The Annals of Probability
1986, Vol. 14, No. 4, 1129–1138

THE CONTRIBUTIONS OF MARK KAC TO MATHEMATICAL PHYSICS

BY COLIN J. THOMPSON

University of Melbourne

By his own admission Mark Kac was a mathematician with a strong interest in physics. He bemoaned the fact that mathematics and physics had entered a period of alienation in the early part of this century and he often expressed the view and hope that the "two great disciplines," as he called them, would come together again. His writings and lectures in mathematical physics have contributed much to the fulfillment of this hope.

Mark's papers in mathematical physics are noted for their unique lucid and informal style, with an emphasis on intuitive and heuristic arguments, rather than on tedious mathematical detail. He was fond of saying that a demonstration is often more valuable than a mathematical proof since, as he put it, a "demonstration is to convince a reasonable man whereas a proof is designed to convince a stubborn one." This belief is reflected in his many articles on topics in mathematical physics and was always apparent in his lectures.

Mark's main contributions to mathematical physics were in the field of statistical mechanics, both classical and quantum, equilibrium and nonequilibrium, but were primarily concerned with model systems in classical equilibrium statistical mechanics which exhibit phase transitions.

The physical phenomenon of phase transitions has been known for a long time and has been extensively studied, both theoretically and experimentally, over the years. Perhaps the simplest example of a phase transition is the condensation of a gas to a liquid under compression at constant pressure for sufficiently low temperatures T. As the temperature is raised one reaches a critical point T_c beyond which there is no abrupt condensation no matter how much it is compressed. In the neighbourhood of such a critical point interesting things happen, such as the phenomenon of critical opalescence caused by large fluctuations in the density. Also, certain physical quantities such as the specific heat and compressibility, become singular as T approaches T_c. Similar phenomena occur in magnetic systems where T_c is now the temperature beyond which the residual or spontaneous magnetization vanishes. For such systems the zero-field specific heat and isothermal susceptibility typically diverge as T approaches T_c.

The theoretical basis for the study of phase transitions lies in the formulation of equilibrium statistical mechanics due to Gibbs [9]. In this formulation the occurrence of a phase transition becomes a precise mathematical problem. Thus, if we have a system described by a Hamiltonian $\mathscr{H}\{\mu\}$, the Gibbs canonical distribution $p\{\mu\}$ for the system in configuration or state μ is given by

$$p\{\mu\} = Z^{-1}\exp(-\mathscr{H}\{\mu\}/kT), \tag{1}$$

Received February 1986; revised March 1986.

where k is Boltzmann's constant, T is the absolute temperature, and Z, the canonical partition function, is given by

$$Z = \int_{\Gamma} \exp(-\mathscr{H}\{\mu\}/kT)\, d\mu, \tag{2}$$

where the integration is taken over the entire configuration space Γ. Quantities of physical interest are then given as averages with respect to the distribution p. For example, the average energy is given by

$$E = \langle \mathscr{H} \rangle = \int_{\Gamma} \mathscr{H}\{\mu\} p\{\mu\}\, d\mu,$$

the specific heat by

$$C = \frac{\partial E}{\partial T},$$

and so forth, and a phase transition will occur if a derived thermodynamic quantity such as C is singular at some critical temperature T_c.

A problem immediately arises with this definition of a phase transition if the system is finite, as of course it is in reality. This is simply because for finite systems, $p\{\mu\}$, Z and averages with respect to $p\{\mu\}$, are analytic functions of T (for $T \neq 0$). In order to give a precise mathematical definition of a phase transition we then need to consider infinite systems which is a reasonable thing to do since in practice there are approximately 10^{23} particles per cubic centimetre.

A standard way of proceeding is to consider a finite system of N particles in a box with volume V and to take the so-called thermodynamic limit; $N \to \infty$ and $V \to \infty$ with the specific volume $v = V/N$ fixed. In this limit, when it exists (for some suitably chosen constant c_N), the free energy per particle is defined by

$$\psi(v, T) = -\lim kTN^{-1}\log(c_N Z).$$

Comparison with the above formulas shows after a moment's reflection that in the thermodynamic limit the specific heat (per particle) is given by

$$c = -T(\partial^2\psi/\partial T^2).$$

The compressibility mentioned previously is similarly given in terms of the free energy ψ by

$$K = v(\partial^2\psi/\partial v^2).$$

There is now, of course, no mathematical reason why c and K need not be singular for some real, positive v and/or T. That is, for the system to have a phase transition. We then *define* a *phase transition* point to be any point of nonanalyticity of the free energy $\psi(v, T)$ occurring for real positive v and/or T.

The above is now universally accepted as the proper mathematical definition of a phase transition. Pre-1940's, however, doubts were commonly expressed as to whether the Gibbs canonical distribution, or ensemble, alone could explain the occurrence of a phase transition. A commonly expressed view was: "How are the molecules to know when they should condense?" Such doubts were finally laid to rest when Onsager [6] "solved" the two-dimensional Ising model in the early 1940's. From an exact evaluation of the partition function for this model, Onsager was able to show, in particular, that the specific heat has a logarithmic divergence at a finite critical temperature. Onsager's 1944 paper on the Ising model remains a real tour de force in mathematical physics and was probably one of the main contributing factors to Mark's subsequent involvement in the subject.

As Mark relates in his "personal reminiscence" article [K95] (Reference citations preceded by K refer to references listed in Publications of Mark Kac, which appears in this issue, pages 1149–1154.) on the work of his friend and colleague Ted Berlin, he was first exposed to the Ising model in the spring of 1947 by his long-time friend and colleague George Uhlenbeck. The exposure may have been along the following lines:

Consider a set of N spins $\mu_i = \pm 1$, $i = 1, 2, \ldots, N$, located on say the vertices of a regular lattice and allow the spins to interact pairwise with coupling constants J_{ij} and individually with an external magnetic field H so that in a given configuration of spins $\{\mu\} = \{\mu_1, \mu_2, \ldots, \mu_N\}$ the Hamiltonian or interaction energy is given by

$$\mathscr{H}\{\mu\} = -\sum_{1 \le i < j \le N} J_{ij}\mu_i\mu_j - H\sum_{i=1}^{N} \mu_i.$$

Now in the Gibbs prescription the canonical partition function is defined by

$$Z = \sum_{\{\mu\}} \exp(-\beta\mathscr{H}\{\mu\}/kT),$$

and in the thermodynamic limit the free energy per spin is given by

$$\psi(H, T) = -\lim_{N \to \infty} kTN^{-1}\log Z.$$

Onsager's great achievement was to obtain an explicit expression for $\psi(0, T)$ for the special case where the spins occupy the vertices of a regular two-dimensional square lattice and only nearest-neighbor spins are allowed to interact. His expression is

$$-\psi(0, T)/kT$$
$$= \log 2 + \frac{1}{2\pi^2}\int\int_0^{\pi} \log\left[\cosh^2 2K - \sinh 2K(\cos\theta_1 + \cos\theta_2)\right] d\theta_1\, d\theta_2,$$

where $K = J/kT$ and J is the nearest-neighbour coupling constant. It will be noted that this expression is singular at $K = K_c = J/kT_c$ given by

$$\sinh 2K_c = 1,$$

and, after some calculation, that the specific heat has a symmetric logarithmic divergence as T approaches T_c.

In terms of numbers of papers published on the problem, the Ising model ranks as probably the most celebrated model in mathematical physics. In spite of the amount of effort devoted to the problem however, exact solutions have not been found to even simple-looking extensions of Onsager's case, such as with nonzero field, with next-nearest-neighbor interactions as well as nearest-neighbor interactions, and of course the three-dimensional model. Many other interesting two-dimensional lattice problems have been solved exactly by Lieb, Baxter, and others, however [1].

As mentioned above, Mark's involvement with the Ising problem began in 1947. In his own words [K95]:

> "It soon became obvious that it was not a problem one solves on the spur of the moment, and, in the best mathematical tradition, not being able to solve the original problem, I looked around for a similar problem which I could solve. I then proceeded to replace the discrete spins by continuous ones distributed according to the Gaussian distribution. In no time I had the free energy per spin calculated, and to my amazement and pleasure the answer looked remarkably like Onsager's."

This was the birth of the so-called Gaussian model. Apart from trivial numerical factors, the free energy has the same form as Onsager's expression but with the argument of the logarithm replaced by $[1 - 2K(\cos\theta_1 + \cos\theta_2)]$ which is very intriguing and to this day still somewhat of a mystery.

The Gaussian model, as noted by Mark, suffers from what he referred to as a "low temperature catastrophe." That is, the free energy is not defined when $K > \frac{1}{4}$.

In searching for a potentially soluble model which was not catastrophic and was still Ising-like, Mark quickly hit upon the idea of replacing the statistical weight of the Ising spins, which can be thought of as concentrated at the vertices $(\pm 1, \pm 1, \ldots, \pm 1)$ of the hypercube inscribed in the sphere

$$\mu_1^2 + \mu_2^2 + \cdots + \mu_N^2 = N,$$

by the uniform distribution on the sphere. That is, the partition function is obtained by integrating the Boltzmann factor $\exp(-\beta\mathscr{H}\{\mu\}/kT)$ over the sphere rather than summing over discrete Ising configurations.

This is the so-called spherical model which also has an extensive literature and ranks alongside the Ising example as one of the most studied models in statistical mechanics.

The usual method of "solving" the spherical model, which was originally devised by Ted Berlin in 1947, utilizes the method of steepest descents (in the thermodynamic limit). The results were "announced" in 1949 and the joint paper "The spherical model of a ferromagnet" [K58] appeared in print in 1952. This was essentially Mark's first published contribution in mathematical physics.

The nice thing about the spherical model is that it can be solved exactly in any dimension and, in principle, for any reasonable interaction potential. For nearest-neighbor interactions the model has no phase transition in one and two dimensions but in three dimensions the model does have a phase transition, from a disordered paramagnetic state to a ferromagnetic state with long-range order.

Mark was particularly intrigued by the fact that the occurrence of a phase transition in the spherical model is due to the "sticking" of the saddle point (in the steepest descents calculation) for low temperatures ($T < T_c$). The more general question of mathematical mechanisms underlying phase transitions was to occupy Mark's thoughts off and on for the remainder of his life. The mechanism for the spherical model turns out to be rather special and in fact the low temperature behaviour of the model is rather unphysical.

Mark published three further papers on the spherical model. The first of these [K118], showed how to evaluate the partition function without using steepest descents, and also provided the first proof of an observation by Gene Stanley [8] that the spherical model could be obtained as a limit of a sequence of so-called n-vector models in which the Ising spins are replaced by n-dimensional unit vectors. Paper [K143] considered certain modifications of the spherical model which are more "Ising-like," and his final paper on the subject [K149], was concerned with computation of correlation functions and their relationship with corresponding quantities for the so-called mean spherical model in which the spherical constraint is satisfied "on average."

The year 1952 saw Mark's second paper in statistical mechanics: the joint paper [K59] with John Ward on "A combinatorial solution of the two-dimensional Ising model." This paper was based on an earlier observation by van der Waerden in 1941 [10] which reduced the problem of evaluating the partition function of the two-dimensional Ising model to the problem of counting closed polygons on the underlying lattice. In this formulation the partition function becomes, essentially, the generating function for the associated combinatorial problem. The idea Mark and John Ward had during their coincidental visit to the Institute for Advanced Study in Princeton during the 1951–52 year was to "do the counting" with the aid of a matrix whose determinant yielded the generating function. There is no problem in associating cycles in the expansion of a determinant with closed oriented polygons but one needs to count unoriented polygons. Mark and John cleverly overcame this difficulty by essentially considering the lattice together with its mirror image. There still remained, however, the problem of "cancelling" the unwanted minus signs in the expansion of the determinant.

The actual process used to construct the appropriate matrix was in fact a shrewd piece of experimental mathematics. I recall Mark telling me that they tried so many matrices that they lost count of the number and merely noted the weight of the paper wasted. As Mark recounted the story he finally got the correct matrix on New Year's Eve 1951. John Ward was apparently out playing tennis and the only person Mark could find to break the good news to was Bram Pais who was all dressed up in a dinner suit and obviously on his way to a party. The paper concluded, in typical Kac fashion, with an acknowledgement to "many of our friends for the healthy pessimism they showed during the early stages of this work."

The "combinatorial approach" is now a standard method of dealing with problems in lattice statistics [9], but these days the Pfaffian of the matrix (which is roughly speaking "half of a determinant"), rather than the determinant, is used in the counting. Mark always found it amusing to recall that in "doubling the lattice," he and John Ward had inadvertently discovered Pfaffians.

In the alternative and more common algebraic approach to the Ising model [9] the partition function is expressed as the trace of a matrix, specifically,

$$Z = \mathrm{Tr}(\mathbf{T}^n)$$

where $\mathbf{T}$ is the so-called transfer matrix. In one dimension n is the number of spins in the chain and $\mathbf{T}$ is a 2×2 matrix. In two dimensions $\mathbf{T}$ is $2^m \times 2^m$ where m is the number of rows and n in this case is the number of columns in the lattice.

To obtain this form for Z one writes the Boltzmann factor as a product over contributions from columns and interactions with their nearest-neighbour columns. Summing over column configurations then becomes equivalent to matrix multiplication and if the last column is connected to the first, one obtains the above expression where $\mathbf{T}$ is the column-to-column transfer matrix. The size of the matrix is determined by the number of spin configurations in each column which is 2^m when there are m rows.

In the thermodynamic limit one has

$$\begin{aligned} -\psi/kT &= \lim_{m,\, n \to \infty} (mn)^{-1} \log Z \\ &= \lim_{m \to \infty} m^{-1} \log \lambda_1, \end{aligned}$$

where λ_1 is the maximum eigenvalue of $\mathbf{T}$. Since the entries of $\mathbf{T}$ are positive and of the form $\exp(\alpha/kT)$ one has, from the Perron–Frobenius theorem, that λ_1 is simple and an analytic function of T (nonzero). In other words, for finite m or equivalently, for a finite-by-infinite strip, there is no phase transition.

In order to see how the nonanalyticity or phase transition develops in the limit $m \to \infty$ it is usual to consider the pair correlation function $\langle \mu_i \mu_j \rangle$ for two spins located at positions i and j on the lattice. If the spins lie in the same row and are separated by a distance $|i - j| = r$, it is not difficult to show that [9]

$$\langle \mu_i \mu_j \rangle \sim a(\lambda_2/\lambda_1)^r \quad \text{as } r \to \infty,$$

where a is some constant and λ_2 is the second largest eigenvalue of $\mathbf{T}$. For finite m, $\lambda_2/\lambda_1 < 1$ and the correlations decay exponentially for all temperatures. In the limit $m \to \infty$ however, long-range order sets in for $T < T_c$ due to the *asymptotic degeneracy* of λ_1, that is,

$$\lambda_2/\lambda_1 \sim 1 - O(e^{-bm}) \quad \text{as } m \to \infty \text{ for } T < T_c.$$

This result was obtained by Onsager [6] and forms a *mathematical mechanism* for the phase transition of the two-dimensional Ising model.

Mark was always intrigued by this result and many of his publications centred on the "mathematical mechanism" theme [K102], [K103], [K104], [K114] and on the role of model systems in understanding phase transitions [K117], [K115], [K127], [K153].

In 1959 Mark invented his own model [K80] consisting of a one-dimensional gas of hard rods attracting one another with an exponentially decaying potential of the form

$$v(x) = \gamma e^{-\gamma|x|}$$

between two rods separated by a distance x. Noting that $v(x)$ is the covariance of an Ornstein–Uhlenbeck process, Mark was able to express the partition function for his model in terms of the spectrum of an integral operator, in much the same way as one expresses the partition function for the Ising model in terms of the spectrum of the transfer matrix. Again, the free energy in the thermodynamic limit is simply expressed in terms of the maximum eigenvalue $\lambda_{\max}$ of the integral operator.

Although an explicit expression for $\lambda_{\max}$ could not be found, Mark noted that in the limit $\gamma \to 0+$, corresponding to a weak long-ranged potential, one recovers the classical van der Waals theory of gas–liquid condensation. This limit was examined in great depth and detail in three classic papers [K91], [K92], [K93] published jointly in 1963–64 with George Uhlenbeck and Per Hemmer where, among other things, the role of asymptotic degeneracy of the spectrum is analysed as a mechanism for the phase transitions. Similar models consisting of Ising spins interacting with combinations of nearest-neighbour and weak long-ranged exponentially decaying interactions were formulated by Mark [K88] and, with various collaborators, studied in the limit $\gamma \to 0+$ [K90], [K110] where one recovers the classical Curie–Weiss theory of magnetism. Correction terms to the classical theories were studied for small γ, particularly in the neighbourhood of the phase transition point; the focus of these studies again aimed at elucidating the underlying mathematical mechanism of the phase transition.

Mark's models instigated renewed interest in the classical theories and their range of validity. Lebowitz and Penrose [4], for example, showed that in general, potentials of the form

$$K(\mathbf{x}) = \gamma^d f(\gamma \mathbf{x})$$

for d-dimensional systems, yield the classical van der Waals theory in the limit $\gamma \to 0+$. Such potentials are now known as *Kac-potentials* in recognition of Mark's original contribution to this aspect of the study of phase transitions.

Mark's other main interest in statistical mechanics was in nonequilibrium phenomena and, more particularly, with the "problem of Boltzmann" of how to reconcile reversibility on the microscopic level with apparent irreversibility on the macroscopic level.

Boltzmann's equation in its simplest form is a nonlinear evolution equation for the probability density $f(\mathbf{v}, t)$ of finding a molecule with velocity $\mathbf{v}$ at time t [9]. A striking consequence of Boltzmann's equation is that the function

$$H(t) = \int_{R^3} f(\mathbf{v}, t) \log f(\mathbf{v}, t)\, d^3v$$

is a nonincreasing function of t. As was pointed out at the time, this result, called "the H-theorem," presents a paradox since on the one hand, the microscopic laws of motion on which Boltzmann's equation is based are invariant under time reversal, whereas on the other hand, Boltzmann's H-theorem singles out a preferred direction in time.

Boltzmann realized that there was not really a paradox since the derivation of his equation was based on a *statistical assumption*, the Stosszahlansatz or assumption of *molecular chaos*, as well as the mechanical assumption that only binary collisions occur.

Stated simply, the assumption of molecular chaos asserts that the probability density for simultaneously finding a molecule with velocity $\mathbf{v}_1$ and another with velocity $\mathbf{v}_2$ at time t is the product $f(\mathbf{v}_1, t)f(\mathbf{v}_2, t)$ of the corresponding single particle densities.

Unfortunately, Boltzmann was unable to convince his critics that probabilistic concepts were an integral part of his arguments and that objections to his equation and H-theorem should not be based solely on mechanical considerations. This apparent failure on Boltzmann's part undoubtedly contributed to his untimely death by his own hand.

The importance of probabilistic arguments in understanding the approach to equilibrium was stressed again by P. and T. Ehrenfest in their celebrated 1912 article [2], and perhaps again at the instigation of George Uhlenbeck, Mark took up the challenge in the mid-1950's.

In typical fashion Mark introduced a model in his 1956 paper "Foundations of kinetic theory" [K64], which in fact was a kind of spherical model, and had the advantage that it could be analysed exactly. What he did essentially, was to show that the model had the "propagation of chaos" property as he called it. That is, he was able to show that if the assumption of molecular chaos held at $t = 0$ then it also held at later times and, moreover, the single particle distribution function for his model satisfied a Boltzmann-like equation.

In another paper published in 1956 [K68], Mark considered a simple one-dimensional version of Ehrenfest's "wind-tree model" consisting of black and white balls moving through randomly distributed "impurities" on a ring such that whenever a ball moved through an impurity it changed colour. The model is invariant under time reversal but Mark was able to show that by performing a suitable statistical average over the impurities, the model irreversibly approached the equilibrium state of equal numbers of black and white balls for almost all initial distributions of black and white colours.

Mark published several more papers on the "Boltzmann problem" theme [K72], [K77], [K124], [K125], [K147] and in fact his last published paper [K170], written jointly with Eugene Gutkin, was on "Propagation of chaos and the Burgers equation."

Apart from the interesting joint work [K97] with Ford and Mazur on "Statistical mechanics of assemblies of coupled oscillators," Mark's contributions to quantum statistical mechanics ([K120], [K126], [K129], [K135], and [K152]) were mainly concerned with the noninteracting or ideal Bose gas. Mark was never happy with the foundations of quantum statistical mechanics but was always intrigued by the mathematical similarities between the ideal Bose gas and his spherical model of a ferromagnet. This was particularly apparent in his joint paper with Bob Ziff and George Uhlenbeck [K152] in which careful consideration was given to fluctuations and surface effects in the ideal Bose gas. Mark also made a number of significant contributions to areas related to quantum mechanics ([K146], [K154], [K155], [K162], and the so-called inverse scattering method ([K123], [K132], [K145]). His joint work with Pierre van Moerbeke ([K131], [K137], [K138], [K140]) on nonlinear evolution equations related to the Toda lattice is also particularly noteworthy.

A number of Mark's publications were on the border between mathematics and physics. Perhaps the most noteworthy of these is [K47] which contains the first of many rigorous derivations of the celebrated "Feynman–Kac formula" in which the solution to the (imaginary time) Schrödinger equation is expressed as a Wiener integral, or path integral in Feynman's formulation of quantum mechanics, which was the motivating force behind Mark's derivation of "the formula."

It is impossible to do complete justice to the enormous influence Mark's work had on the development of statistical mechanics and mathematical physics in general. I have made no attempt to compile a comprehensive list of the many published articles that were based on or influenced by Mark's original work. I have instead listed a few key references which may help the reader delve further into these subjects.

In addition to his published work, Mark made substantial contributions through his lectures and informal discussions with colleagues and scientific associates. A notable example of this is described by Mark in a comment section published in the collected works of George Pólya [7]. There Mark describes discussions he had with T. D. Lee and C. N. Yang, again in Princeton during the 1951–52 year, concerning what was later to become the celebrated Lee–Yang circle theorem [5]. Mark managed to construct a proof for a special case, after which it was only a matter of weeks before Lee and Yang had the proof in the general case.

I saw Mark a few months before he died and his mind was as active as ever although his body was obviously failing him. We talked about many things including some new ideas he had about a possible relationship between the equation for the spherical model saddle point and correction terms in γ-expansions for models with "Kac potentials," and also about some unexpected results obtained by his last Ph.D. student, T. M. Katz, on spherical models with certain Kac-type potentials [3].

Mark Kac will be missed but his legacy will remain strong. His style and contributions to the scientific literature will serve as a model for generations of future aspiring mathematical physicists.

REFERENCES

[1] BAXTER, R. J. (1982). *Exactly Solved Models in Statistical Mechanics*. Academic, London.

[2] EHRENFEST, P. and EHRENFEST, T. (1911). Begriffliche Grundlagen der statisticschen Aerffassung in der Mechanik. *Encyklopädie der Mathematischen Wissenschaften*. Teubner, Bd. 4, Teil 32, Leipzig.

[3] KATZ, T. M. (1985). Spherical models with a Gates–Penrose-type phase transition. *J. Statist. Phys.* **38** 589–602.

[4] LEBOWITZ, J. L. and PENROSE, O. (1966). Rigorous treatment of the van der Waals–Maxwell theory of the liquid–vapour transition. *J. Math. Phys.* **7** 98–113.

[5] LEE, T. D. and YANG, C. N. (1952). Statistical theory of equations of state and phase transitions, II. Lattice gas and Ising model. *Phys. Rev.* **87** 410–419.
[6] ONSAGER, L. (1944). Crystal statistics, I. A two-dimensional model with an order–disorder transition. *Phys. Rev.* **65** 117–149.
[7] PÓLYA, G. (1974). *Collected Works* **2** 424–426. MIT Press, Cambridge, Mass.
[8] STANLEY, H. E. (1968). Spherical model as the limit of infinite spin dimensionality. *Phys. Rev.* **176** 718–721.
[9] THOMPSON, C. J. (1979). *Mathematical Statistical Mechanics*. Princeton Univ. Press.
[10] VAN DER WAERDEN, B. L. (1941). Die lange Reichweite der regelmässigen Atomanordnung in Mischkristallen. *Z. Phys.* **118** 473–488.

MATHEMATICS DEPARTMENT
UNIVERSITY OF MELBOURNE
PARKVILLE, VICTORIA 3052
AUSTRALIA