorie, *Wien. Ber.*, 78: 733–763; repr. in Boltzmann (1909), vol. 2, 289–317.
- (1881) Referat über die Abhandlung von J.C. Maxwell 'Über Boltzmanns Theorem betreffend die mittlere Verteilung der lebendigen Kraft in einem System materieller Punkte (1879) in *Beiblätter zu Annalen der Physik*, 5: 403–417; repr. in Boltzmann (1909), vol. 2, 582–595.
- (1883) Über das Arbeitsquantum, welches bei chemischen Verbindungen gewonnen werden kann, *Wien. Ber.*, 88: 861–896.
- (1885), Über die Eigenschaften monozyklischer und anderer damit verwandter Systeme, *Journal für die reine und angewandte Mathematik*, 98: 68–94; repr. in Boltzmann (1909), vol. 3, 122–152.
- (1895a), Nochmals das Maxwellsche Verteilungsgesetz der Geschwindigkeiten, *Annalen der Physik*, 55: 223–224.
- (1895b), On certain Questions of the Theory of Gases, *Nature*, 51: 413–415; repr. in Boltzmann (1909), vol. 3, 535–544.
- (1896), Entgegnung auf die wärmetheoretischen Betrachtungen des Hrn. E. Zermelo, *Annalen der Physik*, 57: 773–784; repr. in Boltzmann (1909), vol. 3, 567–578.
- (1896–1898), *Vorlesungen über Gastheorie*, 2 vols. (Leipzig, Barth).
- (1898), Über die sogenannte H-Kurve, *Mathematische Annalen*, 50: 325–332; repr. in Boltzmann (1909), vol. 3, 629–637.
- (1909), *Wissenschaftliche Abhandlungen* 3 vols. (Leipzig, Barth).
- (1981–1998), *Gesamtausgabe*, R. Sexl (ed.) (Braunschweig, Vieweg).
- BURBURY, SAMUEL (1894), Boltzmann's minimum function, *Nature*, 51: 78.
- CLAUSIUS, RUDOLPH (1857), Über die Acht der Bewegung, welche wir die Wärme nennen, *Annalen der Physik*, 100: 353–380.

- (1858), Über die mittlere Länge der Wege, welche bei der Molecularbewegung gasförmiger Körper von den einzelnen Molekülen zurückgelegt werden; nebst einigen anderen Bemerkungen über die mechanische Wärmetheorie, *Annalen der Physik*, 105: 239–258.
- (1888), *Die kinetische Theorie der Gase* (Braunschweig, Vieweg).
- EHRENFEST, PAUL, and TATIANA EHRENFEST (1911), Begriffliche Grundlagen der statistischen Auffassung in der Mechanik, *Encyklopädie der mathematischen Wissenschaften*, IV/2.II: 1–90.
- EINSTEIN, ALBERT (1902), Kinetische Theorie des Wärmegleichgewichtes und des zweiten Hauptsatzes der Thermodynamik, *Annalen der Physik*, 9: 417–433.
- (1903), Eine Theorie der Grundlagen der Thermodynamik, *Annalen der Physik*, 11: 170–187.
- (1904), Zur allgemeinen molekularen Theorie der Wärme, *Annalen der Physik*, 14: 354–362.
- GIBBS, JOSIAH WILLARD (1902), *Elementary Principles in Statistical Mechanics Developed with Especial Reference to the Rational Foundation of Thermodynamics* (New York, Scribner's sons).
- KNOTT, C. G. (1911), *Life and Scientific Work of Peter Guthrie Tait* (Cambridge, Cambridge University Press).
- LOSCHMIDT, JOSEPH (1876), Über den Zustand des Wärmegleichgewichtes eines Systemes von Körpern mit Rücksicht auf die Schwerkraft, *Wien. Ber.*, 73: 128–142.
- MAXWELL, JAMES CLERK (1860), Illustrations of the dynamical theory of gases, *Philosophical magazine*, 19: 19–32 and 20: 21–37.
- (1867), On the dynamical theory of gases, *Philosophical Transactions of the Royal Society*, 157: 49–88.
- (1871), *Theory of Heat* (London, Longmans).
- (1875), On the dynamical evidence of the molecular constitution of bodies, *Nature*, 11: 357–359, 374–377.
- (1877), The kinetic theory of gases, *Nature*, 16: 242–246.
- (1878), Diffusion, *Encyclopedia Britannica*; repr. in Maxwell 1890, vol. 2, 625–646.
- (1879), On Boltzmann's theorem on the average distribution of energy in a system of material points, *Transactions of the Cambridge Philosophical Society*, 12: 547–570.
- (1890), *The Scientific Papers of James Clerk Maxwell* 2 vols. (Cambridge, Cambridge University Press).
- (1986), *Maxwell on Molecules and Gases*, ed. E. Garber, S. G. Brush, and C. W. F. Everitt (Cambridge, Mass., MIT Press).
- (1995), *The Scientific Letters and Papers of James Clerk Maxwell* (Cambridge, Cambridge University Press), ii.
- PLANCK, MAX (1900), Zur Theorie des Gesetzes der Energievertheilung im Normalspektrum, Deutsche Physikalische Gesellschaft, *Verhandlungen*, 2: 237–245.
- POINCARÉ, HENRI (1889), Sur le problème des trois corps et les équations de la dynamique, *Acta mathematica*, 13: 1–270.
- VAN DER WAALS, JOHANNES DIDERIK (1873), *Over de Continuïteit van de Gas- en Vloeïstoftoestand* (Leiden, Stijthoff); English translation in *Physical memoirs* (London, Taylor and Francis, 1890).
- ZERMELO, ERNST (1896), Über einen Satz der Dynamik und die mechanische Wärmetheorie, *Annalen der Physik*, 57: 485–494.

STUDIES

- BRUSH, STEPHEN (1965–1972), *Kinetic Theory* (New York: Pergamon Press), 3 vols.
- (1976), *The Kind of Motion we call Heat: A History of the Kinetic Theory of Gases in the 19th century* (Amsterdam, North-Holland), 2 vols.

- (1983), *Statistical Physics and the Atomic Theory of Matter, from Boyle and Newton to Landau and Onsager* (Princeton, Princeton University Press).
- BÜTTNER, JOCHEN, JÜRGEN RENN, and MATTHIAS SCHEMMEL (2003), Exploring the limits of classical physics: Planck, Einstein, and the structure of a scientific revolution, *Studies in History and Philosophy of Modern Physics*, 34: 37–59.
- DARRIGOL, OLIVIER (1988), Statistics and combinatorics in early quantum theory, *Historical Studies in the Physical Sciences*, 19: 17–80; continued in *Historical Studies in the Physical Sciences*, 21 (1991): 237–298.
- (2003), The origins of the entropy concept, in J. Dalibard, B. Duplantier, V. Rivasseau (eds.), *Poincaré Seminar 2003: Bose–Einstein Condensation, Entropy* (Basel: Birkhäuser, 2004), 101–118.
- EVERITT, C. W. F. (1975), *James Clerk Maxwell. Physicist and Natural Philosopher*. Cambridge. Also as Maxwell, James Clerk in *Dictionary of Scientific Biography* (New York, Scribner's sons, 1970–80), ix: 198–230.
- FOX, ROBERT (1971), *The caloric theory of gases from Lavoisier to Regnault* (Oxford, Clarendon Press).
- GALLAVOTTI, GIOVANNI (1994), Ergodicity, ensembles, irreversibility in Boltzmann's and beyond, *Journal of Statistical Physics*, 78: 1571–1589.
- GARBER, ELIZABETH, STEPHEN BRUSH, and C. W. F. EVERITT (1986), Introduction to (Maxwell, 1986).
- KLEIN, MARTIN J. (1970a), *Paul Ehrenfest* (Amsterdam, North-Holland).
- (1970b), Maxwell, his demon, and the second law of thermodynamics, *American Scientist*, 58: 84–97.
- (1972a), Mechanical explanation at the end of the nineteenth century, *Centaurus*, 17: 58–82.
- (1972b), Gibbs, Josiah Willard, *Dictionary of Scientific Biography*, v: 386–393.
- (1973), The development of Boltzmann's statistical ideas, in E. G. D. Cohen and W. Thirring (eds.), *The Boltzmann Equation: Theory and applications* (Vienna, Springer), 53–106.
- PLATO, JAN VON (1991), Boltzmann's ergodic hypothesis, *Archive for the History of Exact Sciences*, 42: 71–89.
- (1994), *Creating Modern Probability: Its Mathematics, Physics and Philosophy in Historical Perspective* (Cambridge, Cambridge University Press).
- PORTER, THEODORE (1986), *The Rise of Statistical Thinking* (Princeton, Princeton University Press).
- PRINCIPE, JOÃO (2008), *La réception française de la mécanique statistique*, thèse de doctorat, Université Paris Diderot.
- RENN, JÜRGEN (1997), Einstein's controversy with Drude and the origin of statistical mechanics: A new glimpse from the 'love letters', *Archive for the History of Exact Sciences*, 51: 315–354.
- SKLAR, LAWRENCE (1993), *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics* (Cambridge, Cambridge University Press).

NOTE

1. An early version of this article was published as 'La nascita della meccanica statistica' in *Enciclopedia Italiana: Storia della scienza*, vol. 7: *L'ottocento* (Rome, Istituto della Enciclopedia Italiana 2003), 496–507.

CHAPTER 26

THREE AND A HALF PRINCIPLES: THE ORIGINS OF MODERN RELATIVITY THEORY

DANIEL KENNEFICK

26.1 INTRODUCTION

In 1900 the field theory of electromagnetism, which owes its origins primarily to the work of James Clerk Maxwell, had been under rapid development for two decades. In the 1880s a number of British physicists, beginning with Oliver Heaviside, had developed Maxwell's work into a successful body of theory which was able to explain a number of important features of electrodynamics. During the 1890s this new theory encountered some difficulties which, as Jed Buchwald, (1985) has shown, were connected with the earlier theory's inattention to the physical nature of the sources of the field, the moving charges themselves. This directed attention towards the problem of microphysics and the nature of the electron and towards a theory of electromagnetism which focused on the reality of charged particles as agents of the field. This was accompanied by a geographical shift away from Britain, whose leading figures came to play a less important role in the development of the theory, to the Continent, in particular to Holland and the German-speaking areas of Europe.

The new Continental theory had important successes, which inspired a hope that was expressed in the term electromagnetic world-view, that all physical phenomena would be expressible in terms of the electromagnetic field. In spite of major achievements by Hertz, Lorentz and others, the new theory still found itself troubled by a