Journal of Statistical Physics, Vol. 110, Nos. 3–6, March 2003 (© 2003)

Some Observations on the Early History of Equilibrium Statistical Mechanics

Cyril Domb[1]

Received January 24, 2002; accepted September 15, 2002

The article begins with some personal comments by the author on the outstanding contributions of Michael Fisher to statistical mechanics and critical phenomena. Its major aim is to trace the contributions of a number of pioneering personalities to the early history of equilibrium statistical mechanics. Four different areas are considered: (1) Classical Statistical Mechanics, (2) Quantum Statistical Mechanics, (3) Interacting Systems, and (4) The Ising Model. The article is concerned with the development and applications of statistical mechanics when certain basic assumptions are made. It does not deal with the justification of these assumptions which is a sophisticated discipline of its own.

KEY WORDS: Statistical mechanics; history; pioneers; critical phenomena; Ising model.

1. INTRODUCTION

Ten years ago when we celebrated Michael Fisher's 60th birthday I contributed an account(1) of Michael's activities at King's College London, before he moved to Cornell in 1966. The decade that followed was undoubtedly the most fruitful period of Michael's career, as well as the most creative and exciting in statistical mechanics in this century. Of supreme importance was Michael's interaction with Ken Wilson, and the seminars on critical phenomena which he ran jointly with Ben Widom at Cornell. Ken Wilson has acknowledged the key role which they played in his development of the Renormalization Group for critical phenomena.(2)

In 1981 the Wolf Prize in Physics was awarded jointly to Michael Fisher, Leo Kadanoff and Ken Wilson for this development. The Wolf Committee was following the precedent of the Nobel Committee for 1962

---

[1] Department of Physics, Bar-Ilan University, 52900 Ramat Gan, Israel; e-mail: dombcy@mail.biu.ac.il

in their award of the Prize for the discovery of the structure of DNA. In the wise words of Sir Lawrence Bragg(3) "It is a source of deep satisfaction to all intimately concerned that, in the award of the Nobel Prize in 1962, due recognition was given to the long, patient investigation by Wilkins at King's College (London) as well as to the brilliant and rapid final solution by Crick and Watson at Cambridge."

An enormous amount of patient and painstaking work by a number of members of the critical phenomena community provided a background vital for Wilson's brilliant innovation. It is not possible to award a prize to a research community, but the Wolf Committee selected the two leading researchers as its appropriate representatives, each of whom had made an essential contribution to Wilson's work.

In 1982 the Nobel Committee decided to ignore the critical phenomena community, and awarded the prize to Ken Wilson alone. I am convinced that the Wolf Committee got it right, and the prestigious Nobel Committee got it wrong.

Many research workers in critical phenomena felt that all major problems had been solved by the RG, and there was little of interest left to explore; they transferred their attention to other fields. Not so Michael, who continues to find areas which have been neglected, and to produce a steady stream of fascinating and significant new results.

I have always maintained that one of my greatest contributions to physics was to wean him away from his charismatic supervisor, Donald Mackay, under whom he prepared a Ph.D. thesis on analogue computers, and to interest him in the unsolved problems of statistical mechanics. Michael kindly informed me that my Inaugural Lecture entitled "Statistical Physics and its Problems" (*Science Progress* **43**:402 (1955)) played a major role in influencing him to change his field. In it I tried to survey briefly the statistical mechanics of non-interacting systems, and to describe the nature of the difficult problems which arise when interactions are taken into account. I paid particular attention to lattice models.

The aim of the present paper is to sketch a historical background to this lecture by identifying the pioneers who laid the foundations on which subsequent generations were able to build. An attempt will be made to assess the specific contribution of each of these "founding fathers."

2. CLASSICAL STATISTICAL MECHANICS

I spent the first part of 1958 as Visiting Professor in the University of Maryland, at the invitation of Elliott Montroll. Soon after I arrived Elliott delivered a lecture at the Bureau of Standards with the intriguing title "101 Years of Statistical Mechanics." In his opening remarks Elliott told us that he had first given the lecture in the previous year to celebrate the centenary of statistical mechanics; and he felt on principle that when he took the trouble to prepare a lecture he was entitled to use it at least twice.

It became evident that Elliott identified the initiation of statistical mechanics with the 1857 paper of Rudolf Clausius entitled "Über die Art der Bewegung welche wir Warme nennen(4)" (English translation "The Nature of the Motion which we call Heat"). In this contention he had the support of Willard Gibbs, who wrote in a tribute to Clausius(5)

> The origin of the kinetic theory of gases is lost in remote antiquity and its completion the most sanguine cannot hope to see. But a single generation has seen it advance from the stage of vague surmises to an extensive and well established body of doctrine. This is mainly the work of three men, Clausius, Maxwell, and Boltzmann, of whom Clausius was the earliest in the field and has been called by Maxwell the principal founder of the science. We may regard his paper (1857) "Über die Art der Bewegung, welche wir Warme nennen" as marking his definite entry into this field, although many points were incidentally discussed in earlier papers.

A good summary of the progress represented by this paper of Clausius has been given by Brush.(6) Clausius was able to give a qualitative description in terms of molecular motions of the solid, liquid and gaseous states of matter, of evaporation from liquid surfaces, and even of the origin of latent heat. He then undertook mathematical calculations relating to the motion of molecules in an ideal gas and estimated their velocities at 0°C for hydrogen, oxygen and nitrogen.

These values were criticized by a Dutch meteorologist, C. H. B. Buys-Ballot,(7) who pointed out that if the molecules moved as fast as Clausius claimed, the mixing of gases by diffusion should be much faster than had been observed. To meet this criticism Clausius wrote a second paper(8) in 1858 introducing molecular collisions into his model, with the new parameter of *mean free path* to characterize the distance traveled by molecules between successive collisions.

There is no doubt that the two Clausius papers represented a major step forward in the Kinetic Theory of Gases. But the role of probability and statistics in the reasoning was marginal; in the first paper the concept of a statistical balance determining equilibrium in evaporation from a liquid, and the need to average over all directions of molecular motion to determine the pressure of an ideal gas; in the second paper probability theory is used in connection with the calculation of the mean free path. The absence of a true statistical model is demonstrated by the assumption that all the molecules of an ideal gas move with the same velocity.

This fault was soon corrected by Maxwell in his address at the British Association meeting(9) in Aberdeen in September 1859.

"The author has established the following results:

1. The velocities of the particles are not uniform, but vary so that they deviate from the mean value by a law well known in the 'method of least squares.'

2. Two different sets of particles will distribute their velocities, so that their *vires vivae* will be equal."

Thus, Maxwell was already enunciating his famous velocity distribution law, and was on the way to the equipartition of energy which he developed shortly afterwards. He went on to consider non-equilibrium properties of gases, like diffusion.

The idea of applying statistical methods to physics was well formulated in a lecture(10) entitled "Molecules" which Maxwell gave to the British Association meeting at Bradford in 1873. I pointed out in my Inaugural Lecture that Maxwell's reasoning could, with one or two minor changes, be put forward with complete validity at the present day. When dealing with a large population the statistician ceases to focus attention on the properties of any one individual and concentrates instead on average properties. He divides the population into groups according to age, height, weight, hair colour or education, and is guided in his analysis (if he is a thinking statistician and not merely a collector of data) by the mathematical theory of probability. But in applying this theory he is usually beset by two worries: is his population really large enough for the method to be valid, and is his population thoroughly homogeneous or is one section influenced by factors which do not apply to the remainder? Neither of these worries exist for the population of molecules, and the population is in this sense a statistician's paradise. Knowing how individual molecules behave and react with one another, can we not apply the technique of the statistician to determine the average properties of the molecular population? These should then be identified with the properties of matter in bulk.

Brush has suggested(11) that Maxwell was alerted to probability and statistics early in his career. Whilst still a 19 year old student in Edinburgh he was strongly influenced by a long article published in June 1850 in the *Edinburgh Review* on Quetelet's work on statistics, and he wrote to his friend Lewis Campbell,(12)

> The true logic for this world is the Calculus of Probabilities ... as human knowledge comes by the senses in such a way that the existence of things external is only inferred from the harmonious (not similar) testimony of the different senses, understanding, acting by the laws of right reason, will assign to different truths (or facts, or testimonies, or what shall I call them) different degrees of probability ... .

Of Maxwell's many contributions to statistical physics we shall single out two for special mention. Maxwell wished to demonstrate that the second law of thermodynamics is not an exact law but statistical in character, with a non-zero probability of being contravened. His 1871 book, *Theory of Heat*, concluded with a section entitled "Limitations of the Second Law of Thermodynamics." In it he introduced "a being whose faculties are so sharpened that he can follow every molecule in its course, such a being, whose attributes are still as essentially finite as our own, would be able to do what is at present impossible to us." By opening or closing a trap-door appropriately to allow fast molecules to move from a vessel A to a vessel B, and slow molecules to move from B to A he could raise the temperature of B and lower the temperature of A in contradiction to the Second Law. The term *Maxwell's demon* was coined by William Thomson in 1874.

This product of Maxwell's fertile imagination has engaged the attention of leading scientists (e.g., L. Brillouin, L. Szilard, D. Gabor, R. Landauer, C. H. Bennett) almost continuously during the 130 years which followed. Just over 10 years ago a reprint collection was published on this topic;(13) and the opening paragraph of a long two-part research paper, published even more recently, provides a good summary of the situation.(14)

> Maxwell's demon is the source of a literature that has no apparent end—the most prestigious scientific journals, including *Physical Review* and *Physical Review Letters*, continue to publish articles on this topic. It is a topic that is considered to be important enough to merit editorials in *Nature*(15) and *New Scientist*.(16)

The second contribution illustrates Maxwell's remarkable intuition. In a lecture delivered to the Chemical Society(17) in 1875, Maxwell calculated the specific heat of a gas of molecules with n degrees of freedom. He found that the ratio γ of C_p the specific heat at constant pressure, to C_v the specific heat at constant volume, should be of the form $(2+n+e)/(n+e)$, where e is a positive quantity of unknown value relating to the stored energy of the molecule. For a monatomic molecule $n=3$, $e=0$ (no stored energy), $\gamma=5/3$, a value which Maxwell described as "too great for any real gas." (When monatomic gases were discovered some years later there was agreement with experiment.) For a diatomic molecule $n=6$, and the largest possible value of γ (corresponding to $e=0$) is 1.33; but the experimental values for several gases are found to be 1.408. Maxwell continued "I have now put before you what I consider to be the greatest difficulty yet encountered by the molecular theory."

Maxwell had put his finger on the major failure of the classical theory of the specific heats of gases; this was only corrected when the quantum theory allowed for *frozen in* modes which do not contribute to the specific heat.

Like Maxwell, the ideas of probability and statistics permeate Boltzmann's writings.(18, 19) It is reasonable to regard Boltzmann as the initiator of statistical mechanics as taught in the 20th century. Boltzmann(20) introduced an assembly consisting of a large number, N, of molecules of an ideal gas, whose interactions may be neglected, but which may be in a gravitational field. A *microstate* of the gas is defined by specifying the energies of individual molecules; classical theory allows these energies to vary continuously, but Boltzmann assumed that they were all multiples $J_1, J_2,..., J_N$ of a small discrete energy e which would ultimately be allowed to tend to zero. The total energy of the assembly E $(= Je)$ is constant, and can be distributed in a large number of different ways among the N molecules. Boltzmann made an important new basic assumption that all microstates (which he called *complexions*) have the same *a priori* probability. A state in which a specific molecule has energy ϵ_r $(= J_r e)$ can be called a *macrostate*. The probability P_r of the macrostate is then proportional to the number of microstates in which the remaining energy $E - \epsilon_r$ is distributed among the other $(N-1)$ molecules. This is a very large number, and using asymptotic approximations, Boltzmann was able to establish the relation which carries his name

$$P_r \propto \exp(-\epsilon_r/kT) \tag{1}$$

Boltzmann felt the need to justify his equal *a priori* probability assumption for microstates, and in 1871 he advanced what afterwards came to be known as the *ergodic hypothesis*(21): that a complex dynamical system with many degrees of freedom will over a long period spend the same amount of time in each microstate. The average value of any property of the system taken over such a long period will be equal to the average value taken over all microstates, and for the classical systems with which he was concerned this meant averaging over the appropriate phase space.

Some years later(22) Boltzmann formulated the method of maximum probability in the treatment of a gas which afterwards became the standard procedure in statistical mechanics. He defined a collection of microstates in which $n_0, n_1,..., n_p$ molecules had energies $(0, \epsilon_0, \epsilon_1, \epsilon_2,..., \epsilon_p)$, showed that it had statistical weight

$$\Omega = \frac{n!}{n_0!\, n_1! \cdots n_p!} \tag{2}$$

The equilibrium state is obtained by maximizing. Ω subject to the restrictions

$$n_0 + n_1 + \cdots + n_p = N \tag{3}$$

$$n_1 + 2n_2 + \cdots + pn_p = J \tag{4}$$

Although the ergodic hypothesis required the sum of all the terms (2), this could be replaced by the maximum term to the order of approximation required. He used the Stirling approximation for $\ln \Omega$, and Lagrange parameters to calculate the equilibrium values of the n_r

$$\bar{n}_r = \bar{n} \exp(-\beta\epsilon_r) \tag{5}$$

where $\bar{n}$ and β are determined by the restrictive conditions (3) and (4).

The connection with thermodynamics was achieved by means of the famous hypothesis for the entropy, S,

$$S = k \ln \Omega \tag{6}$$

Formula (6) is engraved on Boltzmann's tombstone; it was first introduced by Planck (including the new constant k); but the formula is latent in Boltzmann's work. It served as a landmark in enabling physicists to understand the nature of the mysterious entropy which had been introduced by the second law of thermodynamics. Boltzmann was guided by the fact that entropies of independent systems are additive, whereas probabilities are multiplicative.

A very substantial portion of Boltzmann's research was devoted to irreversible processes, (nowadays called non-equilibrium statistical mechanics) with which we are not concerned in this article. But mention should be made of the titanic struggle in which Boltzmann engaged to demonstrate how irreversible macroscopic behaviour can result from reversible microscopic equations of motion.(23)

The term statistical mechanics was first introduced by Gibbs in 1902 in the title of his book(24) "Elementary Principles in Statistical Mechanics." Boltzmann's method could be applied only to an assembly consisting of a large number N, of quasi-independent molecules representing an ideal gas. Gibbs introduced the new concept of an *ensemble* with which he was able to deal with an assembly of interacting molecules representing any physical system.

Instead of a single assembly of fixed energy E, an *ensemble* is a statistical collection of assemblies whose energy follows a probability distribution $g(E)\,dE$. Like Boltzmann, Gibbs accepted the ergodic hypothesis, and equilibrium properties are calculated by averaging over phase space

$$\langle E \rangle = \int E g(E)\, d\Omega, \qquad \langle E^2 \rangle = \int E^2 g(E)\, d\Omega \tag{7}$$

As in standard probability theory it can be shown that for a large variety of models the fluctuations are proportional to $N^{1/2}$, and become negligible compared with N for large N. $\langle E\rangle$ is then identified with the thermodynamic internal energy, U.

The choice of $g(E)$ determines the thermodynamic character of the ensemble. The most useful choice which Gibbs made was

$$g(E) \propto \exp -(E/kT) \tag{8}$$

which he termed the *canonical ensemble*, corresponding to a constant temperature T. This is because the loose connection of two independent canonical ensembles with energies E, E' to form a new canonical ensemble of energy $(E+E')$ will result in no change to the original ensembles;

$$g(E+E') = g(E)\, g(E') \tag{9}$$

Gibbs noted that Boltzmann's treatment, corresponding to a constant energy, is obtained by taking a probability distribution which is zero outside the range $(E, E+dE)$; in modern terminology it is a delta function, $\delta(E)$. He termed this the micro-canonical ensemble. From Boltzmann's hypothesis the thermodynamic function which result from averaging over phase space is

$$\int d\Omega = \exp(S/k) \tag{10}$$

where the entropy S is a function of (U, V). The statistical averages are more difficult to handle in view of the greater difficulty in using thermodynamics with the energy as an independent variable rather than the temperature.

To find the thermodynamic function which corresponds to the canonical ensemble it is necessary to define the pressure

$$\langle P\rangle = -\int \frac{\partial E}{\partial V}\, d\Omega, \tag{11}$$

and other analogous forces (X_i) and to calculate

$$dQ = dU + \langle P\rangle\, dV + \sum_i \langle X_i\rangle\, dx_i, \tag{12}$$

from the first law of thermodynamics. By identifying the result with the second law, Gibbs showed that

$$\int \exp(-E/kT)\, d\Omega = \exp(-F/kT) \tag{13}$$

where $F(T, V)$ is the Helmholtz free energy.

In a series of lectures on statistical thermodynamics which Schrödinger delivered$^{(25)}$ in Dublin in 1944, he introduced an alternative approach to the canonical ensemble. To keep an assembly at a constant temperature it should be loosely connected to a *heat bath*, a large insulated body. The constitution of the heat bath is irrelevant, and it could equally well consist of a very large number, **N**, of identical copies of the original assembly. The whole of the Boltzmann treatment can then be taken over, and Boltzmann's relation (1) applied to our particular assembly

$$g(E) \propto \exp -(E/kT) \tag{14}$$

There is no worry about fluctuations since **N** can legitimately be allowed to tend to infinity.

Schrödinger attributed this neat device to Gibbs, but I have been unable to trace it in any of Gibbs' publications. I suspect that it was an idea which occurred to him whilst pondering over Gibbs' treatment.

In dealing with the thermodynamics of heterogeneous substances Gibbs had found it useful to introduce the *chemical potential*, μ_i, of each species i, which plays a similar role in controlling chemical equilibrium to that of the temperature in controlling thermal equilibrium. Just as it is often convenient to use the temperature as an independent thermodynamic variable rather than the energy, so it is often convenient to use the chemical potential of a given species as a variable rather than its concentration. In formulating the statistical mechanics of heterogeneous bodies, Gibbs introduced the concept of a *grand canonical ensemble* in which an assembly of energy E, containing v_1 molecules of species 1, v_2 molecules of species 2,..., v_r molecules of species r, has probability distribution

$$\exp(\mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_r v_r - E/kT)\, dv_1 \cdots dv_r\, dE \tag{15}$$

Gibbs showed that when the distribution (15) is averaged over a generalized phase space the result is

$$\exp(\Phi/k) \tag{16}$$

where

$$\Phi(T, V, \mu_i) = PV/T \tag{17}$$

is a function from which all thermodynamic properties can be derived.

The contributions of Gibbs to statistical mechanics were fundamental and far-reaching; but like his basic contributions to thermodynamics, a substantial period of time elapsed before they were properly appreciated.

According to Pais,[(26)] Einstein was familiar with a substantial part of Boltzmann's work, but not with the hypothesis (6) which he obtained independently himself; when he became aware of Boltzmann's priority he referred to the formula as "Boltzmann's Principle." But Einstein's approach to the formula differed significantly from that of Boltzmann, and was really complementary to it.

Boltzmann's aim was to establish a theory of matter based on probability and statistics, and the formula served as a bridge back to thermodynamics. Einstein considered instead the calculation of fluctuations in thermodynamic systems and their experimental observation. Re-writing (6) in the form

$$\Omega = \exp(S/k) \tag{18}$$

and making use of thermodynamic information expressing the entropy as a function of state, it is possible to calculate the probability of deviations from equilibrium values. He applied this to the calculation of fluctuations in temperature, density, pressure, etc. in a finite system.

Consider a small region of a large reservoir, which can be considered attached to it in equilibrium. The suffix 0 will be used to refer to the reservoir, and it can be assumed that because it is very large the temperature T_0 and the pressure P_0 of the reservoir are constant; for convenience the number of molecules n in the small region can be kept constant allowing the volume to fluctuate. Consider a fluctuation of the small region, of temperature change ΔT, volume change ΔV, and entropy change ΔS, and internal energy change ΔU. The probability of this fluctuation, w, is given from (18) by

$$w \propto \exp(\Delta S + \Delta S_0)/k \tag{19}$$

where the total system, reservoir + small region is isolated at constant energy and volume. Thus

$$\Delta U + \Delta U_0 = 0, \qquad \Delta V + \Delta V_0 = 0 \tag{20}$$

But

$$\left(\frac{\Delta Q}{T_0}\right)_{\text{reservoir}} = \frac{\Delta U_0 + P_0\,\Delta V_0}{T_0} = -\frac{\Delta U + P_0\,\Delta V}{T_0}$$

Hence

$$w \propto \exp[-\beta(\Delta U - T_0)\,\Delta S + P_0\,\Delta V] \tag{21}$$

The expression (21) is the basic formula from which fluctuations in any thermodynamic variables can be calculated from standard formulae of thermodynamics using appropriate transformations.(27) Coefficients of first order terms must vanish since the system is in equilibrium; usually the coefficients of the second order terms are non-zero and the fluctuations are Gaussian; but where the coefficients vanish the behaviour will be non-Gaussian. When the independent variables are ΔT and ΔV the second order terms are

$$w \propto \exp\left[-\frac{C_{\mathrm{v}}}{2kT^2}(\Delta T)^2 + \frac{1}{2kT}\left(\frac{\partial P}{\partial V}\right)_T (\Delta V)^2 \right], \tag{22}$$

and the fluctuations in volume become large and non-Gaussian when $(\frac{\partial P}{\partial V})_T = 0$; this applies to a van der Waals critical point.

Einstein used these relations in his pioneering investigatons(28) of critical opalescense. (Smoluchowski derived similar results independently.(29))

3. QUANTUM STATISTICAL MECHANICS

It is usual to associate the origin of the quantum theory with Planck's 1900 papers(30) in which he successfully fitted the experimental measurements of black-body radiation by Lummer and Pringsheim(31) with his famous formula

$$U(v, T) = \frac{8\pi v^2}{c^3} \frac{hv}{[\exp(hv/kT) - 1]} \tag{23}$$

Kuhn has argued convincingly(32) from a detailed analysis of the sources that the path to this formula, and the subsequent emergence of quanta of energy, differ substantially from the picture conventionally presented. We shall draw on his analysis to outline the part which thermodynamics and statistical mechanics played in the development.

Thermodynamics was Planck's first love. His detailed thesis in 1879 was based on Clausius' work, and he subsequently followed his own individual path. I can remember when I started graduate studies how much inspiration I derived from his book(33) *Theory of Heat* which was the first of a number of outstanding text-books (e.g., Zemansky,(34) Pippard,(35) Callen(36)) which succeeded in demonstrating the self-contained logic and remarkable power of thermodynamic reasoning. Planck was also familiar with Boltzmann's work, with his statistical treatment of an assembly of loosely coupled systems, and with his formula (6) for the entropy.

By a fortunate coincidence Planck moved to Berlin in 1889 where improved experimental techniques were being used to determine the form of black body radiation. He decided to try to fit the latest experimental curve, using thermodynamic information available regarding black-body radiation. Starting from Wien's law for the temperature dependence of the energy density at wavelength λ,

$$U(\lambda, T) = b\lambda^{-5} \exp(-a/\lambda T); \tag{24}$$

he was led$^{(37)}$ in a search for the relation between U and S, to propose the simple differential equation

$$\frac{\partial^2 S}{\partial U^2} = -\alpha/U^2 \tag{25}$$

α being a constant. But when improved measurements showed weaknesses in Wien's law, he modified (25) to

$$\frac{\partial^2 S}{\partial U^2} = -\alpha/U(\beta + U)^2 \tag{26}$$

and this led to

$$U(\lambda, T) = d\lambda^{-5}/[\exp(D/\lambda T) - 1]. \tag{27}$$

This fitted the data extremely well, and in due course he rewrote (27) in the traditional form (23).

In trying to find a theoretical basis for his formula Planck used Boltzmann's method for an assembly of resonators, but the cell size in phase space was fixed at $h\nu$. Kuhn claims that "nothing in Planck's writing before 1906 reflects the idea of restricted resonator energy." Kuhn also points out that for Planck the most gratifying feature of his research was to have produced a new constant of nature, h; the second constant, k, in his formula he then termed Boltzmann's constant. It is worth noting that the relationship between the Rayleigh–Jeans formula,

$$U(\nu, T) = 8\pi k T \nu^2/c^3$$

derived from classical theory, and agreeing with experiment for small ν, and the Planck formula (23) is much more sophisticated$^{(38)}$ than in the conventional story of the "ultra-violet catastrophe."

In fact it was from Einstein's ideas$^{(39)}$ on particles of light and the photoelectric effect that the concept of a quantum of energy began to emerge. Einstein went on to formulate$^{(40)}$ his theory of the specific heats of solids in which he made it clear that oscillators can possess energy only in multiples of $h\nu$. He was able to account for the experimental fact that specific heats decrease to zero as $T \to 0$, whereas according to classical theory they should remain constant.

Incidentally this also formed the basis for the removal of the "greatest difficulty yet encountered by the molecular theory" noted by Maxwell and mentioned above (p. 479).

With the development of the old quantum theory came the idea that every atomic system has a discrete set of energy levels $\epsilon_0, \epsilon_1, \epsilon_2, \ldots, \epsilon_r$ with degeneracy weights $g_0, g_1, g_2, \ldots, g_r$. The methods of Boltzmann and Gibbs could be applied much more readily and directly to discrete levels, and there was no need to divide continuous phase space into small discrete regions. Planck played a prominent part in this development, and introduced the function$^{(41)}$

$$Z = \sum_{j=0}^{r} g_i \exp(-\epsilon_i/kT) \tag{28}$$

which he called the *Zustandsumme* (sum over states)—hence the letter Z. It is simply related to the Helmholtz free energy F,

$$F = -NkT \ln Z \tag{29}$$

In June 1924 Einstein received a letter from a young Bengali, Satyendra Nath Bose, whose new derivation of Planck's law had been rejected. Einstein was impressed by Bose's paper, *translated the paper personally into German*, and submitted it with a strong positive recommendation. Publication of the paper$^{(42)}$ transformed Bose's career. The derivation of Planck's law was based on a particle picture in which the number of particles is not conserved. Bose was led to Planck's formula, even though he does not seem to have been aware that he was using anything different from Boltzmann statistics.$^{(43)}$

The paper by Bose was followed by two papers by Einstein,$^{(44)}$ extending Bose's treatment to material particles whose number is conserved. He clarified the nature of the statistics that had been used for the *identical particles* showing that the Boltzmann combinatorial factor

$$\Omega = \frac{N!}{n_0!\, n_1! \cdots n_r!}\, g_0^{n_0} g_1^{n_1} \cdots g_r^{n_r} \tag{30}$$

is replaced by

$$\Omega = \prod_{i=0}^{r} \frac{(g_i + n_i - 1)!}{n_i!\,(g_i - 1)!} \tag{31}$$

This leads to the equilibrium values

$$\langle n_i \rangle = \frac{g_i}{\lambda \exp(\epsilon_i / kT) - 1} \tag{32}$$

where λ is determined by

$$\sum_{i=0}^{r} \langle n_i \rangle = N \tag{33}$$

For non-conserved particles $\lambda = 1$ and Bose had used a simple argument to evaluate g_i and to show that (32) leads to Planck's law (23).

In 1925 Pauli$^{(45)}$ introduced his exclusion principle which states that no two electrons can occupy the same quantum energy level. Fermi$^{(46)}$ and Dirac$^{(47)}$ realized independently that this hypothesis would lead to different statistics for electrons, the combinatorial formula (31) being replaced by

$$\Omega = \prod_{i=0}^{r} \frac{g_i!}{n_i!\,(g_i - n_i)!}, \tag{34}$$

by use of which (32) is replaced by

$$\langle n_i \rangle = \frac{g_i}{\lambda \exp(\epsilon_i / kT) + 1} \tag{35}$$

It later became clear that the wave-functions for identical particles in quantum mechanics must be either symmetric (S) or anti-symmetric (A); S particles satisfy (31) and because of the joint work they are reasonably termed Bose–Einstein statistics; A particles satisfy (34) and are similarly termed Fermi–Dirac statistics. Dirac was more exacting, and in his lectures, and his book$^{(48)}$ referred to the former as Bose statistics (presumably because the idea was due to Bose) and to the latter as Fermi statistics (presumably because Fermi's paper was presented several months before his); this also provides a justification for the terms bosons and fermions. Statistics of non-identical particles satisfy (30) and these are reasonably termed Maxwell–Boltzmann statistics, although a more exacting attitude might prefer Boltzmann statistics.

R. H. (later Sir Ralph) Fowler, perhaps the leading world authority on statistical mechanics in the 1920's and 1930's was the author of the first text book on statistical mechanics after that of Gibbs. Fowler started life as a mathematician, and he was dissatisfied with the Boltzmann method which replaced a sum by its maximum term without an estimate of the error involved. With C. G. Darwin he introduced(49) the use of generating functions from combinatorial analysis, selected the coefficient required by a contour integral, and evaluated it by the method of steepest descents; correction terms could readily be calculated. The method was later applied to Bose–Einstein and Fermi–Dirac statistics.

Fowler's book originated in an Adams Prize essay for the years 1923–1924; but statistical mechanics was being applied to so many new problems that the first edition(50) appeared only in 1929 consisting of 563 large pages. By the time the second edition appeared(51) in 1936 it had grown to 864 pages; and with the help of E. A. Guggenheim an adapted version(52) oriented more to the needs of physicists and chemists was published in 1939.

These books were extremely comprehensive in their coverage and widely used. When one talked in Cambridge about consulting Fowler, the intention was almost always the book rather than the man. Fowler was deeply involved in many research projects and in a variety of University activities, and writing the books must have been a great strain. When he passed away in 1944 at the early age of 55 the terse comment by E. A. Moelwyn-Hughes, a well known physical chemist, "Writing Fowler killed Fowler" summed up what many people were thinking.

It was Fowler(53) who borrowed the term *partition function* from the theory of numbers to replace the clumsy *Zustandsumme* of Planck; this was accepted universally, but the *Z* remained as a token reminder.

However, there are some features of Fowler which appear questionable, particularly when the subsequent development of statistical mechanics is taken into account. His attitude towards the Gibbs canonical ensemble is negative. The concept is mentioned only once in the following paragraph:

> There are two distinct starting points from which we may build up with equal success a theoretical model to represent the material systems of our more or less direct experience—the Gibbsian ensemble and the general conservative dynamical system. Of these the Gibbsian ensemble has perhaps the advantage in logical precision, in that the whole of the necessary assumptions can be explicitly introduced in the initial formulation of the "canonical" ensemble. For this reason it should perhaps be preferred, and is preferred by some theoretical physicists. But to others something more than success and logical rigour appears to be necessary for the acceptance of a model which is to account to our aesthetic satisfaction for the properties of matter. A certain "sanity," or physical reality, may be demanded in

> the initial postulates and in the details of the model; particularly in so far as they are to reproduce the well-known properties of matter. To these others the Gibbsian ensemble appears to be weak from this aspect, and they are led—in spite of logical and analytical incompleteness—to prefer the conservative dynamical system of many degrees of freedom as the more satisfactory model from which to derive (or attempt to derive) the properties of matter. This is the model, generalized to include quantized systems, which will be used in this monograph.

By confining himself to the micro-canonical ensemble Fowler was restricting the validity of his treatment to assemblies of quasi-independent systems, and interactions could not be taken into account. This does not seem to have been made very clear.

At a later stage Fowler did convert to the grand canonical ensemble, and wrote a paper(54) in which he introduced the *grand partition function* and enunciated its properties. A chapter of Fowler and Guggenheim is devoted to this topic,(55) and provides a theoretical justification for the Mayer treatment of interactions and condensation which is discussed in the book.

Fowler also engages in a polemic on the introduction of Boltzmann's hypothesis particularly in the approach used by Planck. He feels that the hypothesis should not be introduced *ab initio* by plausibility arguments, and be used as a bridge from statistical mechanics to thermodynamics. The bridge should be established *ab initio*, and Boltzmann's hypothesis derived as a theorem.

From the logical point of view this argument is valid, but Planck's approach has the advantage of giving the student an early insight into the nature of entropy. He can then understand how thermodynamic equilibrium is achieved at temperature T in the competition between the ordering quality of energy and the disordering quality of entropy.

Fortunately Fowler's graduate student, G. S. Rushbrooke, remedied the above defects in his text book(56) published in 1949. Rushbrooke lists books in the bibliography to which he is indebted which include Fowler and Guggenheim, and modestly adds "I only hope that the result of this present text will be that many others are enabled to enjoy these weightier publications, indispensable to serious research workers." Personally I found that although Rushbrooke avoided sophisticated mathematics, he was able to provide the average physics and chemistry student with all that he needs to master the techniques of statistical mechanics.

In 1946, as I started graduate studies, Schrödinger published a small book of 88 pages on statistical thermodynamics to which I have already referred.(25) I read it from cover to cover and found it extremely enlightening. In my view it certainly entitles the author to be regarded as a pioneer in the teaching of quantum statistical mechanics.

4. INTERACTING SYSTEMS

In my Inaugural Lecture I draw attention to the dramatic change in analytic behaviour to be expected when statistical mechanics takes interactions into account, and to the discontinuities in thermodynamic behaviour to be anticipated. For example, one would hope to account for the condensation of a gas, or the melting of a solid, or the behaviour of solid solutions near to their critical temperature. I concerned myself only with exact methods, since Onsager's solution of the 2-dimensional Ising model had cast grave doubts on the value of the different approximations which had been used previously. I suspect that it was the challenge of exact methods which Michael found particularly attractive.

But in a historical survey one must relate to closed form approximations, particularly since they subsequently found a respectable place as *exact* solutions for systems with long range forces. Fortunately I have dealt with them at some length in my book(57) *The Critical Point* for which Michael kindly contributed a Foreword. I can therefore summarize the most important contributions, and give precise references for the reader who seeks more information.

(a) In his thesis in 1873 van der Waals put forward the equation which bears his name, and with the help of Maxwell, a coherent theoretical account could be given of the behaviour of fluids, the relationship between the liquid and vapour states, and the critical point.(58)

(b) In 1907 Pierre Weiss suggested that in a ferromagnet the interaction of the elementary molecular magnets combine to provide an internal field which must be added to the external field to obtain the true field acting on an elementary magnetic dipole. With this hypothesis he was able to account for the basic features of ferromagnetism, the spontaneous magnetization at low temperatures which disappears at the Curie temperature, T_C; and the magnetic susceptibility for temperatures above T_C.(59)

(c) The concept of correlation in fluids was introduced by Ornstein and Zernike to deal with light scattering in fluids and critical opalescence.(60) This important work was largely neglected until Michael Fisher drew attention to it,(61) and showed how it must be modified to comply with Onsager's exact results. (Michael was greatly helped by Uhlenbeck who criticized a first draft in great detail).

(d) In discussing experimental results on order-disorder transitions in alloys, Bragg and Williams introduced a new concept of *long-range order*, correlations between atoms in a crystal which extend to infinity. These correlations decrease with temperature and become zero suddenly at a temperature which they termed the Curie point, by analogy with the Weiss theory of ferromagnetism.(62) I made specific reference to this concept in my Inaugural Lecture, and suggested that it could be applied to melting. This concept was authenticated by Onsager's exact solution.

(e) The idea of Bragg and Williams was extended by Landau to a variety of λ-point transitions. Landau emphasized the role of symmetry in such transitions; the symmetry of the low-temperature phase is characterized by an *order parameter* which becomes zero at the transition point.(63)

The year 1937 saw a development in the statistical mechanics of interacting systems which looked extraordinarily promising. J. E. Mayer developed a perturbation theory for a condensing gas(64) the terms of which were depicted diagrammatically; each diagram represented a multi-dimensional *cluster integral*. For the expansion of the partition function the diagrams were connected graphs; but when he derived the virial expansion for the pressure he found that only multiply connected diagrams survived, and he was able to establish this result rigorously to all orders; the survivors were termed *irreducible cluster integrals*. The second virial coefficient involved only one integral, the third one integral, and the fourth 3 integrals; the situation looked extremely promising. If the asymptotic behaviour of the cluster integrals could be determined a rigorous theory of condensation would result.

The impact was great, and at a meeting in Amsterdam the same year to commemorate the centenary of the birth of van der Waals, Uhlenbeck noted that only 2 of the many papers presented were concerned with the equation of van der Waals; almost all of the remainder concerned themselves with the exciting new developments. Max Born also presented a paper entitled "The Statistical Mechanics of Condensing Systems," *Physica* **4**:1034 (1937) and in a second paper with the same title published (with K. Fuchs) in *Proc. Roy. Soc. A* **166**:391 (1938) he began as follows: "J. E. Mayer has published with some collaborators several papers under the same title as the present one. We consider these papers as a most important contribution to statistical mechanics, and this opinion was shared by the International Conference held in Amsterdam."

Alas the expectations failed to be fulfilled; the proposals for behaviour in the critical region had no real basis and were subsequently falsified by experiment.(65)

The subsequent work of Uhlenbeck and his collaborators,(66) which introduced graph theory into statistical mechanics, clarified the background to the Mayer theory, and provided a major tool in the subsequent series expansion work for crystalline models. Uhlenbeck showed that although the number of multiply-connected graphs started with such modest numbers, they eventually grew exponentially, and for the ninth virial coefficient 194066 multi-dimensional integrals must be evaluated.

5. ISING MODEL

I have also written at length in my book(57) on all stages of the development of the Ising model, so I will again summarize briefly its origin and early development, and refer the reader to the book for further details. In 1925 W. Lenz, who had been looking for a simple model which might serve to explain ferromagnetism, proposed the model to his graduate student E. Ising, and suggested that he investigate whether it possessed a non-zero spontaneous magnetization. Ising was able to solve the problem in one-dimension only where the statistical problem is quite simple, and the solution did not possess a spontaneous magnetization at any temperature above zero. He concluded wrongly that the model would not give rise to a spontaneous magnetization in higher dimensions.(67)

Ising was forced to leave Germany before World War 2 because of his Jewish descent. He eventually settled in Peoria, Illinois where he taught at a local college for the rest of his life. To the best of my knowledge he published no other paper on physical research, or ever attended any of the Conferences in which the Ising model was a topic of discussion. He passed away at an advanced age a year or two ago. Thus, although Ising must be included in the founding fathers, his rating is very weak.

In 1936 Peierls was the first to demonstrate convincingly that the two-dimensional Ising model, at sufficiently low but non-zero temperatures, does possess a spontaneous magnetization.(68) (Michael tells me that he spotted a flaw in the original argument, but was pre-empted by R. B. Griffiths, *Phys. Rev. A* **136**:437 (1964)).

Of importance to subsequent developments was the introduction of the lattice gas model(69) by Cernuschi and Eyring in 1939 in which the cells of a lattice can be either occupied by a *single* molecule, or unoccupied (a hole). Any two neighbouring occupied cells have an attractive interaction; the fact that an occupied cell cannot accept another molecule provides a crude representation of a hard-core repulsion, and the possibility of varying the number of holes allows for change in free volume, one of the essential features of a liquid. The lattice gas model is isomorphic with the Ising model, and pioneered the exploration of the critical behaviour of fluids.

The special simplicity of all one-dimensional chains was highlighted by the introduction of the transfer matrix, whose largest eigenvalue determines the partition function of an infinite chain. Three groups were independently responsible for this innovation. Kramers and Wannier,(70) Lassettre and

Howe,(71) and Montroll.(72) We shall devote particular attention to the Kramers and Wannier papers which contain the following additional features of major significance:

(a) They evaluated the transfer matrix for the Ising model in zero field for a number of $n \times \infty$ strips, and discovered a transformation by which the low temperature values can be converted into high temperature values.

(b) Focusing attention on the fixed point which transforms into itself they conjectured that when n becomes infinite this is the Curie temperature T_C.

(c) Calculating the specific heats of finite strips at T_C for small values of n they found values which increased linearly with $\log n$. Extrapolating to $n = \infty$ they conjectured that the specific heat of the two dimensional lattice becomes infinite at T_C.

All of these results were confirmed by Onsager, and it is legitimate to look on these papers of Kramers and Wannier as the precursor to Onsager.

In addition they derived a small number of terms of high and low temperature series expansions for the partition function. These were used to test the accuracy of the various closed form approximations, and they concluded that even the most sophisticated of these approximations deviates at an early stage from the true solution. This emphasized the importance of exact work.

They did not realize that if longer series expansions were derived, the coefficients themselves can provide information about critical behaviour. This was my own contribution.(73)

This brings us to Onsager's varied and massive contributions which I discuss at length in my book.(74) I referred to some of them in my Inaugural Lecture, and they may well have been the decisive factor which brought Michael into the field. Their brilliance was enhanced rather than dimmed by the RG which needed exact solutions against which it could check its conclusions.

I am sure that Michael will join me in ending with a tribute to Elliott Montroll from whose activities and hospitality we both benefited greatly. Elliott was a mine of information about all that was going on in statistical mechanics; when he learned of any new problem he shared it with all his friends, and Onsager has described the role which Elliott played as a catalyst in interesting him in the Ising problem. His publications do not adequately reflect his contribution to the field, and he can unquestionably be listed with the major pioneers. He was involved in the first diagrammatic expansion of correlation functions;(75) and he was particularly skilled in writing comprehensive review articles. Many research workers learned the essentials of a new field from such a review article.

ACKNOWLEDGMENTS

I am grateful to Michael Fisher for reading the original manuscript with his usual care, and making a number of helpful suggestions.

REFERENCES

1. C. Domb, *Physica A* **177**:1 (1991).
2. K. G. Wilson, *Nobel Lectures in Physics 1981–1990* (World Scientific), p. 100.
3. J. D. Watson, *The Double Helix* (Athaneum Press, 1968). Foreword by Sir Lawrence Bragg.
4. R. Clausius, *Ann. Physik* **100**:353 (1857), English translation, *Phil. Mag.* **14**:108 (1859).
5. J. W. Gibbs, *Proc. Am. Acad.* **16**:458 (1889).
6. S. G. Brush, *Statistical Physics and the Atomic Theory of Matter* (Princeton University Press, 1983), p. 50.
7. C. H. D. Buys-Ballot, *Ann. Phys. Leipzig* **103**:24D (1858).
8. R. Clausius, *Ann. Physik* **105**:239 (1858). English translation, *Phil. Mag.* **17**:81 (1859).
9. *Report of the Twenty-ninth Meeting of the British Association for the Advancement of Science* (London, 1860), part 2:9. Reproduced in *The Scientific Lectures and Papers of James Clerk Maxwell*, P. M. Harman, ed. (Cambridge, 1990), pp. 615–616. The complete paper was published in *Phil. Mag.* January and June 1860. Maxwell's Scientific Papers, Vol. 1, p. 376.
10. *Nature* **8**:437 (1873); Maxwell's Scientific Papers, Vol. 2, p. 361.
11. S. G. Brush, *Statistical Physics and the Atomic Theory of Matter* (Princeton, 1983), p. 59. This book is a useful general reference for the material of this paper.
12. Lewis Campbell and William Garnett, *The Life of James Clerk* (Maxwell, London, 1882), p. 143 (Reprinted Johnson Reprint Corporation, 1969).
13. Harvey S. Leff and Andrew F. Rex, eds., *Maxwell's Demon* (Adam Hilger, Bristol, 1990).
14. John Earman and John D. Norton, *Stud. Hist. Phil. Mod. Phys.* **29**:435 (1998); **30**:1 (1999).
15. J. Maddox, *Nature* **345**:109 (1990).
16. N. Brown, *New Scientist* **35** (July 14, 1990).
17. *Nature* **15** (1876). Maxwell's Scientific Papers, Vol. 2, p. 418.
18. A useful general reference is C. Cercignano, Ludwig Boltzmann, Oxford (1998).
19. It is interesting that this was already noted by Gibbs who wrote in Ref. 6 "In reading Clausius we seem to be reading mechanics; in reading Maxwell, and in much of Boltzmann's most valuable work, we seem rather to be reading in the theory of probabilities."
20. L. Boltzmann, *Wien. Ber.* **58**:517 (1868).
21. L. Boltzmann, *Wien. Ber.* **63**:679 (1871).
22. L. Boltzmann, *Wien. Ber.* **76**:373 (1877).
23. L. Boltzmann, *Ann. Physik* **60**:392 (1897). English translation in S. G. Brush, *Kinetic Theory* (Pergamon Press, Oxford, 1966), Vol. 2, p. 238.
24. J. W. Gibbs, *Elementary Principles in Statistical Mechanics* (Scribner's, New York and Arnold, London, 1902).
25. Published as a book, E. Schrödinger, *Statistical Thermodynamics* (Cambridge, 1946).
26. A. Pais, *Subtle is the Lord, The Science and Life of Einstein* (American Philological Association, 1982), pp. 60–75.
27. The most comprehensive treatment in the literature is in L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon, Oxford, 1958), Ch. 12.
28. A. Einstein, *Ann. Phys. Leipzig* **33**:127 (1910).
29. M. S. Smoluchowski, *Ann. Phys. Leipzig* **25**:205 (1908); *Phil. Mag.* **23**:165 (1912).

30. M. Planck, *Verh. D. Phys. Ges.* **2**:202 (1900); **2**:237 (1900).
31. O. Lummer and E. Pringsheim, *Verh. d. D. Phys. Ges.* **1**:23, 215 (1899).
32. Thomas S. Kuhn, *Black-Body Theory and the Quantum Discontinuity 1894–1912* (Oxford, 1978).
33. *Max Planck Theory of Heat* (Macmillan, London, 1932).
34. Mark W. Zemansky, *Heat and Thermodynamics*, first edn. (McGraw–Hill, 1937).
35. A. B. Pippard, *Classical Thermodynamics* (Cambridge, 1957).
36. H. B. Callen, *Thermodynamics* (Wiley, 1960).
37. M. Planck, *Ann. Phys. Lpz.* **1**:719 (1900)
38. See ref. 32, pp. 144–152.
39. A. Einstein, *Ann. d. Phys.* **17**:132 (1905); **20**:199 (1906).
40. A. Einstein, *Ann. d. Phys.* **22**:800 (1907).
41. M. Planck, *Ann. Phys. Lpz.* **75**:673 (1924).
42. S. N. Bose, *Z. Phys.* **26**:178 (1924).
43. Reference 26, p. 246.
44. A. Einstein, *Sitz. Preuss. Akad. Wiss.*, p. 261 (1924); p. 3 (1925).
45. W. Pauli, *Z. Phys.* **31**:765 (1925).
46. E. Fermi, *Z. Phys.* **36**:902 (1926).
47. P. A. M. Dirac, *Proc. Roy. Soc. A* **112**:661 (1926).
48. P. A. M. Dirac, *Quantum Mechanics*, 4th edn. (Oxford, 1958).
49. C. G. Darwin and R. H. Fowler, *Proc. Camb. Phil. Soc.* **21**:391 (1922).
50. R. H. Fowler, *Statistical Mechanics*, first edn. (Cambridge, 1929).
51. R. H. Fowler, *Statistical Mechanics*, 2nd edn. (Cambridge, 1936).
52. R. H. Fowler and E. A. Guggenheim, *Statistical Thermodynamics* (Cambridge, 1939).
53. Reference 50, p. 30.
54. R. H. Fowler, *Proc. Camb. Phil. Soc.* **34**:382 (1938).
55. Reference 52, Chapter 6.
56. G. S. Rushbrooke, *Introduction to Statistical Mechanics* (Oxford, 1949).
57. C. Domb, *The Critical Point* (Taylor and Francis, London, 1996).
58. Reference 57, Chapter 2.
59. Reference 57, Chapter 3.
60. Reference 57, Chapter 4, pp. 101–112.
61. M. E. Fisher, *J. Math. Phys.* **5**:944 (1964).
62. Reference 57, Chapter 4, pp. 110–112.
63. Reference 57, Chapter 4, pp. 122–124.
64. The same development was suggested independently by J. Yvon, but his work was less well known.
65. Reference 57, Chapter 5, pp. 195–198.
66. See C. Domb, Chapter 1 in C. Domb and M. S. Green, *Phase Transitions and Critical Phenomena*, Vol. 3 (Academic Press, 1974) for a detailed discussion with references.
67. E. Ising, *Z. Phys.* **31**:253 (1925).
68. R. E. Peierls, *Proc. Camb. Phil. Soc.* **32**:477 (1936).
69. F. Cernuschi and H. Eyring, *J. Chem. Phys.* **7**:547 (1939).
70. H. A. Kramers and G. H. Wannier, *Phys. Rev.* **60**:252, 263 (1941).
71. E. N. Lassettre and J. P. Howe, *J. Chem. Phys.* **9**:747 (1941).
72. E. W. Montroll, *J. Chem. Phys.* **9**:706 (1941).
73. C. Domb, Doctoral Thesis (Cambridge, 1949). *Proc. Roy. Soc. A* **196**:36, **199**:39 (1949)
74. Reference 57, Chapter 5, p. 129–147.
75. J. E. Mayer and E. W. Montroll, *J. Chem. Phys.* **9**:2 (1941).