RANDOM MATRICES and the Statistical Theory of Energy Levels

M. L. MEHTA

DEPARTMENT OF PHYSICS AND ASTROPHYSICS
UNIVERSITY OF DELHI
DELHI, INDIA

1967

ACADEMIC PRESS New York and London

COPYRIGHT © 1967, BY ACADEMIC PRESS INC.

ALL RIGHTS RESERVED.
NO PART OF THIS BOOK MAY BE REPRODUCED IN ANY FORM, BY PHOTOSTAT, MICROFILM, OR ANY OTHER MEANS, WITHOUT WRITTEN PERMISSION FROM THE PUBLISHERS.

ACADEMIC PRESS INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS INC. (LONDON) LTD.
Berkeley Square House, London W.1

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 67-23169

PRINTED IN THE UNITED STATES OF AMERICA

Preface

Though random matrices were first encountered in mathematical statistics by Hsu, Wishart, and others, intensive study of their properties in connection with nuclear physics began with the work of Wigner in the 1950's. Much material has accumulated since then, and it was felt that it should be collected. A reprint volume to satisfy this need has been edited by C. E. Porter with a critical introduction (see References); nevertheless, the feeling was that a book containing a coherent treatment of the subject would be welcome.

We make the assumption that the local statistical behavior of the energy levels of a sufficiently complicated system is simulated by that of the eigenvalues of a random matrix. Chapter 1 is a rapid survey of our understanding of nuclear spectra from this point of view. The discussion is rather general, in sharp contrast to the precise problems treated in the rest of the book. In Chapter 2 an analysis of the usual symmetries that a quantum system might possess is carried out, and the joint probability density function for the various matrix elements of the Hamiltonian is derived as a consequence. The transition from matrix elements to eigenvalues is made in Chapter 3 and the standard arguments of classical statistical mechanics are applied in Chapter 4 to derive the eigenvalue density. An unproved conjecture is also stated. In Chapter 5 the method of integration over alternate variables is presented, and an application of the Fredholm theory of integral equations is made to the problem of eigenvalue spacings. The methods developed in Chapter 5 are basic to an understanding of most of the remaining chapters. Chapter 6 deals with the correlations and spacings for less useful cases. A Brownian motion model is described in Chapter 7. Chapters 8 to 11 treat circular ensembles; Chapters 8 to 10 repeat calculations analogous to those of Chapters 4 to 7. The integration method discussed in Chapter 11 originated with Wigner and is being published here for the first time. The theory of non-Hermitian random matrices,

though not applicable to any physical problems, is a fascinating subject and must be studied for its own sake. In this direction an impressive effort by Ginibre is described in Chapter 12. For the Gaussian ensembles the level density in regions where it is very low is discussed in Chapter 13. The investigations of Chapter 16 and Appendices A.29 and A.30 were recently carried out in collaboration with Professor Wigner at Princeton University. Chapters 14, 15, and 17 treat a number of other topics. Most of the material in the appendices is either well known or was published elsewhere and is collected here for ready reference. It was surprisingly difficult to obtain the proof contained in A.21, while A.29, A.30, and A.31 are new.

It is my pleasant duty to thank Professor C. Bloch, Professor F. J. Dyson, and Professor E. P. Wigner to whom I owe so much by way of education and inspiration. I have made use of the cited literature and in particular published as well as unpublished works of E. P. Wigner, F. J. Dyson, and M. Gaudin. I am thankful to the editors of the various publications for allowing me to do so. This book was written in sections at Tata Institute of Fundamental Research, Bombay, the Indian Institute of Technology, Kanpur, Delhi University, Argonne National Laboratory, and Princeton University. The lectures given at the State University of New York at Stony Brook, Long Island, were helpful in the initial stages. I am grateful to all these institutions for their hospitality. My thanks are due to my colleagues H. S. Mani, N. Rosenzweig, and P. K. Srivastava for their critical comments. A few additions and changes were made at almost every stage of the process of publication and I am thankful to the staff of Academic Press for their cooperation.

October, 1967 M. L. MEHTA
Saclay, France

1 / Introduction

1.1. The Need to Study Random Matrices

The experimental nuclear physicists have been and still are collecting vast amounts of data concerning the excitation spectra of various nuclei [Garg *et al.*, 1; Rosen *et al.*, 1]. The ground state and low-lying excited states have been impressively explained in terms of an independent particle model where the nucleons are supposed to move freely in an average potential well [Mayer and Jensen, 1; Kisslinger and Sorenson, 1]. As the excitation energy increases, more and more nucleons are thrown out of the main body of the nucleus, and the approximation of replacing their complicated interactions with an average potential becomes more and more inaccurate. At still higher excitations the nuclear states are so dense and the intermixing is so strong that it is a hopeless task to try to explain the individual states; but when the complications increase beyond a certain point the situation becomes hopeful again, for we are no longer required to explain the characteristics of every individual state but only their average properties, which is much simpler.

The average behavior of the various energy levels is of prime importance in the study of nuclear reactions. In fact, nuclear reactions may be put into two major classes—fast and slow. In the first case a typical reaction time is of the order of the time taken by the incident nucleon to pass through the nucleus. The wavelength of the incident nucleon is much smaller than the nuclear dimensions, and the time it spends inside the nucleus is so short that it interacts with only a few nucleons inside the nucleus. A typical example is the head-on collision with one nucleon in which the incident nucleon hits and ejects a nucleon, thus giving it almost all its momentum and energy. Consequently, in such cases the coherence and interference effects between incoming and outgoing nucleons are strong.

Another extreme is provided by the slow reactions in which the

typical reaction times are two to three orders of magnitude larger. The incident nucleon is trapped and all its energy and momentum are quickly distributed among the various constituents of the target nucleus. It takes a long time before enough energy is again concentrated on a single nucleon to eject it. The compound nucleus lives long enough to forget the manner of its formation, and the subsequent decay is therefore independent of the way in which it was formed.

In the slow reactions, unless the energy of the incident neutron is very sharply defined, a large number of neighboring energy levels of the compound nucleus are involved, hence the importance of an investigation of their average properties, such as the distribution of neutron and radiation widths, level spacings, and fission widths. It is natural that such phenomena, which result from complicated many-body interactions, will give rise to statistical theories. We shall concentrate mainly on the average properties of nuclear levels such as level spacings.

According to quantum mechanics, the energy levels of a system are supposed to be described by the eigenvalues of a Hermitian operator, called the Hamiltonian. The energy-level scheme of a system consists in general of a continuum and a certain, perhaps a large, number of discrete levels. The Hamiltonian of the system should have the same eigenvalue structure and therefore must operate in an infinite dimensional Hilbert space. To avoid the difficulty of working with an infinite dimensional Hilbert space, we make approximations amounting to a truncation keeping only the part of the Hilbert space that is relevant to the problem at hand and either forgetting about the rest or taking its effect in an approximate manner on the part considered. Because we are interested in the discrete part of the energy-level schemes of various quantum systems, we approximate the true Hilbert space by one having a finite, though large, number of dimensions. Choosing a basis in this space, we represent our Hamiltonians by finite dimensional matrices. If we can solve the eigenvalue equation,

$$H\Psi_i = E_i\Psi_i ,$$

we shall get all the eigenvalues and eigenfunctions of the system, and any physical information can then be deduced, in principle, from this knowledge. In the case of the nucleus, however, there are two difficulties. First, we do not know the Hamiltonian and, second, even if we did, it would be far too complicated to attempt to solve it.

Therefore from the very beginning we shall be making statistical hypotheses on H, compatible with the general symmetry properties. Choosing a complete set of functions as basis, we represent the Hamiltonian operators H as matrices. The elements of these matrices are random variables whose distributions are restricted only by the general symmetry properties we might impose on the ensemble of operators. Statistical theory does not predict the detailed level sequence of any one nucleus, but it does describe the general appearance and the degree of irregularity of the level structure that is expected to occur in any nucleus, which is too complicated to be understood in detail.

In classical statistical mechanics a system may be in any one of the many possible states, but one does not ask in which particular state a given system is. Here we shall renounce knowledge of the system itself. As in orthodox statistical mechanics we shall consider an ensemble of Hamiltonians, each of which could describe a different nucleus. There is a reasonable expectation, though no rigorous mathematical proof, that a system under observation will be described correctly by an ensemble average. This expectation is strong because the system might be one of a huge variety of systems, and very few of them will deviate much from a properly chosen ensemble average. On the other hand, our assumption that the ensemble average correctly describes a particular system, say the U^{239} nucleus, is not compelling. In fact, if this particular nucleus turns out to be far removed from the ensemble average, it will show that the U^{239} Hamiltonian possesses specific properties of which we are not aware. This, then, will prompt us to try to discover the nature and origin of these properties [Dyson, 1].

Wigner was the first to propose in this connection the hypothesis alluded to, namely that the local statistical behavior of levels in a simple sequence is identical with the eigenvalues of a random matrix. A simple sequence is one whose levels all have the same spin, parity, and other strictly conserved quantities, if any, which result from the symmetry of the system. The corresponding symmetry requirements are to be imposed on the random matrix. Porter and Rosenzweig were the early workers in the field who analyzed the nuclear experimental data made available by Hughes, Harvey, Rosen, and co-workers and the atomic data compiled by C. E. Moore [1]. They found that the occurrence of two levels close to each other in a simple sequence is a rare event. They also used the computer to generate and diagonalize a large number of random matrices. This Monte Carlo

analysis indicated the correctness of Wigner's hypothesis. In fact, it indicated more; the density and the spacing distribution of eigenvalues of real symmetric matrices are independent of many details of the distribution of individual matrix elements. All that is required is the same distribution for all diagonal elements and that the off-diagonal elements be distributed symmetrically about the zero mean and have the same mean square deviation. This independence is to be expected as well in the case of complex Hermitian or self-dual quaternion matrices, but apart from this numerical evidence and a few heuristic arguments of Wigner no rigorous derivation of this fact has yet been found. The case of the Gaussian distributions of matrix elements is still the only one treated analytically by Hsu, Mehta, Gaudin, Dyson, Bronk, Ginibre, and others, and we have described these developments in great detail in the following pages. From a group-theoretical analysis Dyson [5] found that an irreducible ensemble of matrices, invariant under a symmetry group G, necessarily belongs to one of three classes, named by him orthogonal, unitary, and symplectic. We shall not go into these elegant group-theoretical arguments but shall devote enough space to the study of the circular ensembles introduced by Dyson. It is remarkable that standard thermodynamics can be applied to obtain certain results which otherwise would be difficult to derive. A theory of the Brownian motion of matrix elements has also been created by Dyson thus rederiving a few known results. However, it remains largely a curiosity.

The physical properties of metals depend characteristically on their excitation spectra. In bulk metal at high temperatures the electronic energy levels lie very near to one another and are broad enough to overlap and form a continuous spectrum. As the sample gets smaller, this spectrum becomes discrete, and as the temperature decreases the widths of the individual levels decrease. If the metallic particles are minute enough and at low enough temperatures, the spacings of the electronic energy levels may eventually become much larger than the other energies, such as the level widths and the thermal energy kT. Under such conditions the thermal and the electromagnetic properties of the fine metallic particles may deviate considerably from those of the bulk metal. This circumstance has already been noted by Fröhlich [1] and proposed by him as a test of quantum mechanics. Because it is difficult to control the shapes of such small particles while they are being experimentally produced, the electronic energy levels are seen to be random and the theory for

the eigenvalues of random matrices may be useful in their study.

Random matrices are also encountered in other branches of physics. For example, glass may be considered as a collection of random nets, that is, a collection of particles with random masses exerting random mutual forces, and it is of interest to determine the distribution of frequencies of such nets [Dyson, 6]. A one-dimensional model of glass is considered in Chapter 14.

1.2. A Summary of Statistical Facts about Nuclear Energy Levels

1.2.1. Level Density

As the excitation energy increases, the nuclear energy levels occur on the average at smaller and smaller intervals. In other words, level density increases with the excitation energy. The first question we might ask is how fast does this level density increase for a particular nucleus and what is the distribution of these levels with respect to spin and parity? This is an old problem treated by Bethe [1]. Even a simple model in which the nucleus is taken as a degenerate Fermi gas with equidistant single-particle levels gives an adequate result. It amounts to determining the number of partitions $\lambda(n)$ of a positive integer n into smaller positive integers ν_1, ν_2,...

$$n = \nu_1 + \nu_2 + \cdots, \qquad \nu_1 > 0, \nu_2 > 0, \ldots .$$

For large n this number, according to the Hardy–Ramanujan [1] formula, is given by

$$\lambda(n) \sim \exp[(\tfrac{1}{3}\theta\pi^2 n)^{1/2}],$$

where θ is equal to 1 or 2 according to whether the ν_i are all different or whether some of them are allowed to be equal. With a slight modification due to later work [Lang and Lecouteur, 1; Cameron, 1], Bethe's result gives the level density as

$$\rho(E, j, \pi) \propto (2j+1)(E-\Delta)^{-5/4} \exp\left[-\frac{1}{2\sigma^2} j(j+1)\right] \exp[2a(E-\Delta)^{1/2}],$$

where E is the excitation energy, j is the spin, and π is the parity. The dependence of the parameters σ, a, and Δ on the neutron and proton numbers is complicated and only imperfectly understood.

However, for any particular nucleus a few measurements will suffice to determine them all; the formula will then remain valid for a wide range of energy that contains thousands and even millions of levels.

1.2.2. Distribution of Neutron Widths

An excited level may decay in many ways; for example, by neutron ejection or by giving out a quantum of radiation. These processes are characterized by the corresponding decay widths of the levels. The neutron reduced widths $\Gamma_n^0 = \Gamma_n/E^{1/2}$, in which Γ_n is the neutron width and E is the excitation energy of the level, show large fluctuations from level to level. From an analysis of the then available data Scott [1] and later Thomas and Porter [1] concluded that they had a χ^2-distribution with $\nu = 1$ degree of freedom:

$$P(x) = [\Gamma(\tfrac{1}{2}\nu)]^{-1}\, e^{-(1/2)\nu x}(\tfrac{1}{2}\nu x)^{(1/2)\nu-1} \cdot \tfrac{1}{2}\nu = (2\pi x)^{-1/2}\, e^{-(1/2)x},$$

where $P(x)\,dx$ is the probability that a certain reduced width will lie in an interval dx around the value x. This indicates a Gaussian distribution for the reduced width amplitude

$$\left(\frac{2}{\pi}\right)^{1/2} \exp[-\tfrac{1}{2}(\sqrt{x})^2]\, d(\sqrt{x})$$

expected from the theory. In fact, the reduced width amplitude is proportional to the integral of the product of the compound nucleus wave function and the wave function in the neutron-decay channel over the channel surface. If the contributions from the various parts of the channel surface are supposed to be random and mutually independent, their sum will have a Gaussian distribution with zero mean.

1.2.3. Radiation and Fission Widths†

The total radiation width is almost a constant for particular spin states of a particular nucleus. The total radiation width is the sum of partial radiation widths

$$\Gamma = \sum_{i=1}^{m} \Gamma_i\,.$$

† Bohr [1].

If we assume that each of these $\Gamma_i/\bar{\Gamma}_i$ has a χ^2-distribution with one degree of freedom like the neutron widths and all the $\bar{\Gamma}_i$ are the same, then $\Gamma/\bar{\Gamma}$ will have a χ^2-distribution with m degrees of freedom. Even if the $\bar{\Gamma}_i$ are different, we have

$$\bar{\Gamma} = \sum_i \bar{\Gamma} i$$

and

$$\overline{(\Gamma - \bar{\Gamma})^2} = 2 \sum_i (\bar{\Gamma}_i)^2,$$

so that for m large $\Gamma/\bar{\Gamma}$ has a narrow distribution. It is difficult to measure the partial radiation widths.

Little is known about the fission-width distributions. Some known fission widths of U^{235} have been analyzed [Bohr, 1] and a χ^2-distribution with 2 to 3 degrees of freedom has been found to give a satisfactory fit.

From now on we shall no longer consider neutron, radiation, or fission widths.

1.2.4. Level Spacings

Let us regard level density as a function of the excitation energy as known and consider an interval of energy δE centered at E. This interval is much smaller compared with E, whereas it is large enough to contain many levels; that is,

$$E \gg \delta E \gg D,$$

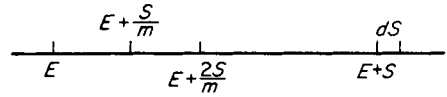
where D is the mean distance between neighboring levels. How are the levels distributed in this interval? Although the level density varies strongly from nucleus to nucleus, the fluctuations in the precise positions of the levels seem not to depend on the nucleus and not even on the excitation energy. As the density of the levels is nearly constant in this interval, we might think that they occur at random positions without regard to one another, the only condition being that their density be a given constant. However, such is not the case. It is true that nuclear levels with different spin and parity or atomic levels with different sets of good quantum numbers seem to have no influence on each other. However, levels with the same set of good quantum numbers show a large degree of regularity. For instance, they rarely occur close together.

A more detailed analysis of the experimental data regarding the above quantities as well as the strength functions may be found in Garg *et al.* [1] and Rosen *et al.* [1].

1.3. Definition of a Suitable Function for the Study of Level Correlations

To distinguish between various possibilities we define the distribution of spacings. Let E_1, E_2,..., E_n be the positions of the successive levels in the interval $\delta E (E_1 \leqslant E_2 \leqslant \cdots)$ and let S_1, S_2,... be their distances apart $S_i = E_{i+1} - E_i$. The average value of S_i is the mean spacing D. We define the relative spacings $t_i = S_i/D$. The probability density function $p(t)$ is defined by the condition that $p(t)\,dt$ is the probability that any t_i will have a value between t and $t + dt$.

For the simple case in which the positions of the energy levels are not correlated the probability that any E_i will fall between E and $E + dE$ is independent of E and is simply $\rho\,dE$, where $\rho = D^{-1}$ is the average number of levels in a unit interval of energy. Let us determine the probability of a spacing S; that is, given a level at E, what is the probability of having no level in the interval $(E, E + S)$ and one level in dS at $E + S$. For this we divide the interval S into m equal parts. Because the levels are independent, the probability of having no level in $(E, E + S)$ is the product of the



probabilities of having no level in any of these m parts. If m is large, so that S/m is small, we can write this as

$$\left(1 - \rho \frac{S}{m}\right)^m \xrightarrow[m\to\infty]{} e^{-\rho S}$$

Moreover, the probability of having a level in dS at $E + S$ is $\rho\,dS$. Therefore, given a level at E, the probability that there is no level in $(E, E + S)$ and one level in dS at $E + S$ is

$$e^{-\rho S} \rho\, dS$$

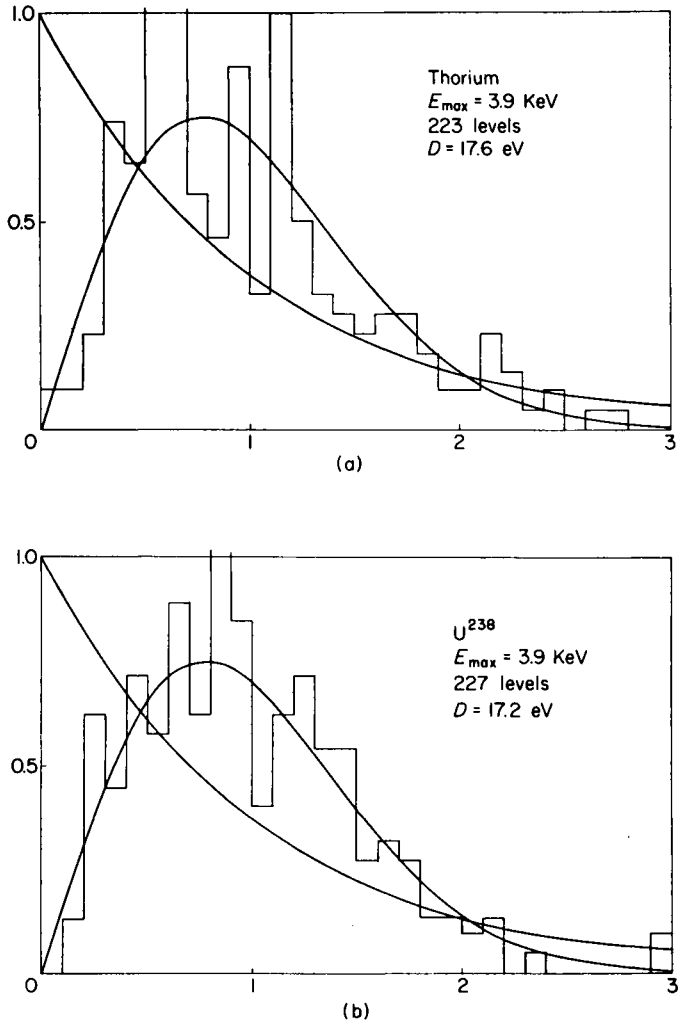


FIG. 1.1. Summary of the experimental data on nuclear level spacings for the elements Th and U^{238}. (*a*) Histogram of the observed density of level spacings as a function of $t = S/D$; the spacing is measured in units of the mean level spacing for thorium. (*b*) The same histogram for the nucleus U^{238}. The two solid curves correspond to the random and orthogonal cases. For details, see (1.1), (5.84), and (5.105). From Garg *et al.* [1].

or in terms of the variable $t = S/D = \rho S$

$$p(t)\,dt = e^{-t}\,dt. \tag{1.1}$$

This is known as the Poisson distribution or the spacing rule for random levels.

That (1.1) is not correct for nuclear levels of the same spin and parity or for atomic levels of the same parity and orbital and spin angular momenta is clearly seen by a comparison with the empirical evidence (Figures 1.1 and 1.2).

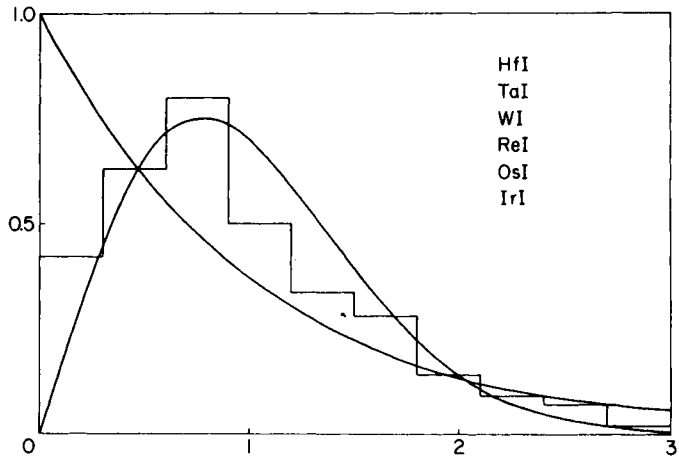


FIG. 1.2. Plot of the density of spacings between odd parity atomic levels of a group of elements in the region of osmium. The levels in each element were separated according to the angular momentum, and separate histograms were constructed for each level series, and then combined. The elements and the number of contributed spacings are HfI, 74; TaI, 180; WI, 262; ReI, 165; OsI, 145; IrI, 131, which lead to a total of 957 spacings. The solid curves correspond to the random and orthogonal cases; (1.1), (5.84), and (5.105). From Porter and Rosenzweig [1].

1.4. Wigner Surmise

When the experimental situation was not yet conclusive, Wigner [3] proposed the following rules for spacing distributions:

1. In the sequence of levels with the same spin and parity, called a simple sequence, the probability density function for a spacing is given by

$$p_W(t) = \frac{\pi}{2}\, t \exp\left(-\frac{\pi}{4}\, t^2\right), \qquad t = \frac{S}{D}. \tag{1.2}$$

2. The levels with different spin and parity are not correlated. The function $p(t)$ for a mixed sequence may be obtained by randomly superimposing the constituent simple sequences (cf. Appendix A.22).

Two simple arguments give rise to Rule 1. As pointed out by Wigner [3] and by Landau and Smorodinsky [1], it is reasonable to expect that, given a level at E, the probability that another level will lie around $E + S$ is proportional to S for very small S. Now, if we extrapolate this to all S's and, in addition, assume that the probabilities in various intervals of length S/m obtained by dividing S into m equal parts are mutually independent, we arrive at

$$p(t)\,dt = \lim_{m\to\infty} \prod_{r=0}^{m-1} \left(1 - \frac{tr}{m}\frac{1}{m}a\right) at\,dt$$

$$= at\,e^{-(1/2)at^2}\,dt. \tag{1.3}$$

The constant a can be determined by the condition that the average value of $t = S/D$ is unity:

$$\int_0^\infty t\,p(t)\,dt = 1. \tag{1.4}$$

Let us, at this point, define the n-point correlation function $R_n(E_1, \ldots, E_n)$ so that $R_n\,dE_1\,dE_2 \cdots dE_n$ is the probability of finding

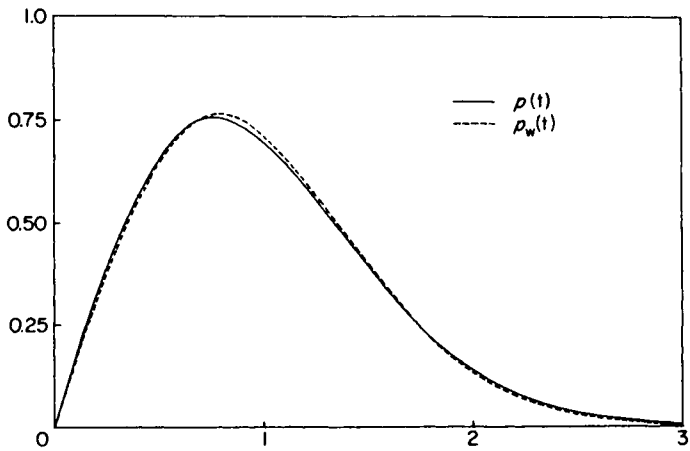


FIG. 1.3. The probability density functions $p(t)$ and $p_W(t)$; (1.2), (5.84), and (5.105)

a level in each of the intervals $(E_1, E_1 + dE_1),\ldots, (E_n, E_n + dE_n)$, all other levels being unobserved. The two simple arguments of Wigner given in the derivation of Rule 1 are equivalent to the following. The two-point correlation function $R_2(E_1, E_2)$ is linear in the variable $|E_1 - E_2|$, and three and higher order correlation functions are negligibly small.

We shall see in Chapter 5 that both arguments are inaccurate, whereas Rule 1 is very near the correct result (Figure 1.3). It is surprising that the two errors made so nearly compensate each other.

1.5. Electromagnetic Properties of Small Metallic Particles

Consider small metallic particles at low temperatures. The number of electrons in a volume V is $n \approx 4\pi p_0^3 V/3h^3$, where p_0 is the Fermi momentum and h is Planck's constant. The energy of an excitation near the Fermi surface is $E_0 \approx p_0^2/2m^*$, where m^* is the effective mass of the electron. The level density at zero excitation is therefore $\sigma = dn/dE_0 \approx 4\pi p_0 V m^*/h^3$, and the average level spacing is the inverse of this quantity $\Delta \approx \sigma^{-1}$. For a given temperature we can easily estimate the size of the metallic particles for which $\Delta \gg kT$, where k is Boltzmann's constant and T is the temperature in degrees Kelvin. For example, a metallic particle of size $10^{-6} - 10^{-7}$ cm contains $10^4 - 10^5$ electrons and, at $T \approx 10°\text{K}$, $\Delta \approx 1$ eV, whereas $kT \approx 10^{-3}$ eV. It is possible to produce particles of this size experimentally and then to sort them out according to their size (e.g., by centrifuging and sampling at a certain radial distance). Thus we have a large number of metallic particles, each of which has a different shape and therefore a different set of electronic energy levels but the same average level spacing, for the volumes are equal. It would be desirable if we could separate (e.g., by applying a nonuniform magnetic field) particles containing an odd number of conduction electrons from those containing an even number. The energy-level schemes for these two types of particle have very different properties (see Chapters 2 and 3).

Given the position of the electronic energies, we can calculate the partition function in the presence of a magnetic field and then use thermodynamic relations to derive various properties such as electronic specific heat and spin paramagnetism. Fröhlich [1] assumed that the energies were equally spaced and naturally obtained the result that all physical quantities decrease exponentially at low temperatures

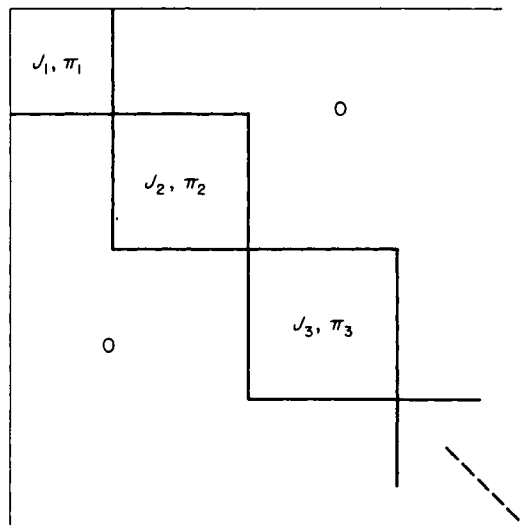
as $e^{-\Delta/kT}$ for $1 \ll \Delta/kT$. Kubo [1] repeated the calculation with the assumption that the energies were random without correlations and that their spacings therefore follow a Poisson law. He arrived at a linear law for the specific heat $\sim kT/\Delta$. The constants are different for particles containing an odd number of electrons from those containing an even number. For spin paramagnetism even the dependence on temperature is different for the two sets of particles. Instead of Fröhlich's equal spacing rule or Kubo's Poisson law, it would perhaps be better to adopt the point of view of Gorkov and Eliashberg [1], which may be justified as follows. The energies are the eigenvalues of a fixed Hamiltonian with random boundary conditions. We may incorporate these boundary conditions into the Hamiltonian by the use of fictitious potentials.

In contrast to nuclear spectra, we have the possibility of realizing in practice all three ensembles considered in various sections of this book. They apply in particular when (a) the number of electrons (in each of the metallic particles) is even and there is no external magnetic field, (b) the number of electrons (in each of the metallic particles) is odd and there is no external magnetic field, (c) there is an external magnetic field $H \gg \Delta/\mu$, where μ is the magnetic moment of the electron.

2 / Gaussian Ensembles. The Joint Probability Density Function for the Matrix Elements

2.1. Preliminaries

In the mathematical model our systems are characterized by their Hamiltonians, which in turn are represented by Hermitian matrices. Let us look into the structure of these matrices. The low-lying energy levels (eigenvalues) are far apart and each may be described by a different set of quantum numbers. As we go to higher excitations, the levels draw closer, and because of their mutual interference most of the approximate quantum numbers lose their usefulness, for they are no longer exact. At still higher excitations the interference is so great that some quantum numbers may become entirely meaningless. However, there may be certain exact integrals of motion, such as total spin or parity, and the quantum numbers corresponding to them



are conserved whatever the excitation may be. If the basic functions are chosen to be the eigenfunctions of these conserved quantities, all Hamiltonian matrices of the ensemble will reduce to the form of diagonal blocks. One block will correspond uniquely to each set of exact quantum numbers. The matrix elements lying outside these blocks will all be zero, and levels belonging to two different blocks will be statistically uncorrelated. As to the levels corresponding to the same block, the interactions are so complex that any regularity resulting from partial diagonalization will be washed out.

We shall assume that such a basis has already been chosen and restrict our attention to one of the diagonal blocks, an $(N \times N)$ Hermitian matrix in which N is a large but fixed positive integer. Because nuclear spectra contain at least hundreds of levels with the same spin and parity, we are interested in the limit of very large N.

With these preliminaries, the matrix elements may be supposed to be random variables and allowed the maximum statistical independence permitted under symmetry requirements. To specify precisely the correlations among various matrix elements we need a careful analysis of the consequences of time-reversal invariance.

2.2. Time-Reversal Invariance†

We begin by recapitulating the basic notions of time-reversal invariance. From physical considerations, the time-reversal operator is required to be antiunitary [Wigner, 1] and can be expressed, as any other antiunitary operator, in the form

$$T = KC, \tag{2.1}$$

where K is a fixed unitary operator and the operator C takes the complex conjugate of the expression following it. Thus a state under time reversal transforms to

$$\psi^R = T\psi = K\psi^*, \tag{2.2}$$

ψ^* being the complex conjugate of ψ. From the condition

$$(\phi, A\psi) = (\psi^R, A^R\phi^R)$$

for all pairs of states ψ, ϕ, and (2.2), we deduce that under time reversal an operator A transforms to

$$A^R = KA^TK^{-1}, \tag{2.3}$$

† Sections 2.2 to 2.5 are based largely on an article by Dyson [1].

where A^T is the transpose of A. A is said to be self-dual if $A^R = A$. A physical system is invariant under time reversal if its Hamiltonian is self-dual, that is, if

$$H^R = H. \tag{2.4}$$

When the representation of the states is transformed by a unitary transformation, $\psi \to U\psi$, T transforms according to

$$T \to UTU^{-1} = UTU^{+} \tag{2.5}$$

or K transforms according to

$$K \to UKU^T. \tag{2.6}$$

Because operating twice with T should leave the physical system unchanged, we have

$$T^2 = \alpha \cdot 1, \qquad |\alpha| = 1, \tag{2.7}$$

where 1 is the unit operator; or

$$T^2 = KCKC = KK^*CC = KK^* = \alpha \cdot 1, \tag{2.8}$$

But K is unitary:

$$K^*K^T = 1.$$

From these two equations we get

$$K = \alpha K^T = \alpha(\alpha K^T)^T = \alpha^2 K.$$

Therefore

$$\alpha^2 = 1 \quad \text{or} \quad \alpha = \pm 1, \tag{2.9}$$

so that the unitary matrix K is either symmetric or antisymmetric. In other words, either

$$KK^* = 1 \tag{2.10}$$

or

$$KK^* = -1. \tag{2.11}$$

These alternatives correspond, respectively, to an integral or a half-odd integral total angular momentum of the system measured

in units of $\hbar$ [Wigner, 1], for the total angular momentum operator $\mathbf{J} = (J_1, J_2, J_3)$ must transform as

$$J_l^R = -J_l, \qquad l = 1, 2, 3. \tag{2.12}$$

For brevity we call the two possibilities the even-spin and odd-spin case, respectively.

2.3. Gaussian Orthogonal Ensemble

Suppose now that the even-spin case holds and (2.10) is valid. Then a unitary operator U will exist such that (cf. Appendix A.23)

$$K = UU^T. \tag{2.13}$$

By (2.6) a transformation $\psi \to U^{-1}\psi$ performed on the states ψ brings K to unity. Thus in the even-spin case the representation of states can always be chosen so that

$$K = 1. \tag{2.14}$$

After one such representation is found, further transformations $\psi \to R\psi$ are allowed only with R a real orthogonal matrix so that (2.14) remains valid. The consequence of (2.14) is that self-dual matrices are symmetric. In the even spin case every system invariant under time reversal will be associated with a real symmetric matrix H if the representation of states is suitably chosen. For even-spin systems with time-reversal invariance the Gaussian orthogonal ensemble E_{1G}, defined below, is therefore appropriate.

Definition 2.1: The Gaussian orthogonal ensemble E_{1G} is defined in the space T_{1G} of real symmetric matrices by two requirements:

1. The ensemble is invariant under every transformation

$$H \to W^T H W \tag{2.15}$$

of T_{1G} into itself, where W is any real orthogonal matrix.

2. The various elements H_{kj}, $k \leqslant j$, are statistically independent.

These requirements, expressed in the form of equations, read as follows:

1. The probability $P(H)\,dH$ that a system of E_{1G} will belong to the volume element $dH = \prod_{k \leqslant j} dH_{kj}$ is invariant under real orthogonal transformations:

$$P(H')\,dH' = P(H)\,dH, \tag{2.16}$$

where

$$H' = W^T H W \tag{2.17}$$

and

$$W^T W = W W^T = 1. \tag{2.18}$$

2. This probability density function $P(H)$ is a product of functions, each of which depends on at most a single variable:

$$P(H) = \prod_{k \leqslant j} f_{kj}(H_{kj}). \tag{2.19}$$

Suppose, next, that we are dealing with a system invariant under space rotations. The spin may now be even or odd. The Hamiltonian matrix H which represents the system commutes with every component of $\mathbf{J}$. If we use the standard representation of the J matrices with J_1 and J_3 real and J_2 pure imaginary, (2.12) may be satisfied by the usual choice [Wigner, 1]

$$K = e^{i\pi J_2} \tag{2.20}$$

for K. With this choice of K, H and K commute and H^R reduces to H^T. Thus a rotation-invariant system is represented by a real symmetric matrix H, and once again the ensemble E_{1G} is appropriate.

2.4. Gaussian Symplectic Ensemble†

In this section we discuss a system to which E_{1G} does not apply, a system with odd-spin, invariant under time reversal, but having no rotational symmetry. In this case (2.11) holds, K cannot be diagonalized by any transformation of the form (2.6), and there is no integral of the motion by which the double-valuedness of the time-reversal operation can be trivially eliminated.

† Dyson [1].

Every antisymmetric unitary operator can be reduced by a transformation (2.6) to the standard canonical form (cf. Appendix A.23)

$$Z = \begin{bmatrix} 0 & -1 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & -1 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \equiv \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \dot{+} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \dot{+} \cdots \tag{2.21}$$

which consists of (2×2) blocks

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

along the leading diagonal; all other elements of Z are zero. We assume that the representation of states is chosen so that K is reduced to this form. The number of rows and columns of all matrices must now be even, for otherwise K would be singular in contradiction to (2.11). It is convenient to denote the order of the matrices by $2N$ instead of N. After one such representation is chosen, for which $K = Z$, further transformations $\psi \rightarrow B\psi$ are allowed, only with B a unitary $(2N \times 2N)$ matrix for which

$$Z = BZB^T. \tag{2.22}$$

Such matrices B form precisely the N-dimensional symplectic group [Weyl, 1], usually denoted by $Sp(N)$.

It is well known [Chevalley, 1; Dieudonné, 1] that the algebra of the symplectic group can be expressed most conveniently in terms of quaternions. We therefore introduce the standard quaternion notation for (2×2) matrices,

$$e_1 = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \qquad e_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \qquad e_3 = \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix}, \tag{2.23}$$

with the usual multiplication table

$$e_1^2 = e_2^2 = e_3^2 = -1, \tag{2.24}$$

$$e_1 e_2 = -e_2 e_1 = e_3, \qquad e_2 e_3 = -e_3 e_2 = e_1, \qquad e_3 e_1 = -e_1 e_3 = e_2. \tag{2.25}$$

Note that in (2.23), as well as throughout the rest of this book, i is the ordinary imaginary unit and not a quaternion unit. The matrices e_1, e_2, and e_3, together with the (2×2) unit matrix

$$1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

form a complete set, and any (2×2) matrix with complex elements can be expressed linearly in terms of them with complex coefficients:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{2}(a+d)\,1 - \frac{i}{2}(a-d)\,e_1 - \frac{1}{2}(b-c)\,e_2 + \frac{i}{2}(b+c)\,e_3 . \quad (2.26)$$

All the $(2N \times 2N)$ matrices will be considered as cut into N^2 blocks of (2×2) and each (2×2) block expressed in terms of quaternions. In general, a $(2N \times 2N)$ matrix with complex elements thus becomes an $(N \times N)$ matrix with complex quaternion elements. In particular, the matrix Z is now

$$Z = e_2 I, \quad (2.27)$$

where I is the $(N \times N)$ unit matrix. It is easy to verify that the rules of matrix multiplication are not changed by this partitioning.

Let us add some definitions. We call a quaternion "real" if it is of the form

$$q = q^{(0)} + \mathbf{q} \cdot \mathbf{e}, \quad (2.28)$$

with real coefficients $q^{(0)}$, $q^{(1)}$, $q^{(2)}$, and $q^{(3)}$. Thus a real quaternion does not correspond to a (2×2) matrix with real elements. Any complex quaternion has a "conjugate quaternion"

$$\bar{q} = q^{(0)} - \mathbf{q} \cdot \mathbf{e}, \quad (2.29)$$

which is distinct from its "complex conjugate"

$$q^* = q^{(0)*} + \mathbf{q}^* \cdot \mathbf{e}. \quad (2.30)$$

A quaternion with $q^* = q$ is real; one with $q^* = -q$ is pure imaginary; and one with $\bar{q} = q$ is a scalar. By applying both types of conjugation together, we obtain the "Hermitian conjugate"

$$q^+ = \bar{q}^* = q^{(0)*} - \mathbf{q}^* \cdot \mathbf{e}. \quad (2.31)$$

A quaternion with $q^+ = q$ is Hermitian and corresponds to the ordinary notion of a (2×2) Hermitian matrix; one with $q^+ = -q$

is anti-Hermitian. The conjugate (Hermitian conjugate) of a product of quaternions is the product of their conjugates (Hermitian conjugates) taken in the reverse order:

$$\overline{(q_1 q_2 \cdots q_n)} = \bar{q}_n \cdots \bar{q}_2 \bar{q}_1 , \tag{2.32}$$

$$(q_1 q_2 \cdots q_n)^+ = q_n{}^+ \cdots q_2{}^+ q_1{}^+ . \tag{2.33}$$

Now consider a general $(2N \times 2N)$ matrix A which is to be written as an $(N \times N)$ matrix Q with quaternion elements q_{kj}; $k, j = 1, 2, \ldots, N$. The standard matrix operations on A are then reflected in Q in the following way:

Transposition,

$$(Q^T)_{kj} = -e_2 \bar{q}_{jk} e_2 . \tag{2.34}$$

Hermitian conjugation,

$$(Q^+)_{kj} = q_{jk}^+ . \tag{2.35}$$

Time reversal,

$$(Q^R)_{kj} = e_2 (Q^T)_{kj} e_2^{-1} = \bar{q}_{jk} . \tag{2.36}$$

The matrix Q^R is called the "dual" of Q. A "self-dual" matrix is one with $Q^R = Q$.

The usefulness of quaternion algebra is a consequence of the simplicity of (2.35) and (2.36). In particular, it is noteworthy that the time-reversal operator K does not appear explicitly in (2.36) as it did in (2.3). By (2.35) and (2.36) the condition

$$Q^R = Q^+ \tag{2.37}$$

is necessary and sufficient for the elements of Q to be real quaternions. When (2.37) holds, we call Q "quaternion real."

A unitary matrix B that satisfies (2.22) is automatically quaternion real. In fact, it satisfies the conditions

$$B^R = B^+ = B^{-1}, \tag{2.38}$$

which define the symplectic group. The matrices H which represent the energy operators of physical systems are Hermitian as well as self-dual:

$$H^R = H, \qquad H^+ = H, \tag{2.39}$$

hence are also quaternion real. From (2.35) and (2.36) we see that the quaternion elements of a self-dual hermitian matrix must satisfy

$$q_{jk}^{+} = \bar{q}_{jk} = q_{kj} \tag{2.40}$$

or $q_{jk}^{(0)}$ must form a real symmetric matrix, whereas $q_{jk}^{(1)}$, $q_{jk}^{(2)}$, and $q_{jk}^{(3)}$ must form real antisymmetric matrices. Thus the number of real independent parameters that define a $(2N \times 2N)$ self-dual Hermitian matrix is

$$\tfrac{1}{2}N(N+1) + \tfrac{1}{2}N(N-1)\cdot 3 = N(2N-1).$$

From this notational excursion, let us come back to the point. Systems having odd-spin, invariance under time-reversal, but no rotational symmetry, must be represented by self-dual, Hermitian Hamiltonians. Therefore the Gaussian symplectic ensemble, as defined below, should be appropriate for their description.

Definition 2.2: The Gaussian symplectic ensemble E_{4G} is defined in the space T_{4G} of self-dual Hermitian matrices by the following properties:

1. The ensemble is invariant under every orthomorphism

$$H \to W^R H W \tag{2.41}$$

of T_{4G} into itself, where W is any symplectic matrix.

2. Various linearly independent components of H are also statistically independent.

These requirements put in the form of equations read as follows:

1. The probability $P(H)\,dH$ that a system E_{4G} will belong to the volume element

$$dH \equiv \prod_{k \leqslant j} dH_{kj}^{(0)} \prod_{\lambda=1}^{3} \prod_{k<j} dH_{kj}^{(\lambda)} \tag{2.42}$$

is invariant under symplectic transformations; that is,

$$P(H')\,dH' = P(H)\,dH \tag{2.43}$$

if

$$H' = W^R H W, \tag{2.44}$$

where

$$WZW^T = Z \tag{2.45}$$

2. The probability density function $P(H)$ is a product of functions each of which depends on a single variable:

$$P(H) = \prod_{k \leqslant j} f_{kj}^{(0)}(H_{kj}^{(0)}) \prod_{\lambda=1}^{3} \prod_{k<j} f_{kj}^{(\lambda)}(H_{kj}^{(\lambda)}). \tag{2.46}$$

2.5. Gaussian Unitary Ensemble

For completeness we discuss briefly a much simpler ensemble, the Gaussian unitary ensemble E_{2G} which applies to systems without invariance under time reversal. Such systems are easily created in principle by putting an ordinary atom or nucleus, for example, into an externally generated magnetic field. The external field is not affected by the time-reversal operation. However, for the unitary ensemble to be applicable, the splitting of levels by the magnetic field must be at least as large as the average level spacing in the absence of the magnetic field. The magnetic field must, in fact, be so strong that it will completely "mix up" the level structure that would exist in zero field; for otherwise our random hypothesis cannot be justified. This state of affairs could never occur in nuclear physics. In atomic or molecular physics a practical application of the unitary ensemble may perhaps be possible.

A system without time-reversal invariance has a Hamiltonian that may be an arbitrary Hermitian matrix not restricted to be real or self-dual. Thus we are led to the following definition.

Definition 2.3: The Gaussian unitary ensemble E_{2G} is defined in the space of Hermitian matrices by the following properties:

1. The probability $P(H)\,dH$ that a system of E_{2G} will belong to the volume element

$$dH = \prod_{k \leqslant j} dH_{kj}^{(0)} \prod_{k<j} dH_{kj}^{(1)}, \tag{2.47}$$

where $H_{kj}^{(0)}$ and $H_{kj}^{(1)}$ are real and imaginary parts of H_{kj}, is invariant under every automorphism

$$H \to U^{-1}HU \tag{2.48}$$

of T_{2G} into itself, where U is any unitary matrix.

2. Various linearly independent components of H are also statistically independent.

In mathematical language these requirements are

1. $$P(H')\,dH' = P(H)\,dH, \tag{2.49}$$

$$H' = U^{-1}HU, \tag{2.50}$$

where U is any unitary matrix.

2. $P(H)$ is a product of functions, each of which depends on a single variable:

$$P(H) = \prod_{k \leqslant j} f_{kj}^{(0)}(H_{kj}^{(0)}) \prod_{k<j} f_{kj}^{(1)}(H_{kj}^{(1)}). \tag{2.51}$$

2.6. Joint Probability Density Function for Matrix Elements

We now come to the question of the extent to which we are still free to specify the joint probability density function $P(H)$. It will be seen that the two postulates of invariance and statistical independence elaborated above fix uniquely the functional form of $P(H)$.

The postulate of invariance restricts $P(H)$ to depend only on a finite number of traces of the powers of H. We state this fact as a lemma [Weyl, 1].

Lemma 2.1. *All the invariants of an* $(N \times N)$ *matrix* H *under nonsingular similarity transformations* A,

$$H \to H' = AHA^{-1},$$

can be expressed in terms of the traces of the first N *powers of* H.

Proof: Because all invariants are symmetric functions of the eigenvalues λ_k, $k = 1, 2, \ldots, N$, of H, and

$$\operatorname{tr} H^j = \sum_{k=1}^{N} \lambda_k{}^j \equiv t_j, \quad \text{say,}$$

we need to show that any symmetric function of λ_k can be expressed in terms of the first N of the t_j. Let the secular equation which determines the λ_k be

$$\det[H - \lambda I] \equiv (-\lambda)^N + \sigma_1(-\lambda)^{N-1} + \cdots + \sigma_N = 0,$$

so that, given the coefficients

$$\sigma_j = \sum_{1 \leqslant k_1 < \cdots < k_j \leqslant N} \lambda_{k_1} \lambda_{k_2} \cdots \lambda_{k_j}, \qquad 1 \leqslant j \leqslant N,$$

all the eigenvalues λ_k are uniquely determined except for their order. Thus any symmetric function of the λ_k can be expressed in terms of the basic functions $\sigma_1, \sigma_2, \ldots, \sigma_N$. To show that the t_j form another such basis it will then be sufficient to express σ_j in terms of t_j. This is achieved by the equation (Appendix A.1)

$$\sigma_r = (r!)^{-1} \det[a_{kj}]_{k,j=1,2,\ldots,r}; \qquad 1 \leqslant r \leqslant N, \tag{2.52}$$

where

$$a_{kj} = \begin{cases} t_{j-k+1}, & \text{if } k \leqslant j, \\ j, & \text{if } k = j+1, \\ 0, & \text{if } k > j+1. \end{cases} \tag{2.53}$$

Incidently, we note that

$$\det[a_{kj}]_{k,j=1,2,\ldots,r} \equiv 0, \quad \text{if } r > N, \tag{2.54}$$

which expresses the traces of all the powers of H in terms of those of the first N powers.

The postulate of statistical independence excludes everything except the traces of the first two powers, and these, too, may occur only in an exponential. To see this we will need the following lemma.

Lemma 2.2. *If three continuous and differentiable functions $f_k(x)$; $k = 1, 2, 3$, satisfy the equation*

$$f_1(xy) = f_2(x) + f_3(y), \tag{2.55}$$

they are necessarily of the form $a \ln x + b_k$ with $b_1 = b_2 + b_3$.

Proof: Differentiating (2.55) with respect to x, we have

$$f_1'(xy) = \frac{1}{y} f_2'(x),$$

which, on integration with respect to y, gives

$$\frac{1}{x} f_1(xy) = f_2'(x) \ln y + \frac{1}{x} g(x), \tag{2.56}$$

where $g(x)$ is still arbitrary. Substituting $f_1(xy)$ from (2.56) into (2.55), we get

$$x f_2'(x) \ln y + g(x) - f_2(x) = f_3(y). \tag{2.57}$$

Therefore the left-hand side of (2.57) must be independent of x; this is possible only if

$$x f_2'(x) = a \qquad \text{and} \qquad g(x) - f_2(x) = b_3 ,$$

that is, only if

$$f_2(x) = a \ln x + b_2 = g(x) - b_3 ,$$

where a, b_2, and b_3 are arbitrary constants.

Now (2.57) gives

$$f_3(y) = a \ln y + b_3$$

and finally (2.55) gives

$$f_1(xy) = a \ln(xy) + (b_2 + b_3). \qquad \text{Q.E.D.}$$

Let us now examine the consequences of the statistical independence of the various components of H. Consider the particular transformation

$$H = \grave{U}^{-1} H' U, \tag{2.58}$$

where

$$U = \begin{bmatrix} \cos\theta & \sin\theta & 0 \cdots 0 \\ -\sin\theta & \cos\theta & 0 \cdots 0 \\ 0 & 0 & 1 \cdots 0 \\ \cdots & \cdots & \cdots \\ 0 & 0 & 0 \cdots 1 \end{bmatrix} \tag{2.59}$$

or, in quaternion notation (provided N is even),

$$U = \begin{bmatrix} \cos\theta - e_2 \sin\theta & 0 \cdots 0 \\ 0 & 1 \cdots 0 \\ \cdots & \cdots \\ 0 & 0 \cdots 1 \end{bmatrix}. \tag{2.60}$$

This U is, at the same time, orthogonal, symplectic, and unitary.

Differentiation of (2.58) with respect to θ gives

$$\begin{aligned} \frac{\partial H}{\partial \theta} &= \frac{\partial U^T}{\partial \theta} H' U + U^T H' \frac{\partial U}{\partial \theta} \\ &= \frac{\partial U^T}{\partial \theta} U H + H U^T \frac{\partial U}{\partial \theta} , \end{aligned} \tag{2.61}$$

and by substituting for U, U^T, $\partial U/\partial\theta$, and $\partial U^T/\partial\theta$ from (2.59) or (2.60) we get

$$\frac{\partial H}{\partial\theta} = AH + HA^T, \tag{2.62}$$

where

$$A = \frac{\partial U^T}{\partial\theta}\, U = \begin{bmatrix} 0 & -1 & 0\cdots 0 \\ 1 & 0 & 0\cdots 0 \\ 0 & 0 & 0\cdots 0 \\ \cdot & \cdot & \cdot\ \cdot\ \cdot \\ 0 & 0 & 0\cdots 0 \end{bmatrix} \tag{2.63}$$

or, in quaternion notation, A is diagonal.

$$A = \begin{bmatrix} e_2 & 0\cdots 0 \\ 0 & 0\cdots 0 \\ \cdot & \cdot\ \cdot\ \cdot \\ 0 & 0\cdots 0 \end{bmatrix}. \tag{2.64}$$

If the probability density function

$$P(H) = \prod f_{kj}^{(\alpha)}(H_{kj}^{(\alpha)}) \tag{2.65}$$

is invariant under the transformation U, its derivative with respect to θ must vanish; that is

$$\sum \frac{1}{f_{kj}^{(\alpha)}} \frac{\partial f_{kj}^{(\alpha)}}{\partial H_{kj}^{(\alpha)}} \frac{\partial H_{kj}^{(\alpha)}}{\partial\theta} = 0. \tag{2.66}$$

Let us write this equation explicitly, say, for the unitary case. Equations (2.62) and (2.66) give

$$\begin{aligned} &\left\{\left[-\frac{1}{f_{11}^{(0)}}\frac{\partial f_{11}^{(0)}}{\partial H_{11}^{(0)}} + \frac{1}{f_{22}^{(0)}}\frac{\partial f_{22}^{(0)}}{\partial H_{22}^{(0)}}\right][2H_{12}^{(0)}] + \frac{1}{f_{12}^{(0)}}\frac{\partial f_{12}^{(0)}}{\partial H_{12}^{(0)}}[H_{11}^{(0)} - H_{22}^{(0)}]\right\} \\ &+ \sum_{i=3}^{N}\left[-\frac{1}{f_{1k}^{(0)}}\frac{\partial f_{1k}^{(0)}}{\partial H_{1k}^{(0)}} H_2^{(0)} + \frac{1}{f_{2k}^{(0)}}\frac{\partial f_{2k}^{(0)}}{\partial H_{2k}^{(0)}} H_{1k}^{(0)}\right] \\ &+ \sum_{i=3}^{N}\left[-\frac{1}{f_{1k}^{(1)}}\frac{\partial f_{1k}^{(1)}}{\partial H_{1k}^{(1)}} H_{2k}^{(1)} + \frac{1}{f_{2k}^{(1)}}\frac{\partial f_{2k}^{(1)}}{\partial H_{2k}^{(1)}} H_{1k}^{(1)}\right] = 0. \end{aligned} \tag{2.67}$$

The braces at the left-hand side of this equation depend on mutually exclusive sets of variables and their sum is zero. Therefore each must be a constant; for example,

$$-\frac{H_{2k}^{(0)}}{f_{1k}^{(0)}}\frac{\partial f_{1k}^{(0)}}{\partial H_{1k}^{(0)}}+\frac{H_{1k}^{(0)}}{f_{2k}^{(0)}}\frac{\partial f_{2k}^{(0)}}{\partial H_{2k}^{(0)}}=C_k^{(0)}. \tag{2.68}$$

On dividing both sides of (2.68) by $H_{1k}^{(0)}\,H_{2k}^{(0)}$ and applying the Lemma 2.2, we conclude that the constant $C_k^{(0)}$ must be zero, that is,

$$\frac{1}{H_{1k}^{(0)}}\frac{1}{f_{1k}^{(0)}}\frac{\partial f_{1k}^{(0)}}{\partial H_{1k}^{(0)}}=\frac{1}{H_{2k}^{(0)}}\frac{1}{f_{2k}^{(0)}}\frac{\partial f_{2k}^{(0)}}{\partial H_{2k}^{(0)}}=\text{constant}$$

$$=-2a, \text{ say}, \tag{2.69}$$

which on integration gives

$$f_{1k}^{(0)}[H_{1k}^{(0)}]=\exp\{-a[H_{1k}^{(0)}]^2\}. \tag{2.70}$$

In the other two cases we also derive a similar equation. Now because the off-diagonal elements come only as squares in the exponential and all invariants are expressible in terms of the traces of powers of H, the function $P(H)$ is an exponential that contains traces of at most the second power of H.

Because $P(H)$ is required to be invariant under more general transformations than we have here considered, one might think that the form of $P(H)$ is further restricted. This, however, is not so, for

$$P(H)=\exp(-a\,\mathrm{tr}\,H^2+b\,\mathrm{tr}\,H+C)$$

$$=e^C\prod_{k\leqslant j}\exp\{-a[H_{kj}^{(0)}]^2+bH_{jj}^{(0)}\}\prod_{k<j,\lambda}\exp\{-a[H_{kj}^{(\lambda)}]^2\} \tag{2.71}$$

is already a product of functions, each of which depends on a separate variable. Moreover, because we require $P(H)$ to be normalizable and real, a must be real and positive and b and c must be real.

Therefore we have proved the following theorem [Porter and Rosenzweig, 1; Wishart, 1].

Theorem 2.1. *In all the above three cases the form of $P(H)$ is automatically restricted to*

$$P(H)=\exp(-a\,\mathrm{tr}\,H^2+b\,\mathrm{tr}\,H+c), \tag{2.72}$$

where a is real and positive and b and c are real.

In the foregoing discussion we have emphasized the postulate of statistical independence of various components of H even at the risk of frequent repetitions. This statistical independence is important in restricting $P(H)$ to the simple form (2.72), and hence makes the subsequent analytical work tractable. However, it lacks a clear physical motivation and therefore looks somewhat artificial.

The main objection to the assumption of statistical independence, leading to (2.72), is that all values of $H_{kj}^{(\lambda)}$ are not equally weighted and therefore do not correspond to all "interactions" being "equally probable." By a formal change Dyson [1–6] has defined his "circular ensembles," which are esthetically more satisfactory to some people and equally easy to work with. We shall come to them in Chapters 8 to 11. On the other hand, Rosenzweig [1] has emphasized the "fixed strength" ensemble. Others [Leff, 1; Fox and Kahn, 1] have arbitrarily tried the so-called "generalized" ensembles. A brief review of these topics is given in Chapter 17.

3 / Gaussian Ensembles. The Joint Probability Density Function for the Eigenvalues†

3.1. Orthogonal Ensemble

The joint probability density function (abbreviated j.p.d.f. later in the chapter) for the eigenvalues $\theta_1, \theta_2, \ldots, \theta_N$ can be obtained from (2.72) by expressing the various components of H in terms of the N eigenvalues θ_j and other mutually independent variables, p_μ, say, which together with the θ_j form a complete set. In an $(N \times N)$ real symmetric matrix the number of independent real parameters which determine all H_{kj} is $\frac{1}{2}N(N+1)$. We may take these as H_{kj} with $k \leqslant j$. The number of extra parameters p_μ needed is therefore

$$l = \tfrac{1}{2}N(N+1) - N = \tfrac{1}{2}N(N-1). \tag{3.1}$$

Because

$$\operatorname{tr} H^2 = \sum_1^N \theta_j^2, \qquad \operatorname{tr} H = \sum_1^N \theta_j, \tag{3.2}$$

the probability that the N roots and the $\frac{1}{2}N(N-1)$ parameters will occur in unit intervals around $\theta_1, \ldots, \theta_N$ and $p_1, p_2, \ldots, p_l$ is, according to (2.72),

$$\Delta(\theta_1, \ldots, \theta_N ; p_1, \ldots, p_l) = \exp\left(-a\sum_1^N \theta_j^2 + b\sum_1^N \theta_j + c\right) J(\theta, p), \tag{3.3}$$

where J is the Jacobian

$$J(\theta, p) = \left| \frac{\partial(H_{11}, H_{12}, \ldots, H_{NN})}{\partial(\theta_1, \ldots, \theta_N, p_1, \ldots, p_l)} \right|. \tag{3.4}$$

Hence the j.p.d.f. of the eigenvalues θ_j can be obtained by integrating (3.3) over the parameters $p_1, \ldots, p_l$. It is usually possible to choose

† This chapter is based largely on Wigner's article [6].

these parameters so that the Jacobian (3.4) becomes a product of a function of θ_j and a function of p_μ. If this is the case, the integration provides the required j.p.d.f. as a product of the exponential in (3.3), the aforementioned function of the θ_j and a constant. The constant can then be absorbed in c in the exponential.

To define the parameters p_μ [Wigner, 6] we recollect that any real symmetric matrix H can be diagonalized by a real orthogonal matrix [Wigner, 2]:

$$H = U\Theta U^{-1} \tag{3.5}$$

$$= U\Theta U^T, \tag{3.5'}$$

where Θ is the diagonal matrix with diagonal elements $\theta_1, \theta_2, \ldots, \theta_N$ arranged in some order, say, $\theta_1 \leqslant \theta_2 \leqslant \cdots \leqslant \theta_N$, and U is a real orthogonal matrix

$$UU^T = U^TU = 1, \tag{3.6}$$

whose columns are the normalized eigenvectors of H. They are, or may be chosen to be, mutually orthogonal. To define U completely we must in some way fix the phases of the eigenvectors—for instance by requiring that the first nonvanishing component be positive. Thus U depends on $\frac{1}{2}N(N-1)$ real parameters and may be chosen to be the U_{kj}, $k > j$. If H has multiple eigenvalues, further conditions are needed to fix U completely. It is not necessary to specify them, for they apply only in regions of lower dimensionality which are irrelevant to the probability density function. At any rate, the $\frac{1}{2}N(N-1)$ parameters p_μ are supposed to characterize the U which is subject to the preceding conditions. Once this is done, the matrix H, which completely determines the Θ and the U subject to the preceding conditions, also determines the θ_j and the p_μ uniquely. Conversely, the θ_j and p_μ completely determine the U and Θ, and hence by (3.5) all the matrix elements of H.

Differentiating (3.6), we get

$$\frac{\partial U^T}{\partial p_\mu} U + U^T \frac{\partial U}{\partial p_\mu} = 0, \tag{3.7}$$

and because the two terms in (3.7) are the Hermitian conjugates of each other,

$$S_\mu = U^+ \frac{\partial U}{\partial p} = -\frac{\partial U^+}{\partial p_\mu} U \tag{3.8}$$

is an antisymmetric matrix.

Also from (3.5) we have

$$\frac{\partial H}{\partial p_\mu} = \frac{\partial U}{\partial p_\mu}\Theta U^T + U\Theta\frac{\partial U^T}{\partial p_\mu}. \tag{3.9}$$

On multiplying (3.9) by U^T on the left and by U on the right, we get

$$U^T\frac{\partial H}{\partial p}U = S^{(\mu)}\Theta - \Theta S^{(\mu)}. \tag{3.10}$$

In terms of its components, (3.10) reads

$$\sum_{j,k}\frac{\partial H_{jk}}{\partial p_\mu}U_{j\alpha}U_{k\beta} = S^{(\mu)}_{\alpha\beta}(\theta_\beta - \theta_\alpha). \tag{3.11}$$

In a similar way, by differentiating (3.5) with respect to θ_γ,

$$\sum_{j,k}\frac{\partial H_{jk}}{\partial \theta_\gamma}U_{j\alpha}U_{k\beta} = \frac{\partial\Theta_{\alpha\beta}}{\partial\theta_\gamma} = \delta_{\alpha\beta}\,\delta_{\alpha\gamma}. \tag{3.12}$$

The matrix of the Jacobian in (3.4) can be written in the partitioned form as

$$[J(\theta, p)] = \begin{bmatrix} \dfrac{\partial H_{jj}}{\partial\theta_\gamma} & \dfrac{\partial H_{jk}}{\partial\theta_\gamma} \\ \dfrac{\partial H_{jj}}{\partial p_\mu} & \dfrac{\partial H_{jk}}{\partial p_\mu} \end{bmatrix}. \tag{3.13}$$

The two columns in (3.13) correspond to N and $\frac{1}{2}N(N-1)$ actual columns: $1 \leqslant j < k \leqslant N$. The two rows in (3.13) correspond again to N and $\frac{1}{2}N(N-1)$ actual rows: $\gamma = 1, 2, \ldots, N$; $\mu = 1, 2, \cdots, \frac{1}{2}N(N-1)$. If we multiply the $[J]$ in (3.13) on the right by the

$$\tfrac{1}{2}N(N+1) \times \tfrac{1}{2}N(N+1)$$

matrix written in the partitioned form as

$$[V] = \begin{bmatrix} (U_{j\alpha}U_{j\beta}) \\ (U_{j\alpha}U_{k\beta}) \end{bmatrix}, \tag{3.14}$$

in which the two rows correspond to N and $\frac{1}{2}N(N-1)$ actual rows, $1 \leqslant j < k \leqslant N$, and the column corresponds to $\frac{1}{2}N(N+1)$ actual columns, $1 \leqslant \alpha \leqslant \beta \leqslant N$, we get by using (3.11) and (3.12)

$$[J][V] = \begin{bmatrix} \delta_{\alpha\beta}\,\delta_{\alpha\gamma} \\ S^{(\mu)}_{\alpha\beta}(\theta_\beta - \theta_\alpha) \end{bmatrix}. \tag{3.15}$$

The two rows on the right-hand side correspond to N and $\frac{1}{2}N(N-1)$ actual rows and the column corresponds to $\frac{1}{2}N(N+1)$ actual columns. Taking the determinant on both sides of (3.15), we have

$$J(\theta, p) \det V = \prod_{\alpha<\beta} (\theta_\beta - \theta_\alpha) \det \begin{bmatrix} \delta_{\alpha\beta}\, \delta_{\alpha\gamma} \\ S^{\mu}_{\alpha\beta} \end{bmatrix}$$

or

$$J(\theta, p) = \prod_{\alpha<\beta} | \theta_\beta - \theta_\alpha | f(p), \tag{3.16}$$

where $f(p)$ is independent of the θ_j and depends only on the parameters p_μ.

By inserting this result in (3.3) and integrating over the variables p_μ we get the j.p.d.f. for the eigenvalues of the matrices of an orthogonal ensemble

$$P(\theta_1, ..., \theta_N) = \exp\left[-\sum_1^N (a\theta_j^2 - b\theta_j - c)\right] \prod_{j<k} | \theta_k - \theta_j |, \tag{3.17}$$

where c is some new constant. Moreover, if we shift the origin of the θ to $b/2a$ and change the energy scale everywhere by a constant factor $\sqrt{2a}$, we may replace θ_j with $(1/\sqrt{2a})\, x_j + b/2a$. By this formal change (3.17) takes the simpler form

$$P_{N1}(x_1, ..., x_N) = C_{N1} \exp\left(-\frac{1}{2}\sum_1^N x_j^2\right) \prod_{j<k} | x_j - x_k |, \tag{3.18}$$

where C_{N1} is a constant.

3.2. Symplectic Ensemble

As the analysis is almost identical in all three cases, we have presented the details for one particular ensemble—the orthogonal one. Here and in the following discussion we indicate briefly the modifications necessary to arrive at the required j.p.d.f. in the other two cases.

Corresponding to the result that a real symmetric matrix can be diagonalized by a real orthogonal matrix, we have the following:

Theorem 3.1. *Given a quaternion-real, self-dual matrix H, there exists a symplectic matrix U such that*

$$H = U\Theta U^{-1} = U\Theta U^{R}, \tag{3.19}$$

where Θ is diagonal, real, and scalar (cf. Appendix A.23).

The fact that Θ is scalar means that it consists of N blocks of the form

$$\begin{bmatrix} \theta_j & 0 \\ 0 & \theta_j \end{bmatrix} \tag{3.20}$$

along the main diagonal. Thus the eigenvalues of H consist of N equal pairs. The Hamiltonian of any system which is invariant under time reversal, which has odd spin, and no rotational symmetry satisfies the conditions of Theorem 3.1. All energy levels of such a system will be doubly degenerate. This is the Kramer's degeneracy [Kramer, 1], and Theorem 3.1 shows how it appears naturally in the quaternion language.

Apart from the N eigenvalues θ_j, the number of real independent parameters p_μ needed to characterize an $N \times N$ quaternion-real, self-dual matrix H is

$$l = 4 \cdot \tfrac{1}{2}N(N-1) = 2N(N-1). \tag{3.21}$$

Equations 3.2 and 3.3 are replaced, respectively, by

$$\operatorname{tr} H^2 = 2\sum_1^N \theta_j^2, \qquad \operatorname{tr} H = 2\sum_1^N \theta_j \tag{3.22}$$

and

$$\Delta(\theta_1, \ldots, \theta_N ; p_1, \ldots, p_l) = \exp\Big[-\sum_1^N (2a\theta_j^2 - 2b\theta_j - c)\Big] J(\theta, p), \tag{3.23}$$

where $J(\theta, p)$ is now given by

$$J(\theta, p) = \frac{\partial(H_{11}^{(0)}, \ldots, H_{NN}^{(0)}, H_{12}^{(0)}, \ldots, H_{12}^{(3)}, \ldots, H_{N-1,N}^{(0)}, \ldots, H_{N-1,N}^{(3)})}{\partial(\theta_1, \ldots, \theta_N, p_1, \ldots, p_{2N(N-1)})}. \tag{3.24}$$

Equation 3.5 is replaced by (3.19); (3.6), (3.7), (3.8), (3.9), and (3.10) are valid if U^T is replaced by U^R. Note that these equations are now in the quaternion language, and we need to separate the four quaternion parts of modified (3.19). For this we let

$$H_{jk} = H_{jk}^{(0)} + H_{jk}^{(1)} e_1 + H_{jk}^{(2)} e_2 + H_{jk}^{(3)} e_3, \tag{3.25}$$

$$S_{\alpha\beta}^{(\mu)} = S_{\alpha\beta}^{(0\mu)} + S_{\alpha\beta}^{(1\mu)} e_1 + S_{\alpha\beta}^{(2\mu)} e_2 + S_{\alpha\beta}^{(3\mu)} e_3, \tag{3.26}$$

and write (3.10) and the one corresponding to (3.12) in the form of partitioned matrices:

$$\begin{bmatrix} \dfrac{\partial H_{jj}^{(0)}}{\partial \theta_\gamma} & \dfrac{\partial H_{jk}^{(0)}}{\partial \theta_\gamma} & \dfrac{\partial H_{jk}^{(1)}}{\partial \theta_\gamma} \cdots \dfrac{\partial H_{jk}^{(3)}}{\partial \theta_\gamma} \\ \dfrac{\partial H_{jj}^{(0)}}{\partial p_\mu} & \dfrac{\partial H_{jk}^{(0)}}{\partial p_\mu} & \dfrac{\partial H_{jk}^{(1)}}{\partial p_\mu} \cdots \dfrac{\partial H_{jk}^{(3)}}{\partial p_\mu} \end{bmatrix} \cdot \begin{bmatrix} v & w \\ A^{(0)} & B^{(0)} \\ \cdot & \cdot \\ \cdot & \cdot \\ A^{(3)} & B^{(3)} \end{bmatrix}$$

$$= \begin{bmatrix} \rho_{\gamma,\alpha\alpha} & \sigma_{\gamma,\alpha\beta}^{(0)} \cdots \sigma_{\gamma,\alpha\beta}^{(3)} \\ \epsilon_{\alpha\alpha}^{(\mu)} & S_{\alpha\beta}^{(0\mu)}(\theta_\beta - \theta_\alpha) \cdots S_{\alpha\beta}^{(3\mu)}(\theta_\beta - \theta_\alpha) \end{bmatrix}, \tag{3.27}$$

$$1 \leqslant j < k \leqslant N, \quad 1 \leqslant \alpha < \beta \leqslant N, \quad 1 \leqslant \gamma \leqslant N, \quad 1 \leqslant \mu \leqslant 2N(N-1),$$

where the matrices $\partial H_{jj}^{(0)}/\partial \theta_\gamma$, v, and ρ are $N \times N$, the matrices $\partial H_{jk}^{(\lambda)}/\partial \theta_\gamma$ and $\sigma_{\gamma,\alpha\beta}^{(\lambda)}$, with $\lambda = 0, 1, 2, 3$, are $N \times \frac{1}{2}N(N-1)$, the $A^{(\lambda)}$ are all $\frac{1}{2}N(N-1) \times N$, the $\partial H_{jj}^{(0)}/\partial p_\mu$ and the $\epsilon_{\alpha\alpha}^{(\mu)}$ are $2N(N-1) \times N$, the w is $N \times 2N(N-1)$, the $\partial H_{jk}^{(\lambda)}/\partial p_\mu$ and the $S_{\alpha\beta}^{(\lambda\mu)}$ are $2N(N-1) \times \frac{1}{2}N(N-1)$ and the matrices $B^{(\lambda)}$ are $\frac{1}{2}N(N-1) \times 2N(N-1)$. The matrices ρ and the σ appear as we separate the result of differentiation of (3.19) with respect to θ_γ into quaternion components. Because Θ is diagonal and scalar, the $\sigma^{(\lambda)}$ are all zero matrices. Moreover, the matrix ρ does not depend on θ_γ, for Θ depends linearly on the θ_γ. The computation of the matrices v, w, $A^{(\lambda)}$, and $B^{(\lambda)}$ is straightforward, but we do not require them. All we need is to note that they are formed of the various components of U, hence do not depend on θ_γ.

Now we take the determinant on both sides of (3.27). The determinant of the first matrix on the left is the Jacobian (3.24). Because the $\sigma^{(\lambda)}$ are all zero, the determinant of the right-hand side breaks into a product of two determinants:

$$\det[\rho_{\gamma,\alpha\alpha}] \det[S_{\alpha\beta}^{(\lambda\mu)}(\theta_\beta - \theta_\alpha)], \tag{3.28}$$

the first one being independent of the θ_γ, whereas the second is

$$\prod_{\alpha<\beta} (\theta_\beta - \theta_\alpha)^4 \det[S_{\alpha\beta}^{(\lambda\mu)}]. \tag{3.29}$$

Thus

$$J(\theta, p) = \prod_{\alpha<\beta} (\theta_\beta - \theta_\alpha)^4 f(p), \tag{3.30}$$

which corresponds to (3.16).

By inserting (3.30) into (3.23) and integrating over the parameters, we obtain the j.p.d.f.

$$P(\theta_1, ..., \theta_N) = \exp\left(-2a \sum_1^N \theta_j^2 + 2b \sum_1^N \theta_j + c\right) \prod_{j<k} (\theta_j - \theta_k). \tag{3.31}$$

As before, we may shift the origin to make $b = 0$ and change the scale of energy to make $a = 1$. Thus the j.p.d.f. for the eigenvalues of matrices in the symplectic ensemble in its simple form is

$$P_{N4}(x_1, ..., x_N) = C_{N4} \exp\left(-2 \sum_1^N x_j^2\right) \prod_{j<k} (x_j - x_k)^4, \tag{3.32}$$

where C_{N4} is a constant.

3.3. Unitary Ensemble

In addition to the real eigenvalues, the number of real independent parameters p_μ needed to specify an arbitrary Hermitian matrix H completely is $N(N-1)$. Equations 3.2 and 3.3 remain unchanged, but (3.4) is replaced by

$$J(\theta, p) = \frac{\partial(H_{11}^{(0)}, ..., H_{NN}^{(0)}, H_{12}^{(0)}, H_{12}^{(1)}, ..., H_{N-1,N}^{(0)}, H_{N-1,N}^{(1)})}{\partial(\theta_1, ..., \theta_N, p_1, ..., p_{N(N-1)})}, \tag{3.33}$$

where $H_{jk}^{(0)}$ and $H_{jk}^{(1)}$ are the real and imaginary parts of H_{jk}. Equations 3.5 to 3.10 are valid if U^T is replaced by U^+. Instead of (3.11) and (3.12), we now have

$$\sum_{j,k} \frac{\partial H_{jk}}{\partial p_\mu} U_{j\alpha}^* U_{k\beta} = S_{\alpha\beta}^{(\mu)}(\theta_\beta - \theta_\alpha), \tag{3.34}$$

$$\sum_{j,k} \frac{\partial H_{jk}}{\partial \theta_\gamma} U_{j\alpha}^* U_{k\beta} = \frac{\partial \Theta_{\alpha\beta}}{\partial \theta_\gamma} = \delta_{\alpha\beta} \delta_{\alpha\gamma}. \tag{3.35}$$

By separating the real and imaginary parts we may write these equations in partitioned matrix notation as

$$\begin{bmatrix} \dfrac{\partial H_{jj}^{(0)}}{\partial \theta_\gamma} & \dfrac{\partial H_{jk}^{(0)}}{\partial \theta_\gamma} & \dfrac{\partial H_{jk}^{(1)}}{\partial \theta_\gamma} \\ \dfrac{\partial H_{jj}^{(0)}}{\partial p_\mu} & \dfrac{\partial H_{jk}^{(0)}}{\partial p_\mu} & \dfrac{\partial H_{jk}^{(1)}}{\partial p_\mu} \end{bmatrix} \cdot \begin{bmatrix} v & w \\ A^{(0)} & B^{(0)} \\ A^{(1)} & B^{(1)} \end{bmatrix}$$

$$= \begin{bmatrix} \rho_{\gamma,\alpha\alpha} & \sigma_{\gamma,\alpha\beta}^{(0)} & \sigma_{\gamma,\alpha\beta}^{(1)} \\ \epsilon_{\alpha\alpha}^{(\mu)} & S_{\alpha\beta}^{(0\mu)}(\theta_\beta - \theta_\alpha) & S_{\alpha\beta}^{(1\mu)}(\theta_\beta - \theta_\alpha) \end{bmatrix},$$

$$1 \leqslant j < k \leqslant N, \qquad 1 \leqslant \alpha < \beta \leqslant N, \tag{3.36}$$

$$1 \leqslant \mu \leqslant N(N-1), \qquad 1 \leqslant \gamma \leqslant N,$$

where $S_{\alpha\beta}^{(0\alpha)}$ and $S_{\alpha\beta}^{(1\mu)}$ are the real and imaginary parts of $S_{\alpha\beta}^{(\mu)}$. The matrices $\partial H_{jj}^{(0)}/\partial\theta_\gamma$, v, and ρ are $N \times N$; the $\partial H_{jk}^{(\lambda)}/\partial\theta_\gamma$ and the $\sigma_{\gamma,\alpha\beta}^{(\lambda)}$ are $N \times \frac{1}{2}N(N-1)$; the $A^{(\lambda)}$ are $\frac{1}{2}N(N-1) \times N$; the $\partial H_{jk}^{(\lambda)}/\partial p_\mu$ and $S_{\alpha\beta}^{(\lambda\mu)}$ are $N(N-1) \times \frac{1}{2}N(N-1)$; the $B^{(\lambda)}$ are $\frac{1}{2}N(N-1) \times N(N-1)$; the $\partial H_{jj}^{(0)}/\partial p_\mu$ and the $\epsilon_{\alpha\alpha}^{(\mu)}$ are $N(N-1) \times N$; and the matrix w is $N \times N(N-1)$. To compute v, w, $A^{(\lambda)}$, ρ, ϵ, $\sigma^{(\lambda)}$, etc., is again straightforward, but we do not need them explicitly. What we want to emphasize is that they are either constructed from the components of U or arise from the differentiation of Θ with respect to θ_j and consequently are all independent of the eigenvalues θ_j. Similarly, $S^{(\mu)}$ is independent of θ_j. One more bit of information we need is that $\sigma^{(0)}$ and $\sigma^{(1)}$ are zero matrices, which can easily be verified.

Thus by taking the determinants on both sides of (3.36) and removing the factors $(\theta_\beta - \theta_\alpha)$ we have

$$J(\theta, p) = \prod_{\alpha<\beta} (\theta_\beta - \theta_\alpha)^2 f(p), \tag{3.37}$$

where $f(p)$ is some function of the p_μ.

By inserting (3.37) into (3.3) and integrating over the parameters p_μ we get the j.p.d.f. for the eigenvalues of matrices in the unitary ensemble

$$P(\theta_1, ..., \theta_N) = \exp\left(-a\sum_1^N \theta_j^2 + b\sum_1^N \theta_j + c\right) \prod_{j<k} (\theta_j - \theta_k)^2, \tag{3.38}$$

and, as before, by a proper choice of the origin and the scale of energy we have

$$P_{N2}(x_1, ..., x_N) = C_{N2} \exp\left(-\sum_1^N x_j^2\right) \prod_{j<k} (x_j - x_k)^2. \tag{3.39}$$

We record (3.18), (3.32), and (3.39) as a theorem.

Theorem 3.2. *The joint probability density function for the eigenvalues of matrices from a Gaussian ensemble is given by*

$$P_{N\beta}(x_1, ..., x_N) = C_{N\beta} \exp\left(-\tfrac{1}{2}\beta \sum_1^N x_j^2\right) \prod_{j<k} |x_j - x_k|^\beta, \tag{3.40}$$

where $\beta = 1$ if the ensemble is orthogonal, $\beta = 4$ if it is symplectic, and $\beta = 2$ if it is unitary. The constant $C_{N\beta}$ is chosen in such a way that the $P_{N\beta}$ is normalized to unity:

$$\int_{-\infty}^{\infty} \cdots \int P_{N\beta}(x_1, ..., x_N)\, dx_1 \cdots dx_N = 1. \tag{3.41}$$

In the following chapters [see (5.36), (6.3), and (6.49)] we calculate $C_{N\beta}$ for the physically interesting cases $\beta = 1, 2$, and 4. For these values of β, $C_{N\beta}$ is given by

$$C_{N\beta}^{-1} = (2\pi)^{(1/2)N} \beta^{-(1/2)N-(1/4)\beta N(N-1)} [\Gamma(1 + \tfrac{1}{2}\beta)]^{-N} \prod_{j=1}^{N} \Gamma(1 + \tfrac{1}{2}\beta j). \tag{3.42}$$

It is possible to understand the different powers of β that appear in (3.40) by a simple mathematical argument based on counting dimensions. The dimension of the space T_{1G} is $\frac{1}{2}N(N + 1)$, whereas the dimension of the subspace T'_{1G}, composed of the matrices in T_{1G} with two equal eigenvalues, is $\frac{1}{2}N(N + 1) - 2$. Because of the single restriction, the equality of two eigenvalues, the dimension should normally have decreased by one; as it is decreased by two it indicates a factor in (3.40) linear in $(x_j - x_k)$. Similarly, when $\beta = 2$, the dimension of T_{2G} is N^2, whereas that of T'_{2G} is $N^2 - 3$. When $\beta = 4$, the dimension of T_{4G} is $N(2N - 1)$, whereas that of T'_{4G} is $N(2N - 1) - 5$ (see Appendix A.2).

4 / Gaussian Ensembles

4.1. The Partition Function†

Consider a gas of N point charges with positions $x_1, x_2, \ldots, x_N$ free to move on the infinite straight line $-\infty < x < \infty$. Suppose that the potential energy of the gas is given by

$$W = \frac{1}{2}\sum_i x_i^2 - \sum_{i<j} \ln | x_i - x_j |. \tag{4.1}$$

The first term in W represents a harmonic potential which attracts each charge independently toward the point $x = 0$; the second term represents an electrostatic repulsion between each pair of charges. The logarithmic function comes in if we assume the universe to be two-dimensional. Let this charged gas be in thermodynamical equilibrium at a temperature T, so that the probability density of the positions of the N charges is given by

$$P(x_1, \ldots, x_N) = C \exp\left(\frac{-W}{kT}\right), \tag{4.2}$$

where k is the Boltzmann constant. We immediately recognize that (4.2) is identical to (3.40), provided β is related to the temperature by

$$\beta = (kT)^{-1}. \tag{4.3}$$

This system of point charges in thermodynamical equilibrium is called the Coulomb gas model, corresponding to the Gaussian ensembles.

Following Dyson [1–3], we can define various expressions that relate to our energy-level series in complete analogy with the classical notions of entropy, specific heat, and the like. These expressions,

† Mehta and Dyson [1].

when computed from the observed experimental data and compared with the theoretical predictions, provide a nice method of checking the theory.

In classical mechanics the joint probability density in the velocity space is a product of exponentials

$$\prod_j \exp(-C_j v_j^2)$$

with constant C_j, and its contribution to the thermodynamic quantities of the model are easily calculated. We simply discard these trivial terms. The nontrivial contributions arise from the partition function

$$\Psi_N(\beta) = \int_{-\infty}^{\infty} \cdots \int e^{-\beta W}\, dx_1 \cdots dx_N \tag{4.4}$$

and its derivatives with respect to β. Therefore it is important to derive an analytical expression for $\Psi_N(\beta)$. Unfortunately, this has not yet been done. Lacking this we put forward the conjecture

Conjecture 4.1. *For any positive integer N and real or complex β we have, identically,*

$$\Psi_N(\beta) = (2\pi)^{(1/2)N}\beta^{-(1/2)N-(1/4)\beta N(N-1)}[\Gamma(1 + \tfrac{1}{2}\beta)]^{-N} \prod_{j=1}^{N} \Gamma(1 + \tfrac{1}{2}\beta j). \tag{4.5}$$

The evidence in favor of Conjecture 4.1 is strong. Equation (4.5) can be verified directly for $\beta = 1, 2$, or 4 [cf. Chapters 5 and 6, especially (5.36), (6.3), and (6.49)]. On the other hand, (4.5) has also been verified for general β in the cases $N = 1$, 2, and 3. When $N = 1$ or 2, the calculation is trivial. When $N = 3$, the verification is by no means trivial, but has been carried through first by Gaudin [2]. An independent and much simpler verification of (4.5) for this case, $N = 3$, was later given by C. L. Mehta [1]. We reproduce both these methods in Appendix A.3.

Finally, we can prove that if Conjecture 4.1 is true for $\beta = 2k$, where k is an arbitrary positive integer, it is true for all complex values of β. The argument is based on the fact that the energy W given by (4.1) is bounded from below. More precisely,

$$W \geqslant W_0 = \tfrac{1}{4}N(N-1)(1 + \ln 2) - \frac{1}{2}\sum_{j=1}^{N} j \ln j, \tag{4.6}$$

and this minimum is attained when the positions of the charges coincide with the zeros of the Hermite polynomial $H_N(x)$ (cf. Appendix A.4).

Therefore we can write

$$\Psi_N(\beta) = \int_0^Y P(y)\, y^\beta\, dy, \tag{4.7}$$

where

$$Y = e^{-W_0}, \tag{4.8}$$

and $P(y)$ is a positive weight function. In other words, $\Psi_N(\beta)$ is a moment function defined over a finite interval $(0, Y)$. This function must possess special analytical properties [Shohat and Tamarkin, 1]. It must be analytic in the half-plane (real $\beta > 0$) and in this region must satisfy the inequality

$$|\, \Psi_N(\beta)| < C|\, Y^\beta\,|. \tag{4.9}$$

Now the function

$$\psi_N(\beta) = (2\pi)^{(1/2)N}\beta^{-(1/2)N-(1/4)\beta N(N-1)}[\Gamma(1 + \tfrac{1}{2}\beta)]^{-N} \prod_{j=1}^{N} \Gamma(1 + \tfrac{1}{2}\beta j) \tag{4.10}$$

certainly satisfies these conditions. It has singularities only on the negative real axis and its behavior for large $|\, \beta\,|$ is

$$\begin{aligned}\psi_N(\beta) &\sim (2\pi)^{(1/2)N}\beta^{-(1/2)N-(1/4)\beta N(N-1)}[(\tfrac{1}{2}\beta)^{(1/2)\beta+1/2}\, e^{-(1/2)\beta}\sqrt{2\pi}]^{-N} \\ &\quad \times \prod_{j=1}^{N} [(\tfrac{1}{2}\beta j)^{(1/2)\beta j+1/2}\, e^{-(1/2)\beta j}\sqrt{2\pi}] \\ &\sim Y^\beta(2\pi)^{(1/2)N}\beta^{-(1/2)N}(N!)^{1/2}.\end{aligned} \tag{4.11}$$

The function

$$\Delta_N(\beta) = Y^{-2\beta}[\Psi_N(2\beta) - \psi_N(2\beta)] \tag{4.12}$$

is thus regular and bounded in the half-plane real $\beta > 0$.

At this point we can apply a theorem of Carlson [Tischmarsh, 1].

Carlson's Theorem. *If a function of β is regular and bounded in the half-plane* real $\beta > 0$ *and is zero for* $\beta = 1, 2, 3, \ldots$, *then it is identically zero.*

Applying this theorem to the function $\Delta_N(\beta)$, we deduce that (4.5) holds for all complex values of β if it holds for all even integers

$\beta = 2k$. In other words, Conjecture 4.1 is equivalent to the following, apparently weaker, statement:

$$\int_{-\infty}^{\infty}\cdots\int \exp\left(-k\sum_1^N x_i^2\right)\prod_{i<j}(x_i - x_j)^{2k}\,dx_1\cdots dx_N$$

$$= (2\pi)^{(1/2)N}(2k)^{-(1/2)N-(1/2)kN(N-1)}(k!)^{-N}\prod_{j=1}^{N}(kj)!. \tag{4.13}$$

It is possible to go further and reduce Conjecture 4.1 to a finite algebraic identity. For this purpose we introduce the notation

$$D_j = \left(\frac{\partial}{\partial x_j}\right)_{x_j=0}, \tag{4.14}$$

with the understanding that all differentiations are to be carried out before the variables x_j are set equal to zero. The identity

$$\left(\frac{k}{\pi}\right)^{1/2}\int_{-\infty}^{\infty} e^{-kx_j^2}x_j^n\,dx_j = \left[\frac{i}{(2k)^{1/2}}D_j\right]^n e^{-(1/2)x_j^2}, \tag{4.15}$$

which is easily verified (Appendix A.5), enables us to replace all integrations in (4.13) by differentiations at the point $x_j = 0$. Equation 4.13 then takes the form

$$\left[\prod_{i<j}(D_i - D_j)^{2k}\right]\exp\left(-\frac{1}{2}\sum_1^N x_j^2\right) = (-1)^{(1/2)N(N-1)k}(k!)^{-N}\prod_{j=1}^{N}(kj)!. \tag{4.16}$$

Replacing the exponential with its power-series expansion, we see that all terms $(\frac{1}{2}\sum x_j^2)^l$ give zero on differentiation if $l < \frac{1}{2}N(N-1)k$. On the other hand, if $l > \frac{1}{2}N(N-1)k$, the differentiations will leave a homogeneous polynomial of order $l - \frac{1}{2}N(N-1)k$ in the variables $x_1, x_2, \ldots, x_N$ and, on setting $x_j = 0$; $j = 1, 2, \ldots, N$, once again we obtain no contribution. Only the term with $l = \frac{1}{2}N(N-1)k$ contributes and (4.16) is a finite combinational identity:

$$\left[\prod_{1\le i<j\le N}\left(\frac{\partial}{\partial x_i} - \frac{\partial}{\partial x_j}\right)^{2k}\right]\left[\left(\frac{1}{2}\sum_{j=1}^{N} x_j^2\right)^{(1/2)N(N-1)k}\right]$$

$$= [\tfrac{1}{2}N(N-1)k]!\,(k!)^{-N}\prod_{j=1}^{N}(kj)! \tag{4.17}$$

A somewhat neater form is obtained by interchanging the roles of $\partial/\partial x_j$ and x_j. Thus we obtain a statement that is equivalent to Conjecture 4.1.

Conjecture 4.2. *The identity*

$$\left(\frac{1}{2}\sum_{j=1}^{N}-\frac{\partial^2}{\partial x_j^2}\right)^{(1/2)N(N-1)k}\prod_{1\leqslant i<j\leqslant N}(x_i-x_j)^{2k}$$

$$=(\tfrac{1}{2}N(N-1)k)!\,(k!)^{-N}\prod_{j=1}^{N}(kj)! \tag{4.18}$$

holds for all positive integers N and k.

Once the partition function is known, other thermodynamic quantities such as free energy, energy, entropy, and specific heat can be calculated by elementary differentiation. Because all the known properties are identical to those of the circular ensembles, studied at length in Chapters 8, 9, and 10, we do not insist on this point here.

4.2. The Asymptotic Formula for the Level Density†

Since the expression (3.40) for $P(x_1, \ldots, x_N)$, the probability that the eigenvalues will lie in unit intervals around $x_1, x_2, \ldots, x_N$ is valid for all values of x_i, the density of levels

$$\sigma_N(x)=N\int_{-\infty}^{\infty}\cdots\int P(x_1, x_2, \ldots, x_N)\,dx_2\cdots dx_N \tag{4.19}$$

can be calculated for any N by actual integration [Mehta and Gaudin, 1]. The details of this tedious calculation are not given here, for an expression for $\sigma_N(x)$, derived by a different method, appears in Chapter 5.

However, if one is interested in the limit of large N, as we certainly are, these complications can be avoided by assuming that the corresponding Coulomb gas is a classical fluid with a continuous macroscopic density. More precisely, this amounts to the following two assumptions:

† Wigner [4].

1. The potential energy W given by (4.1) can be approximated by the functional

$$W(\sigma) = \frac{1}{2}\int_{-\infty}^{\infty} dx\, x^2\, \sigma(x) - \frac{1}{2}\iint_{-\infty}^{\infty} dx\, dy\, \sigma(x)\, \sigma(y) \ln |x - y|. \quad (4.20)$$

2. The level density $\sigma(x)$ will be such as to minimize the expression (4.20), consistent with the requirements

$$\int_{-\infty}^{\infty} dx\, \sigma(x) = N \quad (4.21)$$

and

$$\sigma(x) \geqslant 0. \quad (4.22)$$

The first integral in (4.20) reproduces the first sum in (4.1) accurately in the limit of large N. The same is not true of the second integral, for it neglects the two-level correlations, which may be expected to extend over a few neighboring levels; however, because the total number of levels is large their effect may be expected to be small. The factor $\frac{1}{2}$ in the second term of (4.20) comes from the condition $i < j$ in (4.1).

The problem of finding the stationary points of the functional $W(\sigma)$, (4.20), with the restriction (4.21), leads us to the integral equation

$$-\tfrac{1}{2}x^2 + \int_{-\infty}^{\infty} dy\, \sigma(y) \ln |x - y| = C, \quad (4.23)$$

where C is a Lagrange constant. Actually (4.23) has to hold only for those values of x for which $\sigma(x) > 0$. One cannot add a negative increment to $\sigma(x)$, where $\sigma(x) = 0$, and therefore the functional differentiation is not valid; hence (4.23) cannot be derived for such values of x. It is not difficult to solve (4.23) [Mushkhelishvili, 1]. This will not be done here, but the solution will be given and then verified.

Differentiation of (4.23) with respect to x eliminates C. Before carrying it out, we must replace the integral with

$$\lim_{\epsilon \to 0} \left(\int_{-\infty}^{x-\epsilon} dy + \int_{x+\epsilon}^{\infty} dy \right) \sigma(y) \ln |x - y|. \quad (4.24)$$

When (4.24) is differentiated with respect to x, the terms arising from the differentiation of the limits drop out and only the derivative

of $\ln|x-y|$ remains. The integral becomes a principal value integral and (4.23) becomes

$$P\int_{-\infty}^{\infty}\frac{\sigma(y)}{x-y}\,dy = x. \tag{4.25}$$

Conversely, if (4.25) is satisfied by some $\sigma(y)$ and this σ is an even function, then it will satisfy (4.23) also. We try

$$\begin{aligned}\sigma(y) &= C(A^2-y^2)^{1/2}, && |y|<A,\\ &= 0, && |y|>A.\end{aligned} \tag{4.26}$$

Elementary integration gives

$$\int\frac{(A^2-y^2)^{1/2}}{x-y}\,dy = x\sin^{-1}\frac{y}{A} - (A^2-y^2)^{1/2} + (A^2-x^2)^{1/2}\ln\left[\frac{A(x-y)-x(A^2-y^2)^{1/2}-y(A^2-x^2)^{1/2}}{A(x-y)-x(A^2-y^2)^{1/2}+y(A^2-x^2)^{1/2}}\right]. \tag{4.27}$$

Taking the principal value of (4.27) between the limits $(-A, A)$, we find that only the first term gives a nonzero contribution, which is πx. Hence (4.25) gives

$$C = \frac{1}{\pi} \tag{4.28}$$

and (4.21) gives

$$\frac{1}{\pi}\frac{\pi}{2}A^2 = N. \tag{4.29}$$

Thus

$$\sigma(x) = \begin{cases}\dfrac{1}{\pi}(2N-x^2)^{1/2}, & |x|<(2N)^{1/2},\\ 0, & |x|>(2N)^{1/2}.\end{cases} \tag{4.30}$$

This is the so-called "semicircle law" first derived by Wigner.

Actually the two-level correlation function can be calculated (cf. Chapters 5 and 6) and the above intuitive arguments put to test. Instead, we shall derive an exact expression for the level-density valid for any N (cf. Section 5.4.1). The limit $N\to\infty$ can then be taken to obtain the "semicircle law."

We have noted in Section 4.1 that without any approximation whatever the energy W attains its minimum value when the points $x_1, x_2, \ldots, x_N$ are the zeros of the Nth-order Hermite polynomial. The postulate of classical statistical mechanics then implies that in the limit of very large N the level density is the same as the density of zeros of the Nth-order Hermite polynomial. This later problem has been investigated by many authors, and we may conveniently refer to the relevant mathematical literature [Szegö, 1].

In Appendix A.29 we present an argument, due essentially to Wigner, which shows that the eigenvalue density obeys the semicircle law for a much wider class of matrices than those considered here. For the spacing distribution no such argument has yet been found.

5 / The Gaussian Orthogonal Ensemble

5.1. General Remarks

In this chapter we shall study the statistical properties of N points distributed on an infinite straight line with the joint probability density function

$$P(x_1, ..., x_n) = C_{N1} \exp\left(-\frac{1}{2}\sum_1^N x_j^2\right) \prod_{i<j} |x_i - x_j|. \tag{5.1}$$

In Chapter 3 it was shown that the probability density function (5.1) holds for the eigenvalues of a real symmetric matrix chosen at random from the Gaussian orthogonal ensemble. It was suggested that if the zero and the unit of the energy scale are properly chosen the series of points x_1, x_2, ..., x_N should provide a good model for the statistical behavior of the fluctuations of energy levels of a sufficiently complicated system.

The main objective of the analysis is to calculate the *n-level correlation function* [Dyson, 3]

$$R_n(x_1, ..., x_N) = \frac{N!}{(N-n)!} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} P(x_1, ..., x_N)\, dx_{n+1} \cdots dx_N, \tag{5.2}$$

which measures the probability of finding a level (regardless of labeling) in each of the unit intervals around the points x_1, x_2, ..., x_n, the positions of the remaining levels being unobserved. In particular, $R_1(x)$ will give the over-all level density. Each function R_n for $n > 1$ contains terms of various kinds describing the grouping of n levels into various subgroups or clusters. For practical purposes it is convenient to work with the *n-level cluster function* defined by

$$T_n(x_1, ..., x_n) = \sum_G (-1)^{m-n}(m-1)! \prod_{j=1}^{m} R_{G_j}(x_k, \text{ with } k \text{ in } G_j). \tag{5.3}$$

Here G stands for any division of the indices $(1, 2, \ldots, n)$ into subgroups $(G_1, G_2, \ldots, G_m)$. For example,

$$\begin{aligned}
T_1(x) &= R_1(x), \\
T_2(x_1, x_2) &= -R_2(x_1, x_2) + R_1(x_1) R_1(x_2), \\
T_3(x_1, x_2, x_3) &= R_3(x_1, x_2, x_3) - R_1(x_1) R_2(x_2, x_3) - \cdots - \cdots + \\
&\quad + 2R_1(x_1) R_1(x_2) R_1(x_3). \\
T_4(x_1, x_2, x_3, x_4) &= -R_4(x_1, x_2, x_3, x_4) + \\
&\quad + [R_1(x_1) R_3(x_2, x_3, x_4) + \cdots + \cdots + \cdots] \\
&\quad + [R_2(x_1, x_2) R_2(x_3, x_4) + \cdots + \cdots] \\
&\quad - 2[R_2(x_1, x_2) R_1(x_3) R_1(x_4) + \cdots + \cdots + \cdots + \cdots + \cdots] \\
&\quad + 6R_1(x_1) R_1(x_2) R_1(x_3) R_1(x_4),
\end{aligned}$$

where in the last equation the first bracket contains four terms, the second contains three terms, and the third contains six terms. Equation 5.3 is a finite sum of products of the R functions, the first term in the sum being $(-1)^n R_n(x_1, \ldots, x_n)$ and the last being $(n-1)!\, R_1(x_1) \ldots R_1(x_n)$.

We would be particularly pleased if these functions R_n and T_n turn out to be functions only of the differences $|x_i - x_j|$. Unfortunately, this is not true in general. Even the level density R_1 will turn out to be a semicircle rather than a constant or an exponential (cf. Section 4.2). However, as long as we remain in the region of constant density, we can see that the functions R_n and T_n satisfy this requirement.

It was this unsatisfactory feature of the Gaussian ensembles that led Dyson [1] to define the circular ensembles discussed in Chapters 8 to 11.

The inverse of (5.3) (cf. Appendix A.6) is

$$R_n(x_1, \ldots, x_n) = \sum_G (-1)^{n-m} \prod_{j=1}^{m} T_{G_j}(x_k, \text{ with } k \text{ in } G_j). \tag{5.4}$$

Thus each set of functions R_n and T_n is easily determined in terms of the other. The advantage of the cluster functions is that they have the property of vanishing when any one (or several) of the separations $|x_i - x_j|$ becomes large in comparison with the local mean level spacing. The function T_n describes the correlation properties of a single cluster of n levels, isolated from the more trivial effects of lower order correlations.

Of special interest for comparison with experiment are those features of the statistical model that tend to definite limits as $N \to \infty$. The cluster functions are convenient also from this point of view. While taking the limit $N \to \infty$, we must measure the energies in units of the mean level spacing D and introduce the variables

$$\xi_j = \frac{x_j}{D}. \tag{5.5}$$

The ξ_j then form a statistical model for an infinite series of energy levels with mean spacing $D = 1$. The cluster functions

$$Y_n(\xi_1, \xi_2, \ldots, \xi_n) = \lim_{N\to\infty} D^n T_n(x_1, x_2, \ldots, x_n) \tag{5.6}$$

are well defined and finite everywhere. In particular,

$$Y_1(\xi) = 1, \tag{5.7}$$

whereas $Y_2(\xi, \eta)$ defines the shape of the neutralizing charge cloud induced by each particle around itself when the model is interpreted as a classical Coulomb gas (see Chapter 4).

In this chapter it is shown that all cluster functions are calculable in principle. The Y_n are exhibited as coefficients in the expansion of a certain determinant. However, the elementary algebra that is required for the extraction of higher Y_n is tedious. Explicit evaluations will be made only for the one- and two-level functions $T_1(x)$ and $T_2(x, y)$; when x and y are in the region of maximum and constant density, the limit $Y_2(\xi, \eta)$ is given, it being a function only of the difference $|\xi - \eta|$. *The two-level form factor* defined as the Fourier transform of Y_2

$$b(k) = \int_{-\infty}^{\infty} Y_2(r)\, e^{2\pi i k r}\, dr \tag{5.8}$$

is also given, for many important properties of the level distribution, such as mean square values, depend only on it.

5.2. The Method of Integration over Alternate Variables†

In any exact calculation with the function (5.1) the first serious difficulty is its unfavorable symmetry caused by the presence of the

† Mehta [1], Bruijn [1].

absolute-value sign. This difficulty can be overcome by integrating over alternate variables. Because the method is frequently applied in the discussion that follows, a somewhat detailed description is given here.

Let $u(x)$ and $v(x)$ be any two functions defined for the real values of x. Consider the average value

$$\rho_N = \left\langle \prod_{\text{alt}} u(x_i) \prod_{\text{alt}}' v(x_j) \right\rangle \tag{5.9}$$

taken with respect to the probability density (5.1). Here $\prod_{\text{alt}}$ means a product taken over a set of $(\frac{1}{2}N)$ alternate points x_i as they lie on the real axis, and $\prod'_{\text{alt}}$ denotes the product taken over the set of remaining alternate points x_j. From the symmetry of (5.1) we may write

$$\rho_N = \tfrac{1}{2}[\rho_N(u, v) + \rho_N(v, u)] \tag{5.10}$$

with

$$\begin{aligned}\rho_N(u, v) &= N! \int_{R(-\infty, x_1, \ldots, x_N, \infty)} \cdots \int \left[\prod_i u(x_{2i-1})\, v(x_{2i})\right] P(x_1, \ldots, x_N)\, dx_1 \cdots dx_N \\ &= C_{N1} N! \int_R \cdots \int \left[\prod_i u(x_{2i-1})\, v(x_{2i})\right] \exp\left(-\frac{1}{2}\sum_i x_i^2\right) \\ &\quad \cdot \prod_{i>j} (x_i - x_j)\, dx_1 \cdots dx_N ,\end{aligned} \tag{5.11}$$

where the region of integration R is $-\infty < x_1 \leqslant \cdots \leqslant x_N < \infty$. Note that the absolute value sign in (5.11) is no longer required.

It is convenient to write the integrand in (5.11) as a determinant containing oscillator wave functions. To do this we first write the product $\prod_{i>j} (x_i - x_j)$ as a Vandermonde determinant, its jth row being

$$x_1^{j-1}, x_2^{j-1}, \ldots, x_N^{j-1}, \tag{5.12}$$

j varying from 1 to N. Multiplying the jth row by 2^{j-1} and adding to it an appropriate linear combination of the other rows with lower powers of the variables, we may replace this jth row by

$$H_{j-1}(x_1), H_{j-1}(x_2), \ldots, H_{j-1}(x_N), \tag{5.13}$$

where H_j is the Hermite polynomial of order j. Now we may take the exponentials inside and multiply the jth row with the normalization factor $(2^{j-1}(j-1)!\sqrt{\pi})^{-1/2}$, changing this row in turn to

$$\varphi_{j-1}(x_1), \varphi_{j-1}(x_2), \ldots, \varphi_{j-1}(x_N), \tag{5.14}$$

where the φ_j are the normalized oscillator wave functions

$$\begin{aligned}\varphi_j(x) &= (2^j j!\sqrt{\pi})^{-1/2} e^{-(1/2)x^2} H_j(x) \\ &= (2^j j!\sqrt{\pi})^{-1/2} e^{(1/2)x^2}\left(-\frac{d}{dx}\right)^j e^{-x^2}.\end{aligned} \tag{5.15}$$

In each of these manipulations the value of the determinant is changed only by a multiplicative factor which can be taken out. Thus (5.11) reads

$$\begin{aligned}\rho_N(u, v) = C_{N1} N! \prod_{j=0}^{N-1} (2^{-j} j!\sqrt{\pi})^{1/2} \int_R \cdots \int dx_1 \cdots dx_N \\ \left[\prod_i u(x_{2i-1})\, v(x_{2i})\right] \det[\varphi_{i-1}(x_j)]_{i,j=1,\ldots,N},\end{aligned} \tag{5.16}$$

where, written in full,

$$\det[\varphi_{i-1}(x_j)]_{i,j=1,\ldots,N} \equiv \begin{vmatrix} \varphi_0(x_1) & \cdots & \varphi_0(x_N) \\ \varphi_1(x_1) & \cdots & \varphi_1(x_N) \\ \cdots & \cdots & \cdots \\ \varphi_{N-1}(x_1) & \cdots & \varphi_{N-1}(x_N) \end{vmatrix}. \tag{5.17}$$

As we have already emphasized, the difficulty in the evaluation of (5.16) is that the integrand is not symmetric. However, if the integration over, say, $x_1, x_3, \ldots$, that is, variables with odd indices, is carried out, the remaining integrand becomes a symmetric function of the variables $x_2, x_4, \ldots$.

Integration over x_1 replaces the first column in the determinant by the column $[F_0(x_2), F_1(x_2), \ldots, F_{N-1}(x_2)]$ where the functions $F_j(x)$ are defined by

$$F_j(x) = \int_{-\infty}^{x} u(y)\, \varphi_j(y)\, dy. \tag{5.18}$$

Now integration over x_3 replaces the third column by the column

$$[F_0(x_4) - F_0(x_2), F_1(x_4) - F_1(x_2), \ldots, F_{N-1}(x_4) - F_{N-1}(x_2)]$$

but because we already have a column of $F_j(x_2)$ functions we may drop the terms with negative signs. Thus at each integration over $x_3, x_5, \ldots$ the corresponding columns are replaced by columns of single F_j functions. In the case of odd N the last column is replaced by pure numbers $F_j(\infty)$, and we have

$$\rho_N(u, v) = C_{N1} N! 2^{-(1/4)N(N-1)} \pi^{N/4} \prod_{j=0}^{N-1} (j!)^{1/2} \cdot \int\cdots\int_{R(-\infty, x_2, x_4, \ldots, \infty)} dx_2\, dx_4 \cdots \left[\prod_i v(x_{2i})\right] \det[F_{i-1}(x_{2j}), \varphi_{i-1}(x_{2j})], \quad i = 1, \ldots, N, \quad j = 1, \ldots, \left[\frac{N}{2}\right]. \tag{5.19}$$

As the integrand in (5.19) is now symmetric in the remaining variables, we can integrate separately and independently over them on the whole interval $-\infty, \infty$ and divide by $[(1/2)N]!$ where $[(1/2)N]$ is the largest integer less than or equal to $(1/2)N$.

To avoid minor complications we take N even, $N = 2m$. The symmetry of the integrand and the independence of the variables in (5.19) allow us to express it as a pfaffian (cf. Appendix A.7):

$$\rho_N(u, v) = C_{N1} N!\, 2^{-(1/4)N(N-1)} \pi^{(1/4)N} \prod_{j=0}^{N-1} (j!)^{1/2} \prod_{j=0}^{m-1} \left(\frac{8}{2j+1}\right)^{1/2} (\det[f_{ij}]_{i,j=0,1,\ldots,2m-1})^{1/2}, \tag{5.20}$$

where the numbers f_{ij} are given by

$$f_{2i,2j} = \iint dy\, dx\, u(y)\, v(x)[\varphi_{2i}(y)\, \varphi_{2j}(x) - \varphi_{2i}(x)\, \varphi_{2j}(y)], \tag{5.21}$$

$$f_{2i,2j+1} = -f_{2j+1,2i} = \left(\frac{2j+1}{8}\right)^{1/2} \iint dy\, dx\, u(y)\, v(x)[\varphi_{2i}(y)\, \varphi_{2j+1}(x) - \varphi_{2i}(x)\, \varphi_{2j+1}(y)] \tag{5.22}$$

and

$$f_{2i+1,2j+1} = \left(\frac{2i+1}{8}\right)^{1/2} \left(\frac{2j+1}{8}\right)^{1/2} \iint dy\, dx\, u(y)\, v(x) [\varphi_{2i+1}(y)\, \varphi_{2j+1}(x) - \varphi_{2i+1}(x)\, \varphi_{2j+1}(y)], \tag{5.23}$$

the range of integrations in these equations being $-\infty < y \leqslant x < \infty$. From the relation

$$\sqrt{2}\,\varphi_n'(x) = \sqrt{n}\,\varphi_{n-1}(x) - \sqrt{n+1}\,\varphi_{n+1}(x) \tag{5.24}$$

we derive

$$\begin{aligned} f_{2i,2j+1} - \left(\frac{2j}{2j-1}\right)^{1/2} f_{2i,2j-1} &= -\frac{1}{2} \iint\limits_{-\infty<y\leqslant x<\infty} dy\,dx\,u(y)\,v(x)[\varphi_{2i}(y)\,\varphi_{2j}'(x) - \varphi_{2i}(x)\,\varphi_{2j}'(y)] \\ &= g_{ij}\,, \text{ say,} \end{aligned} \tag{5.25}$$

and

$$\begin{aligned} f_{2i+1,2j+1} - \left(\frac{2i}{2i-1}\right)^{1/2} f_{2i-1,2j+1} - \left(\frac{2j}{2j-1}\right)^{1/2} f_{2i+1,2j-1} \\ + \left(\frac{2i}{2i-1}\right)^{1/2}\left(\frac{2j}{2j-1}\right)^{1/2} f_{2i-1,2j-1} \\ = \frac{1}{4} \iint\limits_{-\infty<y\leqslant x<\infty} dy\,dx\,u(y)\,v(x)[\varphi_{2i}'(y)\,\varphi_{2j}'(x) - \varphi_{2i}'(x)\,\varphi_{2j}'(y)] \\ = \mu_{ij}\,, \text{ say.} \end{aligned} \tag{5.26}$$

On the other hand, we have

$$\begin{aligned} \pi^{(1/4)N} \prod_{j=0}^{N-1} (j!)^{1/2} &= \prod_{j=0}^{m-1} [\pi(2j)!\,(2j+1)!]^{1/2} \\ &= \prod_{j=0}^{m-1} [\pi^{1/2}(2j+1)^{1/2}\,\Gamma(2j+1)] \\ &= \prod_{j=0}^{m-1} [2^{2j}\Gamma(j+\tfrac{1}{2})\,\Gamma(j+1)(2j+1)^{1/2}] \\ &= 2^{m(m-1)} \prod_{j=0}^{m-1} (2j+1)^{1/2} \prod_{g=1}^{2m} \left[\frac{2}{g}\,\Gamma(1+\frac{1}{2}g)\right]. \end{aligned} \tag{5.27}$$

Now, to reduce the determinant in (5.20), we replace its $(2j+2)$-th column with

$$(2j+2)\text{ th column} - \left(\frac{2j}{2j-1}\right)^{1/2} (2j)\text{ th column},$$

for $j = m - 1, m - 2, \ldots, 0$, in turn. Next, we repeat the same procedure for the rows. Thus, using (5.25), (5.26), and (5.27), we get

$$\rho_N^2(u, v) = \left[C_{N1} 2^{(3/2)N} \prod_{j=1}^{N} \Gamma(1 + \tfrac{1}{2}j)\right]^2 \det \begin{bmatrix} \lambda_{ij} & g_{ij} \\ -g_{ji} & \mu_{ij} \end{bmatrix}_{i,j=0,1,\ldots,m-1}, \tag{5.28}$$

where

$$\lambda_{ij} = f_{2i,2j} \tag{5.29}$$

and g_{ij} and μ_{ij} are defined by (5.25) and (5.26), respectively.

Equation 5.28 is a general result. If, moreover,

$$u(-x)\, v(-y) = u(x)\, v(y), \tag{5.30}$$

there are further simplifications, for we then have

$$\lambda_{ij} = \mu_{ij} = 0, \tag{5.31}$$

and the pfaffian form in (5.28) reduces to a determinant:

$$\rho_N(u, v) = C_{N1} 2^{(3/2)N} \left[\prod_{j=1}^{N} \Gamma(1 + \tfrac{1}{2}j)\right] \det[g_{ij}]_{i,j=0,1,\ldots,m-1}, \tag{5.32}$$

with

$$g_{ij} = \iint\limits_{-\infty < y < x < \infty} dy\, dx\, u(y)\, v(x)\, \varphi_{2i}(x)\, \varphi'_{2j}(y). \tag{5.33}$$

5.3. The Normalization Constant

Let

$$u(x) = v(x) = 1 \tag{5.34}$$

in (5.9). Equation 5.30 is now satisfied, and (5.33) gives

$$g_{ij} = \delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases} \tag{5.35}$$

Consequently, from (5.32) we get

$$C_{N1}^{-1} = 2^{(3/2)N} \prod_{j=1}^{N} \Gamma(1 + \tfrac{1}{2}j). \tag{5.36}$$

This verifies (4.5) for the case $\beta = 1$.

Substituting the value of C_{N1} from (5.36) in (5.28), we get

$$\rho_N{}^2(u, v) = \det \begin{bmatrix} \lambda_{ij} & g_{ij} \\ -g_{ji} & \mu_{ij} \end{bmatrix}_{i,j=0,1,\ldots,m-1}, \tag{5.37}$$

and if, in addition, (5.30) is satisfied we get from (5.32)

$$\rho_N(u, v) = \det[g_{ij}]_{i,j=0,1,\ldots,m-1}. \tag{5.38}$$

Instead of calculating two pfaffians $\rho_N(u, v)$ and $\rho_N(v, u)$ and then taking their arithmetic mean, it is much better to have a single pfaffian from the very beginning. This is achieved by the following lemma.

Lemma 5.1. *Let the function $f(x_1, \ldots, x_{2m})$ be antisymmetric under any of the interchanges $x_{2j-1} \leftrightarrow x_{2j}$, $j = 1, \ldots, m$. Then*

$$\int\cdots\int_{R(x_1;\, x_2;\, \ldots;\, x_{2m})} f(x_1, \ldots, x_{2m})\, dx_1 \cdots dx_{2m}$$

$$= \int\cdots\int_{R'(x_1,x_2;\, x_3,x_4;\, \ldots;\, x_{2m-1},x_{2m})} f(x_1, \ldots, x_{2m}) \prod_{j=1}^{m} [\tfrac{1}{2}\epsilon(x_{2j} - x_{2j-1})]\, dx_1 \cdots dx_{2m}, \tag{5.39}$$

where the region of integration R on the left is

$$a \leqslant x_1 \leqslant x_2 \leqslant \cdots \leqslant x_{2m} \leqslant b,$$

whereas the region R' on the right is $a \leqslant (x_1, x_2) \leqslant (x_3, x_4) \leqslant \cdots \leqslant b$. The sign function $\epsilon(x)$ appearing in the right-hand integrand is

$$\epsilon(x) = \begin{cases} 1, & \text{if } x > 0, \\ 0, & \text{if } x = 0, \\ -1, & \text{if } x < 0. \end{cases} \tag{5.40}$$

Let the expression on the left in (5.39) be denoted by I. To prove the validity of (5.39), we interchange the variables x_1 and x_2 in I. Taking account of the antisymmetry of f under this interchange, we get

$$I = -\int\cdots\int_{R(x_2;\, x_1;\, x_3;\, \ldots;\, x_{2m})} dx_1 \cdots dx_{2m} f(x_1, \ldots, x_{2m}).$$

Taking the arithmetic mean of both sides of this equation and (5.39), we obtain

$$I = \int\cdots\int_{a \leqslant (x_1, x_2) \leqslant x_3 \leqslant \cdots \leqslant x_{2m} \leqslant b} dx_1 \cdots dx_{2m} f(x_1, \ldots, x_{2m})[\tfrac{1}{2}\epsilon(x_2 - x_1)].$$

A repetition of this procedure for the variables x_3, x_4 will give

$$I = \int\cdots\int_{a \leqslant (x_1, x_2) \leqslant (x_3, x_4) \leqslant x_5 \leqslant \cdots \leqslant b} dx_1 \cdots dx_{2m} f(x_1, \ldots, x_{2m}) \prod_{j=1}^{2} [\tfrac{1}{2}\epsilon(x_{2j} - x_{2j-1})].$$

Proof is now obvious.

As a particular case of (5.39), we write

$$\begin{aligned}
\rho_N &= \int\cdots\int_{-\infty < x_1 \leqslant x_2 \leqslant \cdots \leqslant x_{2m} < \infty} dx_1 \cdots dx_{2m}\, P_{N1}(x_1, \ldots, x_{2m}) \\
&\quad \times \frac{1}{2}\Big[\prod_{j=1}^{m} u(x_{2j-1})\, v(x_{2j}) + \prod_{j=1}^{m} u(x_{2j})\, v(x_{2j-1})\Big] \\
&= \int\cdots\int_{\substack{-\infty < (x_1, x_2) \leqslant \cdots \\ \leqslant (x_{2m-1}, x_{2m}) < \infty}} dx_1 \cdots dx_{2m}\, P_{N1}(x_1, \ldots, x_{2m}) \\
&\quad \times \frac{1}{2}\Big\{\prod_{j=1}^{m} [\tfrac{1}{2}\epsilon(x_{2j} - x_{2j-1})\, u(x_{2j-1})\, v(x_{2j})] \\
&\quad + \prod_{j=1}^{m} [\tfrac{1}{2}\epsilon(x_{2j} - x_{2j-1})\, u(x_{2j})\, v(x_{2j-1})]\Big\} \\
&= \int\cdots\int_{\substack{-\infty < (x_1, x_2) < \cdots \\ < (x_{2m-1}, x_{2m}) < \infty}} dx_1 \cdots dx_{2m}\, P_{N1}(x_1, \ldots, x_{2m}) \\
&\quad \times \prod_{j=1}^{m} [\tfrac{1}{2}\epsilon(x_{2j} - x_{2j-1})\, u(x_{2j-1})\, v(x_{2j})].
\end{aligned}$$

Now all the previous considerations in this section can be applied to this form of ρ_N . Thus $\rho_N{}^2$ is given by (5.37):

$$\rho_N{}^2 = \det \begin{bmatrix} \lambda_{ij} & g_{ij} \\ -g_{ji} & \mu_{ij} \end{bmatrix}_{i,j=0,1,\ldots,m-1}, \tag{5.41}$$

where now

$$\lambda_{ij} = \frac{1}{2} \int\int_{-\infty}^{\infty} dy\, dx\, u(y)\, v(x)\, \epsilon(x-y)[\varphi_{2i}(y)\, \varphi_{2j}(x) - \varphi_{2i}(x)\, \varphi_{2j}(y)], \tag{5.42}$$

$$g_{ij} = -\frac{1}{4} \int\int_{-\infty}^{\infty} dy\, dx\, u(y)\, v(x)\, \epsilon(x-y)[\varphi_{2i}(y)\, \varphi'_{2j}(x) - \varphi_{2i}(x)\, \varphi'_{2j}(y)], \tag{5.43}$$

and

$$\mu_{ij} = \frac{1}{8} \int\int_{-\infty}^{\infty} dy\, dx\, u(y)\, v(x)\, \epsilon(x-y)[\varphi'_{2i}(y)\, \varphi'_{2j}(x) - \varphi'_{2i}(x)\, \varphi'_{2j}(y)]. \tag{5.44}$$

If the functions $u(x)$, $v(x)$ satisfy (5.30), then (5.31) holds and

$$\rho_N = \det[g_{ij}]_{i,j=0,1,\ldots,m-1}\,; \tag{5.45}$$

where g_{ij} is given by (5.43).

5.4. One- and Two-Level Correlation Functions

If we can evaluate ρ_N in a closed form while the functions u and v remain arbitrary (and this is essential here), we will obtain all the correlation functions by functional differentiations. However, as Dyson [3] pointed out, we do not need to know ρ_N in any greater detail then its power-series expansion when u and v are in the neighborhood of unity. So let us write $u(x) = v(x) = 1 + a(x)$ in (5.9).

Equation 5.37 then becomes

$$\rho_N{}^2(1+a, 1+a) = \det \begin{bmatrix} \lambda_{ij} & g_{ij} \\ -g_{ji} & \mu_{ij} \end{bmatrix}_{i,j=0,1,\ldots,m-1}. \tag{5.46}$$

From (5.25), (5.26), and (5.29), we have

$$g_{ij} \equiv \delta_{ij} + \nu_{ij}$$

$$= \delta_{ij} - \frac{1}{4}\int_{-\infty}^{\infty}\int dy\, dx\, A(x,y)[\varphi_{2i}(y)\,\varphi'_{2j}(x) - \varphi_{2i}(x)\,\varphi'_{2j}(y)], \quad (5.47)$$

$$\lambda_{ij} = \frac{1}{2}\int_{-\infty}^{\infty}\int dy\, dx\, A(x,y)[\varphi_{2i}(y)\,\varphi_{2j}(x) - \varphi_{2i}(x)\,\varphi_{2j}(y)], \quad (5.48)$$

and

$$\mu_{ij} = \frac{1}{8}\int_{-\infty}^{\infty}\int dy\, dx\, A(x,y)[\varphi'_{2i}(y)\,\varphi'_{2j}(x) - \varphi'_{2i}(x)\,\varphi'_{2j}(y)], \quad (5.49)$$

where

$$A(x,y) = \epsilon(x-y)[a(x) + a(y) + a(x)\,a(y)] \quad (5.50)$$

and δ_{ij} is the Kronecker delta:

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

All the cluster functions $T_n(x_1,\ldots,x_n)$ can, in principle, be determined by expanding the two sides of the identity (5.46) in powers of $a(x)$. According to (5.2), (5.4), (5.9), and (5.46) (cf. Appendix A.6),

$$\rho_N = \sum_{n=0}^{\infty}\frac{1}{n!}\int_{-\infty}^{\infty}\cdots\int R_n(x_1,\ldots,x_n)\prod_1^n a(x_i)\,dx_i \quad (5.51)$$

$$\rho_N = \exp\Big[\sum_{n=1}^{\infty}\frac{(-1)^n}{n!}\int_{-\infty}^{\infty}\cdots\int T_n(x_1,\ldots,x_n)\prod_1^n a(x_i)\,dx_i\Big]. \quad (5.52)$$

The pfaffian for ${\rho_N}^2$ in (5.46) can be expanded (cf. Appendix A.7). The result is a series beginning with the terms

$$\rho_N = 1 + \sum_i \nu_{ii} + \frac{1}{2}\sum_{i,j}(\nu_{ii}\nu_{jj} - \nu_{ij}\nu_{ji} + \lambda_{ij}\mu_{ji})$$

$$+ \frac{1}{6}\sum_{i,j,k}[\nu_{ii}\nu_{jj}\nu_{kk} - 3\nu_{ii}(\nu_{jk}\nu_{kj} - \lambda_{jk}\mu_{jk}) + 2\nu_{ij}\nu_{jk}\nu_{ki} - 6\nu_{ij}\lambda_{jk}\mu_{ki}] + \cdots, \quad (5.53)$$

the remaining terms being of the order a^3 and higher. Because $a(x)$ is arbitrary, each T_n can be picked out as the coefficient of $a(x_1) \cdots a(x_n)$ in the logarithm of the series (5.53),

$$\ln \rho_n = \sum_i \nu_{ii} - \frac{1}{2} \sum_{i,j} (\nu_{ij}\nu_{ji} - \lambda_{ij}\mu_{ji}) + \frac{1}{3} \sum_{i,j,k} (\nu_{ij}\nu_{jk}\nu_{ki} - 3\nu_{ij}\lambda_{jk}\mu_{ki}) + \cdots. \tag{5.54}$$

The summation over the indices i, j, k runs from 0 to $m - 1$.

5.4.1. Level Density

Taking $n = 1$ this procedure gives the level density

$$\sigma_{2m}(x) = \left(\frac{\delta}{\delta a(x)} \sum_{i=0}^{m-1} \nu_{ii}\right)_{a=0} = \sum_{i=0}^{m-1} \left[\varphi_{2i}^2(x) - \varphi_{2i}'(x) \int_0^x \varphi_{2i}(y)\, dy\right], \tag{5.55}$$

or, using (5.24), we get

$$\sigma_{2m}(x) = \sum_{i=0}^{2m-1} \varphi_i^2(x) + \sqrt{m}\, \varphi_{2m-1}(x) \int_0^x \varphi_{2m}(y)\, dy. \tag{5.56}$$

Equation 5.56 was previously obtained [Mehta and Gaudin, 1] by straightforward integrations.

In the limit $m \to \infty$ the summation in (5.56) tends to the "semicircle law" [Wigner, 4]

$$\sigma_{2m}(x) \to \sigma(x) = \begin{cases} \frac{1}{\pi} (4m - x^2)^{1/2}, & |x| < (4m)^{1/2}, \\ 0, & |x| > (4m)^{1/2}, \end{cases} \tag{5.57}$$

whereas the remaining isolated term becomes negligible (cf. Appendix A.8).

5.4.2. Two-Level Functions

For $n = 2$, the theory of Section 5.4 and (5.54) gives

$$T_2(x, y) = -\left[\frac{\delta^2}{\delta a(x)\, \delta a(y)} \ln \rho_N\right]_{a=0}$$

$$= -\tfrac{1}{2}\epsilon(x-y)\sum_i [\varphi_{2i}(x)\,\varphi'_{2i}(y) - \varphi_{2i}(y)\,\varphi'_{2i}(x)]$$

$$- \Big\{\sum_i [\varphi_{2i}(x)\,\varphi'_{2i}(y) - \varphi_{2i}(y)\,\varphi'_{2i}(x)]\Big\}$$

$$\times \Big\{\sum_i \Big[\varphi_{2i}(x)\int_0^y \varphi_{2i}(z)\,dz - \varphi_{2i}(y)\int_0^x \varphi_{2i}(z)\,dz\Big]\Big\}$$

$$+ \Big\{\sum_i \Big[\varphi_{2i}(x)\,\varphi_{2i}(y) - \varphi'_{2i}(x)\int_0^y \varphi_{2i}(z)\,dz\Big]\Big\}$$

$$\times \Big\{\sum_i \Big[\varphi_{2i}(x)\,\varphi_{2i}(y) - \varphi'_{2i}(y)\int_0^x \varphi_{2i}(z)\,dz\Big]\Big\} \tag{5.58}$$

where all sums in (5.58) run from $i = 0$ to $i = m - 1$ and $\epsilon(x)$ is defined by (5.40). By using (5.24) we can transform the last term in (5.58) to a more suitable form:

$$\Big[\sum_{i=0}^{2m-1} \varphi_i(x)\,\varphi_i(y) + \sqrt{m}\,\varphi_{2m-1}(x)\int_0^y \varphi_{2m}(z)\,dz\Big]$$

$$\times \Big[\sum_{i=0}^{2m-1} \varphi_i(x)\,\varphi_i(y) + \sqrt{m}\,\varphi_{2m-1}(y)\int_0^x \varphi_{2m}(z)\,dz\Big]. \tag{5.59}$$

Equation 5.58 gives T_2 for any arbitrary $N = 2m$, x, and y. Let us now restrict x, y to a region of highest, hence constant, level density and take the limit as $m \to \infty$; in other words, if

$$\frac{2\sqrt{m}}{\pi}\,x = \xi, \qquad \frac{2\sqrt{m}}{\pi}\,y = \eta, \tag{5.60}$$

then, as $m \to \infty$, we want ξ and η to remain finite. With this restriction (cf. Appendix A.9), we get

$$\lim_{m\to\infty} \sum_{i=0}^{2m-1} \varphi_i(x)\,\varphi_i(y) = \frac{2\sqrt{m}}{\pi}\,s(r), \tag{5.61}$$

$$\lim_{m\to\infty} \sum_{i=0}^{m-1} [\varphi_{2i}(x)\,\varphi'_{2i}(y) - \varphi_{2i}(y)\,\varphi'_{2i}(x)]\,\epsilon(x-y) = \Big[\frac{2\sqrt{m}}{\pi}\Big]^2 \frac{d}{dr}\,s(r), \tag{5.62}$$

and

$$\lim_{m\to\infty} \sum_{i=0}^{m-1} \Big[\varphi_{2i}(x)\int_0^y \varphi_{2i}(z)\,dz - \varphi_{2i}(y)\int_0^x \varphi_{2i}(z)\,dz\Big]\,\epsilon(x-y) = -\int_0^r s(z)\,dz, \tag{5.63}$$

where

$$s(r) = \frac{\sin \pi r}{\pi r}, \tag{5.64}$$

$$r = |\xi - \eta| = \frac{2\sqrt{m}}{\pi} |x - y|. \tag{5.65}$$

Thus in the limit $m \to \infty$ we get

$$\begin{aligned} Y_2(\xi, \eta) &= \left[\frac{1}{2} - \int_0^r s(z)\, dz\right]\left[\frac{d}{dr} s(r)\right] + [s(r)]^2 \\ &= \left[\int_r^\infty s(z)\, dz\right]\left[\frac{d}{dr} s(r)\right] + [s(r)]^2. \end{aligned} \tag{5.66}$$

The behavior of $Y_2(r)$ for small and large r is given by

$$Y_2(r) = 1 - \tfrac{1}{6}\pi^2 r + \tfrac{1}{60}\pi^4 r^3 - \tfrac{1}{1\,3\,5}\pi^4 r^4 + \cdots \tag{5.67}$$

and

$$Y_2(r) = \frac{1}{\pi^2 r^2} - \frac{1 + \cos^2 \pi r}{\pi^4 r^4} + \cdots, \tag{5.68}$$

respectively. The Fourier transform of Y_2 yields the two-level form factor according to (5.8) (cf. Appendix A.10):

$$\begin{aligned} b(k) &= 1 - 2|k| + |k| \ln(1 + 2|k|), \qquad |k| \leqslant 1, \\ &= -1 + |k| \ln\left(\frac{2|k| + 1}{2|k| - 1}\right), \qquad |k| \geqslant 1. \end{aligned} \tag{5.69}$$

This has the behavior

$$b(k) = 1 - 2|k| + 2k^2 + \cdots, \tag{5.70}$$

$$b(k) = \frac{1}{12k^2} + \frac{1}{80k^4} + \cdots, \tag{5.71}$$

for small and large $|k|$, respectively. At $k = \pm 1$, where the analytic form of $b(k)$ changes, not only $b(k)$ but also its first two derivatives are continuous. Discontinuity occurs only in the third derivative. This is connected with the fact that the oscillating term in $Y_2(r)$, according to (5.68), is of the order r^{-4} for large r.

The oscillating term in (5.68) is of considerable interest, for it indicates the presence of an incipient crystal-lattice structure of long-range order in a series of eigenvalues. Even at large separations, two eigenvalues feel the natural periodicity of the lattice and have a slight preference for separation, which is an integer multiple of the mean level spacing. Unfortunately, the r^{-4} dependence of this effect makes it unobservable in practice.

5.4.3. Two-Level Functions for the Alternate Series

The method of integration over alternate variables provides information not only about the total eigenvalue probability density but also the separate densities of odd-numbered or even-numbered levels. For example, we may take in (5.9)

$$u(x) = 1, \qquad v(x) = 1 + a(x) \qquad \text{or} \qquad u\,x) = 1 + a(x), \qquad v(x) = 1.$$

Then

$$\rho_N\,(\text{alt}) = \tfrac{1}{2}[\rho_N\,(\text{odd}) + \rho_N\,(\text{even})], \tag{5.72}$$

where

$$\begin{aligned} \rho_N\,(\text{odd}) &= \rho_N(1 + a(x), 1) \\ \text{and} \qquad \rho_N\,(\text{even}) &= \rho_N(1, 1 + a(x)) \end{aligned} \tag{5.73}$$

The function $\rho_N(\text{odd})[\rho_N(\text{even})]$ is given by (5.53), (5.47), (5.48), and (5.49), where $A(x, y)$ is now replaced by $\epsilon(x - y)\,a(y)$ [by $\epsilon(x - y)\,a(x)$]. Alternatively, we may use (5.41) to (5.44), where $u(x) = 1 + a(x)$ and $v(x) = 1$. The result is

$$\rho_N{}^2\,(\text{alt}) = \det \begin{bmatrix} \lambda_{ij} & g_{ij} \\ -g_{ji} & \mu_{ij} \end{bmatrix}, \tag{5.74}$$

where

$$\begin{aligned} g_{ij} &= \delta_{ij} + \nu_{ij} \\ &= \delta_{ij} + \frac{1}{2}\int_{-\infty}^{\infty} a(x)\,\varphi_{2i}(x)\,\varphi_{2j}(x)\,dx \\ &\quad - \frac{1}{4}\iint_{-\infty}^{\infty} a(x)\,\epsilon(x - y)\,\varphi_{2i}(y)\,\varphi'_{2j}(x)\,dx\,dy, \end{aligned} \tag{5.75}$$

$$\lambda_{ij} = -\frac{1}{2}\iint_{-\infty}^{\infty} a(x)\,\epsilon(x - y)[\varphi_{2i}(x)\,\varphi_{2j}(y) - \varphi_{2i}(y)\,\varphi_{2j}(x)]\,dx\,dy \tag{5.76}$$

and

$$\mu_{ij} = \frac{1}{4}\int_{-\infty}^{\infty} a(x)[\varphi_{2i}(x)\,\varphi'_{2j}(x) - \varphi'_{2i}(x)\,\varphi_{2j}(x)]\,dx. \tag{5.77}$$

The expression for ρ involves correlation functions of the alternate eigenvalue series. The analysis proceeds as before, only the term $[\delta^2/\delta a(x)\,\delta a(y)]\sum_i \nu_{ii}$ is now missing from (5.58). The results are the following.

In an infinite eigenvalue series with mean spacing $D = 1$ let

$$\tfrac{1}{2}[1 - Y_2^e(\xi, \eta)]\,d\xi\,d\eta \tag{5.78}$$

be the probability of finding two eigenvalues in the intervals $(\xi, \xi + d\xi)$, $(\eta, \eta + d\eta)$, both of which belong to the same alternate series. Then

$$Y_2^e(\xi, \eta) = [s(r)]^2 - \Big[\int_0^r s(z)\,dz\Big]\Big[\frac{d}{dr}\,s(r)\Big]; \tag{5.79}$$

the notation is that of (5.64). For small and large r we find

$$Y_2^e(r) = 1 - \frac{\pi^4 r^4}{135} + \cdots \tag{5.80}$$

and

$$Y_2^e(r) = -\frac{\cos \pi r}{2r} + \frac{1 + (\pi/2)\sin \pi r}{\pi^2 r^2} - \frac{1 + \cos^2 \pi r}{\pi^4 r^4} + \cdots, \tag{5.81}$$

respectively. The corresponding two-level form factor (cf. Appendix A.10) is

$$b^e(k) = \begin{cases} 2 - 2|\,k\,| + |\,k\,| \ln(|(2|\,k\,| - 1)|), & |\,k\,| \leqslant 1, \\ 0, & |\,k\,| \geqslant 1. \end{cases} \tag{5.82}$$

The long-range order of the eigenvalue series appears much more strongly in (5.81) than in (5.68), and it shows clearest of all in the singularity of the Fourier transform $b^e(k)$ at $|\,k\,| = \frac{1}{2}$. This behavior of the alternate eigenvalues proves that the long-range crystalline structure of the level series is real. In a one-dimensional gas the operation of merely picking out alternate atoms for examination could not create long-range order if it had not been present to start with.

5.5. Level Spacings†

So far we have studied the probability frequencies R_n of finding n levels at a given set of positions with no restriction over the positions of the other levels. Now we come to the problem of the distribution of spacings. The probability density function is defined by the statement that $p(t)\,dt$ is the probability that a spacing, that is, the distance between any two adjacent levels, will lie between t and $t + dt$; this distance is measured in units of the mean level spacing. The distribution function $F(t)$ is defined by

$$F(t) = \int_0^t p(y)\,dy; \tag{5.83}$$

it is the probability that a given spacing will be less than or equal to t. If $E(t)$ is the probability that a randomly chosen interval of length t will be empty of energy levels, it is easy to see (cf. Appendix A.11) that

$$p(t) = \frac{d^2 E(t)}{dt^2}. \tag{5.84}$$

In what follows we calculate $E(t)$ when the eigenvalues lie in the region of constant density. This calculation can be considerably simplified if the excluded interval is taken symmetrically about the origin. Actually, it matters little, for the correlations among various levels depend only on their mutual separations as long as the levels lie in the vicinity of the origin.

This last statement about the correlations being unaffected by small displacements of the origin is corroborated, for example, by fixing two levels symmetrically about the origin, integrating others outside the interval between them and thus verifying (5.84), or by a direct calculation of the derivatives of $p(t)$ at $t = 0$ without any such symmetry of the excluded interval (M. L. Mehta [1], sections 5 and 7).

Let us then choose

$$u(x) = v(x) = \begin{cases} 0, & \text{if } -\theta \leqslant x \leqslant \theta, \\ 1, & \text{if } |x| > \theta, \end{cases} \tag{5.85}$$

in (5.9). Equation 5.30 is satisfied and therefore we obtain from (5.38)

$$E_m(\theta) = \det[g_{ij}]_{i,j=0,1,\ldots,m-1}, \tag{5.86}$$

† M. L. Mehta [1], Gaudin [1].

where from (5.85) and (5.33)

$$g_{ij} = \delta_{ij} - \int_{-\theta}^{\theta} \varphi_{2i}(x)\, \varphi_{2j}(x)\, dx. \tag{5.87}$$

To find the limit of (5.86) Gaudin [1] devised an ingenious application of the Fredholm theory of integral equations. Consider the integral equation

$$\lambda f(x) = \int_{-\theta}^{\theta} K_m(x, y) f(y)\, dy, \tag{5.88}$$

where $K_m(x, y)$ is the continuous, real, symmetric kernel:

$$K_m(x, y) = \sum_{i=0}^{m-1} \varphi_{2i}(x)\, \varphi_{2i}(y). \tag{5.89}$$

The eigenfunction $f(y)$ is of the form

$$f(y) = \sum_{i=0}^{m-1} c_i\, \varphi_{2i}(y). \tag{5.90}$$

Substituting (5.89) and (5.90) into (5.88) and remembering that $\varphi_0, \varphi_2, \ldots, \varphi_{2m-2}$ are linearly independent functions, we obtain the system of linear equations

$$\lambda c_i = \sum_{j=0}^{m-1} c_j \int_{-\theta}^{\theta} \varphi_{2i}(y)\, \varphi_{2j}(y)\, dy, \qquad i = 0, 1, \ldots, m-1. \tag{5.91}$$

This system will have a nonzero solution if and only if λ satisfies the following equation:

$$\det \left[\lambda \delta_{ij} - \int_{-\theta}^{\theta} \varphi_{2i}(y)\, \varphi_{2j}(y)\, dy \right]_{i,j=0,1,\ldots,m-1} = 0. \tag{5.92}$$

This algebraic equation in λ has m real roots; let them be $\lambda_0, \lambda_1, \ldots, \lambda_{m-1}$, so that

$$\det \left[\lambda \delta_{ij} - \int_{-\theta}^{\theta} \varphi_{2i}(y)\, \varphi_{2j}(y)\, dy \right]_{i,j=0,1,\ldots,m-1} \equiv \prod_{i=0}^{m-1} (\lambda - \lambda_i). \tag{5.93}$$

Comparing (5.93) with (5.86) and (5.87), we see that

$$E_m(\theta) = \prod_{i=0}^{m-1} (1 - \lambda_i), \tag{5.94}$$

where λ_i are the eigenvalues of the integral equation (5.88).

As $m \to \infty$, the quantity that tends to a finite limit is not K_m but Q_m, defined by

$$Q_m(\xi, \eta) = \left(\frac{2\sqrt{m}}{\pi}\right)^{-1} K_m(x, y), \tag{5.95}$$

with

$$\xi = \frac{2\sqrt{m}}{\pi} x, \qquad \eta = \frac{2\sqrt{m}}{\pi} y, \tag{5.96}$$

and (cf. Appendix A.9)

$$\lim_{m\to\infty} Q_m(\xi, \eta) = Q(\xi, \eta), \tag{5.97}$$

$$Q(\xi, \eta) = \frac{1}{2}\left[\frac{\sin(\xi - \eta)\,\pi}{(\xi - \eta)\,\pi} + \frac{\sin(\xi + \eta)\,\pi}{(\xi + \eta)\,\pi}\right]. \tag{5.98}$$

The limiting integral equation is then

$$\lambda f_1(\xi) = \int_{-(1/2)t}^{(1/2)t} Q(\xi, \eta) f_1(\eta)\, d\eta, \tag{5.99}$$

where

$$f_1(\xi) = f\left(\frac{\pi\xi}{2\sqrt{m}}\right) \quad \text{and} \quad t = 2\theta\frac{2\sqrt{m}}{\pi} = (\text{spacing}) \div$$

(mean spacing at the origin). (5.100)

Because $Q(\xi, \eta)$ is even in ξ, all solutions $f_1(\xi)$ are necessarily even. Hence all solutions of (5.99) are also solutions of

$$\lambda g(\xi) = \int_{-(1/2)t}^{(1/2)t} \frac{\sin(\xi \pm \eta)\,\pi}{(\xi \pm \eta)\,\pi} g(\eta)\, d\eta, \tag{5.101}$$

and conversely all even solutions of (5.101) are solutions of (5.99).

We note that the kernel $\sin(\xi + \eta)\,\pi/(\xi + \eta)\,\pi$ is the square of another symmetric kernel $t^{-1/2}e^{2\pi i\xi\eta/t}$, for

$$\int_{-(1/2)t}^{(1/2)t} t^{-1/2}\, e^{2\pi i\xi z/t} t^{-1/2}\, e^{2\pi i z\eta/t}\, dz = \frac{\sin(\xi + \eta)\,\pi}{(\xi + \eta)\,\pi}.$$

The kernel $e^{2\pi i\xi\eta/t}$ is equivalent to the kernel $\cos(2\pi\xi\eta/t)$ if we restrict ourselves to even solutions. The latter kernel is symmetric and real; therefore all eigenvalues corresponding to even solutions of

$$\mu g(\xi) = t^{-1/2} \int_{-(1/2)}^{(1/2)t} e^{2\pi i\xi\eta/t}\, g(\eta)\, d\eta \tag{5.102}$$

are real. A change of scale $\xi = \frac{1}{2}tx$, $\eta = \frac{1}{2}ty$ brings the limits of integration to ± 1 and (5.102) takes the form

$$\gamma f(x) = \frac{1}{2}\int_{-1}^{1} e^{(1/2)i\pi xyt} f(y)\, dy \tag{5.103}$$

with

$$f(x) = g(\tfrac{1}{2}tx) \qquad \text{and} \qquad \gamma = t^{-1/2}\mu.$$

If we obtain a complete set of solutions of (5.103), we would have done the same for (5.102) and (5.100) with the correspondence

$$\lambda = \mu^2 = t\gamma^2. \tag{5.104}$$

Out of these the even functions form a complete set of solutions of (5.99) and (5.98). Equation 5.94 then gives

$$E(t) = \prod_{i=0}^{\infty} (1 - t\gamma_{2i}^2), \tag{5.105}$$

where γ_{2i}, $i = 0, 1, 2,\ldots$ are the eigenvalues of (5.103) corresponding to even eigenfunctions. In other words, γ_{2i} are the eigenvalues of

$$\gamma f(x) = \int_{0}^{1} \cos(\tfrac{1}{2}\pi xyt) f(y)\, dy. \tag{5.106}$$

A careful examination of (5.103) shows that its solutions are the spheroidal functions [Robin, 1] that depend on the parameter t. These functions are defined as the solutions of the differential equation

$$(L - l)f(x) \equiv \left[(x^2 - 1)\frac{d^2}{dx^2} + 2x\frac{d}{dx} + \frac{\pi^2}{4}t^2x^2 - l\right] f(x) = 0, \tag{5.107}$$

which are regular at the points $x = \pm 1$. In fact it is easy to verify that the self-adjoint differential operator L commutes with the kernel $\exp(\frac{1}{2}i\pi xyt)$ defined over the interval $(-1, 1)$; that is,

$$\int_{-1}^{1} \exp(\tfrac{1}{2}i\pi xyt) L(y) f(y)\, dy = L(x)\int_{-1}^{1} \exp(\tfrac{1}{2}i\pi xyt) f(y)\, dy, \tag{5.108}$$

provided

$$(1 - x^2)f(x) = 0 = (1 - x^2)f'(x), \qquad |x| \to 1. \tag{5.109}$$

Equation 5.109 implies that $f(x)$ is regular at $x = \pm 1$. Hence (5.103) and (5.107) both have the same set of eigenfunctions. Once the eigenfunctions are known, the corresponding eigenvalues can be computed from (5.103) by setting, for example, $x = 0$:

$$\gamma_{2i} = \tfrac{1}{2}[f_{2j}(0)]^{-1} \int_{-1}^{1} f_{2j}(y)\, dy. \tag{5.110}$$

The spheroidal functions $f_0, f_2, f_4, \ldots$ form a complete set of even functions integrable over $(-1, 1)$. They are therefore the complete set of even functions for the kernel $e^{(1/2)i\pi xyt}$ and the kernel $\sin(\xi + \eta)\,\pi/(\xi + \eta)\,\pi$ and consequently the complete set of eigenfunctions for the limiting kernel $Q(\xi, \eta)$. Equations 5.105 and 5.110 give then the function $E(t)$, where $f_{2j}(x)$ are the spheroidal functions, that is, the solutions of the differential equation (5.107).

A few remarks about this analysis are in order.

1. We have put aside the question of convergence. In fact, for fixed ξ and η, $Q_m(\xi, \eta)$ tends to $Q(\xi, \eta)$ uniformly [Goursat, 1] with respect to ξ and η in any finite interval $|\xi|, |\eta| \leqslant \frac{1}{2}t$. Hence the Fredholm determinant of the kernel $Q_m(\xi, \eta)$ converges [Goursat, 1] to the Fredholm determinant of the limiting kernel $Q(\xi, \eta)$; that is,

$$\lim_{m\to\infty} E_m(\theta) = E(t). \tag{5.111}$$

2. For small values of t the spheroidal functions lie near the Legendre functions, and hence can be expanded in terms of them:

$$f_{2j}(x) = \sum_p d_{2p}(t, j)\, P_{2p}(x). \tag{5.112}$$

For example, to the 0th order we get from (5.110)

$$\gamma_{2j} = \frac{1}{2}\int_{-1}^{1} P_{2j}(x)dx/P_{2j}(0) \tag{5.113}$$

$$= 0, \quad \text{if} \quad j \neq 0$$

and

$$\gamma_0 = 1, \tag{5.114}$$

which gives

$$E(t) = 1 - t + \cdots.$$

Thus we get to the seventh power in t [Gaudin, 1; M. L. Mehta, 1; Kahn, 1]

$$E(t) = 1 - t + \frac{\pi^2}{36} t^3 - \frac{\pi^4}{1200} t^5 + \frac{\pi^4}{8100} t^6 + \frac{\pi^6}{70560} t^7 + \cdots, \quad (5.115)$$

which yields the probability distribution function

$$F(t) = 1 + \frac{d\,E(t)}{dt}$$

$$= \frac{\pi^2}{12} t^2 - \frac{\pi^4}{240} t^4 + \frac{\pi^4}{1350} t^5 + \frac{\pi^6}{10080} t^6 + \cdots \quad (5.116)$$

and the probability frequency function

$$p(t) = \frac{d^2E}{dt^2} = \frac{\pi^2}{6} t - \frac{\pi^4}{60} t^3 + \frac{\pi^4}{270} t^4 + \frac{\pi^6}{1680} t^5 + \cdots. \quad (5.117)$$

From (5.117) we can obtain the derivatives of $p(t)$ at $t = 0$. Thus the slope of the probability density function at the origin is

$$\left(\frac{dp}{dt}\right)_{t=0} = \frac{\pi^2}{6}. \quad (5.118)$$

The slope of the "Wigner surmise"

$$p_W(t) = \frac{\pi}{2} te^{-(\pi/4)t^2} \quad (1.2)$$

at the origin is $\pi/2$ and not $\pi^2/6$, as given by (5.118). Hence the "Wigner surmise" cannot be exact.

The coefficients of t and t^3 in (5.117) were computed elsewhere [M. L. Mehta, 1] by a different method.

3. The form (5.105) of $E(t)$ as an infinite product is useful for two reasons: (a) the spheroidal functions f_j are frequently encountered in other problems as well, and therefore are already extensively tabulated [Stratton *et al.*, 1], and (b) the infinite product converges rapidly. For these two reasons the numerical computation of $E(t)$ is easy. [A table of values of $E(t)$ and its first two derivatives is reproduced in Appendix A.12.]

5.6. Bounds for the Distribution Function $F(t)$†

In the expression (5.86) for $E_m(\theta)$ we can write

$$g_{ij} = 2\int_\theta^\infty \varphi_{2i}(x)\,\varphi_{2j}(x)\,dx. \tag{5.87'}$$

We then apply Gram's result (cf. Appendix A.13) to write

$$\det\left[\int_\theta^\infty \varphi_{2i}(x)\,\varphi_{2j}(x)\,dx\right]_{i,j=0,1,\ldots,m-1}$$

$$= \frac{1}{m!}\int_\theta^\infty\cdots\int \{\det[\varphi_{2i}(x_j)]_{\substack{i=0,1,\ldots,m-1\\ j=1,2,\ldots,m}}\}^2\,dx_1\cdots dx_m \tag{5.119}$$

In Section 5.2 we expressed the product $\prod_{i<j}(x_i - x_j)$ in the form of a determinant containing the oscillator wave functions φ_j. Following the same procedure step by step but in the reverse direction, we can convert the integrand in (5.119) back to the form

$$\exp\left(-\sum_1^m x_i^2\right)\prod_{i<j}(x_i^2 - x_j^2)^2,$$

except for some multiplicative factors. Collecting these factors, we obtain from (5.86), (5.87') and (5.119)

$$E_m(\theta) = \frac{2^m}{m!}\left\{\prod_{j=1}^{2m}[\Gamma(\tfrac{1}{2}j)]^{-1}\right\}\int_\theta^\infty\cdots\int dx_1\cdots dx_m$$

$$\times\exp\left(-\sum_1^m x_i^2\right)\prod_{1\leqslant i<j\leqslant m}(x_i^2 - x_j^2)^2. \tag{5.120}$$

Differentiation with respect to θ gives

$$\frac{dE_m(\theta)}{d\theta} = -\frac{2^m}{(m-1)!}\prod_1^{2m}\{\Gamma(\tfrac{1}{2}j)\}^{-1}\int_\theta^\infty\cdots\int dx_1\cdots dx_{m-1}$$

$$\times e^{-\theta^2}\exp\left(-\sum_1^{m-1} x_i^2\right)\prod_1^{m-1}(x_i^2 - \theta^2)^2\prod_{1\leqslant i<j\leqslant m-1}(x_i^2 - x_j^2)^2. \tag{5.121}$$

† M. L. Mehta [1].

Introducing new variables y_i defined by

$$x_i^2 = y_i^2 + \theta^2, \tag{5.122}$$

we may write

$$\sqrt{\pi}\, 2^{2(m-1)} \frac{m!(m-1)!}{(2m)!} \frac{dE_m}{d\theta} = -I_m(\theta)\, e^{-m\theta^2}, \tag{5.123}$$

where

$$I_m(\theta) = 2^{m-2}\sqrt{\pi} \prod_1^{2m-1} \{\Gamma(1 + \tfrac{1}{2}j)\}^{-1} \int_\theta^\infty \cdots \int dy_1 \cdots dy_{m-1}$$

$$\times \exp\left(-\sum_1^{m-1} y_i^2\right) \prod_{1 \leqslant i < j \leqslant m-1} (y_i^2 - y_j^2)^2 \prod_{i=1}^{m-1} [y_i^5 (y_i^2 + \theta^2)^{-1/2}]. \tag{5.124}$$

Applying Gram's result once again (cf. Appendix A.13), we write $I_m(\theta)$ as a determinant:

$$I_m(\theta) = \frac{\sqrt{\pi}}{2}(m-1)! \prod_1^{2m-1} [\Gamma(1 + \tfrac{1}{2}j)]^{-1} \det[\eta_{i+j}]_{i,j=1,2,\ldots,m-1}, \tag{5.125}$$

where

$$\eta_j(\theta) = 2\int_0^\infty e^{-y^2} y^{2j+1} (y^2 + \theta^2)^{-1/2}\, dy. \tag{5.126}$$

Inserting a power series expansion of $\eta_j(\theta)$ in (5.125), we obtain a power series expansion of $I_m(\theta)$ (cf. Appendix A.14):

$$I_m(\theta) = 1 - \tfrac{1}{3}(m-1)\,\theta^2 + \tfrac{1}{30}(m-1)(7m+1)\,\theta^4 + \cdots, \tag{5.127}$$

and taking the limit of (5.123) as $m \to \infty$, $\theta \to 0$ in such a way that $(4/\pi)\,\theta\,\sqrt{m} = t$ is finite we have

$$\frac{dE}{dt} = -I(t) \exp\left[-\left(\frac{\pi t}{4}\right)^2\right], \tag{5.128}$$

where

$$I(t) = \lim_{m \to \infty} I_m\left(\frac{\pi t}{4\sqrt{m}}\right). \tag{5.129}$$

The form (5.125), expressing $I_m(\theta)$ as a determinant, is convenient for calculations, as in arriving at (5.127), whereas the integral form (5.124) is useful for finding bounds for $I(t)$.

It is easy to prove (cf. Appendix A.15) that for all positive values of y_i

$$1 - \frac{1}{2}\sum_{i=1}^{m}\frac{\theta^2}{{y_i}^2} \leqslant \prod_{i=1}^{m}[y_i(y_i^2 + \theta^2)^{-1/2}] \leqslant 1. \tag{5.130}$$

The expansion (5.127) and the limiting procedure (5.129) then give the inequalities

$$1 - \frac{1}{3}\left(\frac{\pi t}{4}\right)^2 \leqslant I(t) \leqslant 1. \tag{5.131}$$

It follows from (5.128), (5.116), and (5.131) that we can obtain rigorous lower and upper bounds for the distribution function of the spacings $F(t)$:

$$F_L(t) \leqslant F(t) \leqslant F_R(t), \tag{5.132}$$

where

$$F_L(t) = 1 - \exp\left(-\frac{\pi^2 t^2}{16}\right) \tag{5.133}$$

and

$$F_R(t) = 1 - \left(1 - \frac{1}{3}\frac{\pi^2}{16}t^2\right)\exp\left(-\frac{\pi^2 t^2}{16}\right). \tag{5.134}$$

Because the differences $F - F_L$ and $F_R - F$ are everywhere positive, the difference between unity and the approximate values $\langle t_L\rangle$ and $\langle t_R\rangle$ obtained by substituting F_L and F_R for F into

$$\langle t\rangle = \int_0^\infty t\frac{dF}{dt}\,dt = \int_0^\infty [1 - F(t)]\,dt = 1 \tag{5.135}$$

provides a good estimation of the accuracy of the corresponding approximations to F. We obtain

$$\begin{aligned} \langle t_L\rangle - 1 &= \frac{2}{\sqrt{\pi}} - 1 \approx 0.1284, \\ 1 - \langle t_R\rangle &= 1 - \frac{5}{3\sqrt{\pi}} \approx 0.0597. \end{aligned} \tag{5.136}$$

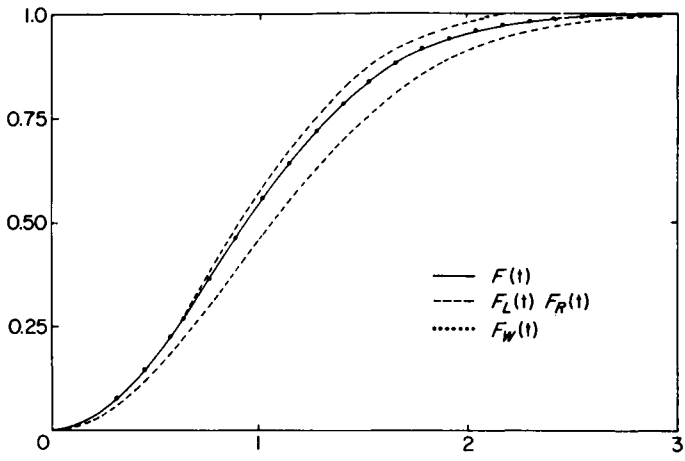


FIG. 5.1. The distribution functions of the spacings $F(t)$ and $F_W(t)$ lying between $F_L(t)$ and $F_R(t)$.

For visual comparison, Figure 5.1 is a plot of the functions F_L, F, F_R, and the Wigner surmise

$$F_W(t) = 1 - \exp\left(-\frac{\pi}{4}t^2\right), \tag{5.137}$$

whereas Figure 1.3 in Chapter 1 is a plot of $p(t)$ and the Wigner curve

$$p_W(t) = \frac{\pi}{2}t \exp\left(-\frac{\pi}{4}t^2\right). \tag{5.138}$$

6 / The Gaussian Unitary and Symplectic Ensembles

6.1. The Gaussian Unitary Ensemble†

In Chapter 5 we studied in detail the properties of the eigenvalues of random real symmetric matrices chosen from an orthogonal ensemble. It is of some interest to compare the results obtained there with the corresponding results for the eigenvalues of random Hermitian matrices chosen from a unitary ensemble. According to (3.39), the joint probability density function for the eigenvalue of matrices from the unitary ensemble is given by

$$P_{N2}(x_1, \ldots, x_N) = C_{N2} \exp\left(-\sum_1^N x_j^2\right) \prod_{1 \leqslant k < j \leqslant N} (x_k - x_j)^2. \tag{6.1}$$

This would be a model for the energy levels of a complex system without invariance under time reversal.

To obtain correlation functions and the like with (6.1) is quite simple. As in Section 5.2, we express the product of differences as the Vandermonde determinant and introduce the normalized oscillator wave functions

$$\varphi_j(x) = (2^j j!\, \sqrt{\pi})^{-1/2}\, e^{(1/2)x^2} \left(-\frac{d}{dx}\right)^j e^{-x^2},$$

so that

$$P_{N2}(x_1, \ldots, x_N) = C_{N2} \prod_{j=0}^{N-1} [2^{-j}(2^j j!\, \sqrt{\pi})^{1/2}]^2 \{\det[\varphi_k(x_j)]_{\substack{k=0,1,\ldots,N-1 \\ j=1,2,\ldots,N}}\}^2. \tag{6.2}$$

To determine C_{N2}, we integrate over all the variables, and by expanding the square of the determinant we see that all the cross

† E. P. Wigner [6].

terms drop out in the integral and every square term contributes unity, for

$$\int_{-\infty}^{\infty} \varphi_k(x)\,\varphi_j(x)\,dx = \delta_{kj} = \begin{cases} 1, & k = j, \\ 0, & k \neq j. \end{cases}$$

Because the number of square terms is $N!$, we finally get

$$1 \equiv C_{N2} 2^{-(1/2)N(N-1)} N! \pi^{(1/2)N} \prod_{j=0}^{N-1} j!, \tag{6.3}$$

which agrees with (3.42) for $\beta = 2$.

Substituting C_{N2} from (6.3) into (6.2), we get

$$P_{N2}(x_1, \ldots, x_N) = \frac{1}{N!} \{\det[\varphi_{k-1}(x_j)]_{k,j=1,2,\ldots,N}\}^2. \tag{6.4}$$

6.1.1. Correlation and Cluster Functions

To obtain the n-level correlation function we must integrate (6.4) over all variables except $x_1, x_2, \ldots, x_n$, and for this it is convenient to make the Laplace expansion of the determinant with respect to its first n columns, which depend on the variables $x_1, x_2, \ldots, x_n$:

$$\det[\varphi_k(x_j)] = \sum_{k_1 \leq k_2 \leq \ldots \leq k_n} \begin{vmatrix} \varphi_{k_1}(x_1) & \cdots & \varphi_{k_1}(x_n) \\ \cdots & \cdots & \cdots \\ \varphi_{k_n}(x_1) & \cdots & \varphi_{k_n}(x_n) \end{vmatrix} \Phi_{k_1 \cdots k_n}(x_{n+1}, \ldots, x_N), \tag{6.5}$$

where $\Phi_{k_1 \cdots k_n}$ is the algebraic complement of

$$\det[\varphi_{k_l}(x_j)]_{l,j=1,2,\ldots,n}.$$

On squaring and integrating, we see that all cross terms drop out and each square term takes the same coefficient $(N - n)!$. Thus for the n-level correlation function we obtain

$$\begin{aligned} R_n^{(u)}(x_1, \ldots, x_n) &\equiv \frac{N!}{(N-n)!} \int_{-\infty}^{\infty} \cdots \int P_{N2}(x_1, \ldots, x_N)\, dx_{n+1} \cdots dx_N \\ &= \sum_{k_1 \leq k_2 \leq \cdots \leq k_n} \begin{vmatrix} \varphi_{k_1}(x_1) & \cdots & \varphi_{k_1}(x_n) \\ \cdots & \cdots & \cdots \\ \varphi_{k_n}(x_1) & \cdots & \varphi_{k_n}(x_n) \end{vmatrix}^2 \\ &= \frac{1}{n!} \sum_{k_1, k_2, \ldots, k_n} \{\det[\varphi_{k_l}(x_j)]_{l,j=1,\ldots,n}\}^2. \end{aligned} \tag{6.6}$$

By applying Gram's result (cf. Appendix A.13), we write (6.6) in the more elegant form

$$R_n^{(u)}(x_1, \ldots, x_n) = \det[K_N(x_i, x_j)]_{i,j=1,2,\ldots,n}, \tag{6.7}$$

where

$$K_N(x, y) = \sum_{j=0}^{N-1} \varphi_j(x)\, \varphi_j(y). \tag{6.8}$$

Putting $n = 1$ in (6.6) or (6.7), we get the level density

$$\sigma_N(x) = K_N(x, x) = \sum_{j=0}^{N-1} \varphi_j^2(x). \tag{6.9}$$

As $N \to \infty$, (6.9) tends (cf. Appendix A.8) to the semicircle law,

$$\sigma_N(x) \to \sigma(x) = \begin{cases} \dfrac{1}{\pi}(2N - x^2)^{1/2}, & x^2 < 2N, \\ 0 & x^2 > 2N. \end{cases} \tag{6.10}$$

Putting $n = 2$ in (6.7), we get the two-level correlation function

$$R_2^{(u)}(x, y) = \sigma_N(x)\, \sigma_N(y) - [K_N(x, y)]^2, \tag{6.11}$$

hence the two-level cluster function

$$T_{2u}(x, y) = [K_N(x, y)]^2 = \Big[\sum_{j=0}^{N-1} \varphi_j(x)\, \varphi_j(y)\Big]^2. \tag{6.12}$$

Taking the limit as $N \to \infty$, this equation and the definition (5.6) (cf. Appendix A.9) give

$$\begin{aligned} Y_{2u}(\xi, \eta) &= \lim_{N\to\infty} \Big[\frac{\pi}{(2N)^{1/2}} \sum_{j=0}^{N-1} \varphi_j(x)\, \varphi_j(y)\Big]^2 \\ &= [s(r)]^2 = \Big[\frac{\sin \pi r}{\pi r}\Big]^2, \end{aligned} \tag{6.13}$$

with

$$r = |\, \xi - \eta\, |, \qquad \xi = \frac{1}{\pi}(2N)^{1/2} x, \qquad \eta = \frac{1}{\pi}(2N)^{1/2} y.$$

The two-level form factor is

$$b_u(k) = \int_{-\infty}^{\infty} Y_{2u}(r)\, e^{2\pi i k r}\, dr$$

$$= \begin{cases} 1 - |k|, & |k| \leqslant 1, \\ 0, & |k| \geqslant 1. \end{cases} \tag{6.14}$$

It must be noted that the three and higher level correlation functions can all be expressed in terms of the level density and the two-level function $K_N(x, y)$, as is evident from (6.7). This is a particular feature of the unitary ensemble.

6.1.2. Level Spacings

It is quite simple to integrate (6.4) over all variables from $-\infty$ to $-\theta$ and again from θ to ∞. If we apply Gram's result (cf. Appendix A.13), we get

$$E_{uN}(\theta) = \int_{\text{out}} \cdots \int P_{N2}(x_1, \ldots, x_N)\, dx_1 \cdots dx_N$$

$$= \det \left[\delta_{ij} - \int_{-\theta}^{\theta} \varphi_i(x)\, \varphi_j(x)\, dx \right]_{i,j=0,1,\ldots,N-1}, \tag{6.15}$$

where "out" indicates that all the variables are integrated over the entire real axis except for the interval $[-\theta, \theta]$.

Equation 6.15 corresponds to (5.86) and (5.87) and gives the probability that no eigenvalues will be found in the interval $[-\theta, \theta]$.

Because

$$\delta_{ij} - \int_{-\theta}^{\theta} \varphi_i(x)\, \varphi_j(x)\, dx$$

is zero whenever $i + j$ is odd, the determinant in (6.15) factorizes as a product of two determinants:

$$E_{uN}(\theta) = E_m(\theta)\, E'_m(\theta), \tag{6.16}$$

where $E_m(\theta)$ is given by (5.86)

$$E_m(\theta) = \det \left[\delta_{ij} - \int_{-\theta}^{\theta} \varphi_{2i}(x)\, \varphi_{2j}(x)\, dx \right]_{i,j=0,1,\ldots,m-1} \tag{5.86}$$

and $E'_m(\theta)$ is formed correspondingly from the odd oscillator functions

$$E'_m(\theta) = \det\left[\delta_{ij} - \int_{-\theta}^{\theta} \varphi_{2i+1}(x)\,\varphi_{2j+1}(x)\,dx\right]_{i,j=0,1,\ldots,m-1}, \tag{6.17}$$

where m is the largest integer equal to or smaller than $\frac{1}{2}N$.

Now we may repeat the analysis of Chapter 5 and write

$$E_{uN}(\theta) = \Big[\prod_{k=1}^{m}(1-\lambda_k)\Big]\Big[\prod_{k=1}^{m}(1-\mu_k)\Big], \tag{6.18}$$

where λ_k are the eigenvalues of the integral equation (5.88)

$$\lambda f(x) = \int_{-\theta}^{\theta}\Big[\sum_{j=0}^{m-1}\varphi_{2j}(x)\,\varphi_{2j}(y)\Big]f(y)\,dy$$

and μ_k are the eigenvalues of the integral equation

$$\mu f(x) = \int_{-\theta}^{\theta}\Big[\sum_{j=0}^{m-1}\varphi_{2j+1}(x)\,\varphi_{2j+1}(y)\Big]f(y)\,dy. \tag{6.19}$$

The passage to the limit $N \to \infty$ is exactly the same as in Section 5.5, except that we now need the even as well as the odd spheroidal functions. The final result is

$$E_{uN}(\theta) \to E_u(t) = E(t)\,E'(t), \tag{6.20}$$

with

$$E(t) = \prod_{j=0}^{\infty}(1 - t\lambda_j^2) \tag{5.105}$$

and

$$E'(t) = \prod_{j=0}^{\infty}(1 - t\mu_j^2), \tag{6.21}$$

here λ_j and μ_j are the eigenvalues of the integral equations

$$\lambda f(y) = \int_0^1 \cos\left(\frac{\pi}{2}tyz\right)f(z)\,dz \tag{5.106}$$

and

$$\mu f(y) = \int_0^1 \sin\left(\frac{\pi}{2}tyz\right)f(z)\,dz, \tag{6.22}$$

whereas t is the level spacing measured in units of the mean level spacing at the origin:

$$t = (2\theta) \div \frac{\pi}{(2N)^{1/2}}. \tag{6.23}$$

The probability distribution (the probability density) function can be obtained by differentiating $E_u(t)$ once (twice):

$$F_u(t) = 1 + \frac{dE_u}{dt}, \tag{6.24}$$

$$p_u(t) = \frac{d^2E_u}{dt^2}. \tag{6.25}$$

As in the derivation of (5.84) we have assumed here that the various correlations are unaffected by small displacements of the origin. If we do a direct calculation by fixing one (two) level(s) and integrating the remaining levels outside a certain interval edged by the fixed level(s), we get back (6.24), while (6.25) is replaced by

$$p_u(t) = 4\frac{dE}{dt}\frac{dE'}{dt}. \tag{6.25'}$$

Thus (6.25) and (6.25') are both correct if the following is an identity:

$$E\frac{d^2E'}{dt^2} + E'\frac{d^2E}{dt^2} = 2\frac{dE}{dt}\frac{dE'}{dt}. \tag{6.25''}$$

A proof of (6.25'') is reproduced in Appendix A.31.

A table of numerical values of the various quantities is reproduced in Appendix A.12. For small values of t the quantities $E'(t)$, $E_u(t)$, $F_u(t)$, and $p_u(t)$ may also be expanded in powers of t. A few terms are listed in the same appendix.

6.2. Gaussian Symplectic Ensemble

The joint probability density function for the eigenvalues of a self-dual Hermitian matrix taken from a Gaussian symplectic ensemble was derived in Chapter 3 [cf. (3.32)]:

$$P_{N4}(x_1, \ldots, x_N) = C_{N4} \exp\left(-2\sum_1^N x_i^2\right) \prod_{i<j} (x_i - x_j)^4, \tag{6.26}$$

which forms a model for the energy levels of a complex system with half odd integral spin and time-reversal invariance but no rotational symmetry.

The statistical properties of N points with (6.26) as their joint probability density function have not been investigated in detail for two reasons:

1. No experimental situation is yet known in which the symplectic ensemble is applicable.

2. For the corresponding circular ensembles, which we study next (Chapters 8 to 11), a remarkable relation (Theorem 9.1) reduces the study of the symplectic ensemble to that of the orthogonal ensemble. We strongly suspect a similar relation for the Gaussian ensembles because all the other known statistical properties in the two sets of ensembles are identical when proper limits are taken. Such a relation has not yet been discovered. In view of this lack, we now discuss a method of dealing with integrals that contain the fourth power of the product of differences.

6.2.1. The Determination of the Normalization Constant

It is convenient to express the fourth power of the product of differences as a "confluent alternant," a limiting form obtained from the simple alternant when variables become equal in groups. It is shown in Appendix A.16 that

$$\prod_{1\leqslant i<j\leqslant N}(x_i - x_j)^4 = \det[x_i^{\,j}, jx_i^{j-1}]; \tag{6.27}$$

here and in the rest of this section the indices i and j vary over the respective sets of integers:

$$\begin{aligned} i &= 1, 2, \ldots, N, \\ j &= 0, 1, \ldots, 2N-1. \end{aligned} \tag{6.28}$$

As in Chapter 5 we introduce Hermite polynomials for the new variables

$$y_i = x_i\sqrt{2} \tag{6.29}$$

to obtain

$$\begin{aligned} \det[x_i^{\,j}, jx_i^{j-1}] &= \left[\frac{1}{\sqrt{2}}\right]^{(1/2)N(N-1)4}\left(\prod_j 2^{-j}\right)\det[H_j(y_i), H_j'(y_i)] \\ &= 2^{-N(3N-2)}\det[H_j(y_i), H_j'(y_i)], \end{aligned} \tag{6.30}$$

where the prime denotes the derivative. In terms of the oscillator wave functions (5.15), this gives

$$\exp\left(-2\sum_1^N x_i^2\right)\prod_{1\leqslant i<l\leqslant N}(x_i-x_l)^4 = a\det[\varphi_j(y_i),\sqrt{2j}\,\varphi_{j-1}(y_i)], \tag{6.31}$$

where

$$\begin{aligned} a &= 2^{-N(3N-2)}\prod_j (2^j j!\sqrt{\pi})^{1/2} \\ &= 2^{-N(2N-3/2)}\pi^{(1/2)N}\prod_{l=0}^{N-1}[(2l)!\,(2l+1)^{1/2}]. \end{aligned} \tag{6.32}$$

The average value of $\prod_i u(x_i)$ is, therefore, given by

$$\begin{aligned} \left\langle\prod_i u(x_i)\right\rangle = C_{N4}\,a\,2^{-(1/2)N}\int_{-\infty}^{\infty}\cdots\int\prod_i\left[dy_i u\left(\frac{y_i}{\sqrt{2}}\right)\right] \\ \times\det[\varphi_j(y_i),(2j)^{1/2}\varphi_{j-1}(y_i)]. \end{aligned} \tag{6.33}$$

Performing the integrations over the y_i in (6.33), we get (cf. Appendix A.7)

$$\left\langle\prod_i u(x_i)\right\rangle = C_{N4}\,a\,2^{-(1/2)N}N!\,(\det[g_{jj'}])^{1/2}, \tag{6.34}$$

where

$$\begin{aligned} g_{jj'} &= \int_{-\infty}^{\infty} dy[(2j')^{1/2}\varphi_j(y)\varphi_{j'-1}(y) - (2j)^{1/2}\varphi_{j-1}(y)\varphi_{j'}(y)]\,u\left(\frac{y}{\sqrt{2}}\right) \\ &= \int_{-\infty}^{\infty}\sqrt{2}\,dx[(2j')^{1/2}\varphi_j(x\sqrt{2})\varphi_{j'-1}(x\sqrt{2}) \\ &\qquad - (2j)^{1/2}\varphi_{j-1}(x\sqrt{2})\varphi_{j'}(x\sqrt{2})]\,u(x). \end{aligned} \tag{6.35}$$

We may add to any row (column) a linear combination of the other rows (columns) without changing the value of the determinant. We replace $g_{jj'}$ with $f_{jj'}$, which in turn is replaced by $h_{jj'}$, defined successively by the following equations:

$$\begin{aligned} f_{2i-1,j} &= g_{2i-1,j}, \\ f_{2i-2,j} &= g_{2i-2,j} + \left(\frac{2i-2}{2i-3}\right)^{1/2} f_{2i-4,j} \end{aligned} \tag{6.36}$$

and

$$h_{j,2i-1} = f_{j,2i-1}\,,$$
$$h_{j,2i-2} = f_{j,2i-2} + \left(\frac{2i-2}{2i-3}\right)^{1/2} h_{j,2i-4}\,. \tag{6.37}$$

Equation 6.34 then becomes

$$\left\langle \prod_i u(x_i) \right\rangle = C_{N4}\, a\, 2^{-(1/2)N} N!\, (\det[h_{jj'}])^{1/2}, \tag{6.38}$$

where from (6.36) and (6.37)

$$h_{2i-1,2i'-1} = g_{2i-1,2i'-1}\,, \tag{6.39}$$

$$h_{2i-2,2i'-1} = \sum_{l=1}^{i} \left[\frac{\Gamma(i)\,\Gamma(l-\frac{1}{2})}{\Gamma(l)\,\Gamma(i-\frac{1}{2})}\right]^{1/2} g_{2l-2,2i'-1} \tag{6.40}$$

and

$$h_{2i-2,2i'-2} = \sum_{l=1}^{i} \sum_{l'=1}^{i'} \left[\frac{\Gamma(i)\,\Gamma(l-\frac{1}{2})}{\Gamma(l)\,\Gamma(i-\frac{1}{2})} \frac{\Gamma(i')\,\Gamma(l'-\frac{1}{2})}{\Gamma(l')\,\Gamma(i'-\frac{1}{2})}\right] g_{2l-2,2l'-2}\,. \tag{6.41}$$

Writing

$$h_{2i-2,2i'-1} = [2(2i'-1)]^{1/2}\,(\delta_{ii'} + \nu_{ii'}), \tag{6.42}$$

$$h_{2i-1,2i'-1} = [2(2i'-1)\,2(2i-1)]^{1/2}\,\mu_{ii'} \tag{6.43}$$

and

$$h_{2i-2,2i'-2} = \lambda_{ii'}\,, \tag{6.44}$$

we have

$$\left\langle \prod_i u(x_i) \right\rangle = C_{N4}\, a\, 2^{-(1/2)N} N! \prod_i [2(2i-1)]^{1/2} \times \left\{ \det \begin{bmatrix} \lambda_{ii'} & (\delta_{ii'} + \nu_{ii'}) \\ -(\delta_{ii'} + \nu_{i'i}) & \mu_{ii'} \end{bmatrix} \right\}^{1/2}. \tag{6.45}$$

If we write $u(x) = 1$, we get

$$g_{jj'} = (2j')^{1/2}\,\delta_{j,j'-1} - (2j)^{1/2}\,\delta_{j',j-1} \tag{6.46}$$

so that

$$h_{2i-1,2i'-1} = h_{2i-2,2i'-2} = 0,$$
$$h_{2i-2,2i'-1} = [2(2i'-1)]^{1/2}\,\delta_{ii'}\,, \tag{6.47}$$

or

$$\lambda_{ii'} = \mu_{ii'} = \nu_{ii'} = 0. \tag{6.48}$$

Equations 6.45, 6.48, and 6.32 then give the normalization constant

$$\begin{aligned} C_{N4}^{-1} &= a \cdot 2^{-(1/2)N} N! \prod_i [2(2i-1)]^{1/2} \\ &= (2\pi)^{(1/2)N} 4^{-(1/2)N - N(N-1)} 2^{-N} \prod_1^N (2i)! \end{aligned} \tag{6.49}$$

Thus (3.42) is verified for $\beta = 4$. Substituting this value of C_{N4} into (6.45), we finally obtain

$$\left\langle \prod_i u(x_i) \right\rangle = \left\{ \det \begin{bmatrix} \lambda_{ii'} & (\delta_{ii'} + \nu_{ii'}) \\ -(\delta_{ii'} + \nu_{i'i}) & \mu_{ii'} \end{bmatrix} \right\}^{1/2}. \tag{6.50}$$

By writing $u(x) = 1 + a(x)$ and expanding in powers of $a(x)$, as in Section 5.4, we may obtain, in principle, all the correlation (cluster) functions. For example, the level density and the two-level correlation function are given by (5.55), (5.58), and (5.54), where ν_{ij}, λ_{ij}, and μ_{ij} are now given by (6.42), (6.43), and (6.44). However, because the limiting process $N \to \infty$ is more tedious in this case, it is hard to get at the quantities of interest.

7 / Brownian Motion Model†

7.1. Stationary Ensembles

In Chapter 4 we exploited the idea that the probability $P(x_1,..., x_N)$, (3.40) for the eigenvalues of a random matrix to lie in unit intervals around the points $x_1,..., x_N$,

$$P(x_1,..., x_N) = C_{N\beta}e^{-\beta W}, \tag{7.1}$$

$$W = \frac{1}{2}\sum_1^N x_j^2 - \sum_{i<j} \ln | x_i - y_j |, \tag{7.2}$$

is identical with the probability density of the positions of N unit charges free to move on the infinite straight line $-\infty < x < \infty$ under the influence of forces derived from the potential energy (7.2), according to the laws of classical mechanics, in a state of thermodynamical equilibrium at a temperature given by

$$kT = \beta^{-1}. \tag{7.3}$$

This system of point charges in thermodynamical equilibrium is called the stationary Coulomb gas model or simply the Coulomb gas model, which corresponds to the Gaussian ensembles.

7.2. Nonstationary Ensembles

In this chapter we present an idea of Dyson, generalizing the notion of a matrix ensemble in such a way that the Coulomb gas model acquires meaning not only as a static model in timeless thermodynamical equilibrium but as a dynamical system that may be in an arbitrary nonequilibrium state changing with time. The word "time" in this

† Dyson [4].

chapter always refers to a fictitious time which is a property of the mathematical model and has nothing to do with real physical time.

When we try to interpret Coulomb gas as a dynamical system, we naturally consider it first as an ordinary conservative system in which the charges move as Newtonian particles and exchange energy with one another only through the electric forces arising from the potential (7.2). We then have to give meaning to the velocity of each particle and to regulate the behavior of the random matrix H in such a way that the eigenvalues have the normal Newtonian property of inertia. No reasonable way of doing this has yet been found. Perhaps there is no such way.

After considerable and fruitless efforts to develop a Newtonian theory of ensembles, Dyson [4] discovered that the correct procedure is quite different and much simpler. The x_j should be interpreted as positions of particles in Brownian motion [Chandrasekhar, 1; Uhlenbeck and Ornstein, 1; Wang and Uhlenbeck, 1]. This means that the particles have no well-defined velocities nor do they possess inertia. Instead, they feel frictional forces resisting their motion. The gas is not a conservative system, for it is constantly exchanging energy with its surroundings through these frictional forces. The potential (7.2) still operates on the particles in the following way. The particle at x_j experiences an external electric force

$$E(x_j) = -\frac{\partial W}{\partial x_j} = -x_j + \sum_{\substack{i \\ (i \neq j)}} \frac{1}{x_j - x_i} \tag{7.4}$$

in addition to the local frictional force and the constantly fluctuating force giving rise to the Brownian motion.

The equation of motion of the Brownian particle at x_j may be written as

$$\frac{d^2x_j}{dt^2} = -f\frac{dx_j}{dt} + E(x_j) + A(t), \tag{7.5}$$

where f is the friction coefficient and $A(t)$ is a rapidly fluctuating force. For $A(t)$ we postulate the usual properties [Uhlenbeck and Ornstein, 1]

$$\langle A(t_1)\, A(t_2) \cdots A(t_{2n+1})\rangle = 0, \tag{7.6}$$

$$\langle A(t_1)\, A(t_2) \cdots A(t_{2n})\rangle = \sum_{\text{pairs}} \langle A(t_i)\, A(t_j)\rangle \langle A(t_k)\, A(t_l)\rangle \cdots, \tag{7.7}$$

and

$$\langle A(t_1)\, A(t_2)\rangle = \frac{2kT}{f}\,\delta(t_1 - t_2), \tag{7.8}$$

where the summation in (7.7) extends over all distinct ways in which the $2n$ indices can be divided into n pairs.

There is nothing new in the integration of the Langevin equation (7.5). After long enough time for the effect of the initial velocity to become negligible, let x_1 , x_2 ,..., x_N be the positions of the particles at time t. At a later time $t + \delta t$ let these positions be changed to $x_1 + \delta x_1$, $x_2 + \delta x_2$,..., $x_N + \delta x_N$. The δx_j , $j = 1, 2,..., N$, will in general be different for every member of the ensemble. They are random variables. Using (7.6), (7.7), and (7.8) we find that to the first order in the small quantities

$$f\langle \delta x_j\rangle = E(x_j)\,\delta t, \tag{7.9}$$

$$f\langle (\delta x_j)^2\rangle = 2kT\,\delta t, \tag{7.10}$$

and all other ensemble averages, for example, $\langle \delta x_j\, \delta x_l\rangle$, $\langle (\delta x_j)^2\, \delta x_l\rangle$, $\langle (\delta x_j)^3\rangle$, are of a higher order in δt.

An alternative description of Brownian motion is obtained by deriving the Fokker–Planck or Smoluchowski equation. Let $P(x_1 , x_2 ,..., x_N ; t)$ be the time-dependent joint probability density that the particles will be at the positions x_j at time t. Assuming that the future evolution of the system is completely determined by its present state, with no reference to its past (that is, the process is a Markov process), we obtain

$$P(x_1 ,..., x_N ; t + \delta t) = \int\cdots\int P(x_1 - \delta x_1 ,..., x_N - \delta x_N ; t)$$

$$\times\, \psi(x_1 - \delta x_1 ,..., x_N - \delta x_N ; \delta x_1 ,..., \delta x_N ; \delta t)\, d(\delta x_1)\cdots d(\delta x_N), \tag{7.11}$$

where ψ under the integral sign is the probability that the positions of the particles will change from $x_1 - \delta x_1$,..., $x_N - \delta x_N$ to x_1 ,..., x_N in a time interval δt. Expanding both sides of (7.11) in a power series of δx_j , δt, using (7.9) and (7.10), and going to the limit $\delta t \to 0$, we get [Uhlenbeck and Ornstein, 1]

$$f\frac{\partial P}{\partial t} = \sum_{j=1}^{N}\left\{kT\frac{\partial^2 P}{\partial x_j} - \frac{\partial}{\partial x_j}[E(x_j)\, P]\right\}. \tag{7.12}$$

Equation 7.12 describes the development of the Coulomb gas with time. If we start from an arbitrary initial probability density P at time $t = t_0$, a unique solution of (7.12) will exist for all $t \geqslant t_0$. Any solution of this sort we call a *time-dependent Coulomb gas model.*

Equation 7.12 implies in turn (7.9) and (7.10). To see this we multiply both sides of (7.12) by x_j and integrate over all x_i. Making the usual assumptions that $P(x_1, ..., x_N ; t)$, as well as its derivatives, vanish quite fast on the boundary, we get on partial integration

$$f \frac{d}{dt} \langle x_j \rangle = \langle E(x_j) \rangle, \tag{7.13}$$

where

$$\langle F \rangle = \int F(x_1, ..., x_N) P(x_1, ..., x_N ; t) \, dx_1 \cdots dx_N$$

is the ensemble average of F. Starting at the positions $x_1, ..., x_N$ and executing the motion for a small time interval δt, we find that (7.13) is the same as (7.9). Similarly, by multiplying by x_j^2 and integrating (7.12) we have

$$f \frac{d}{dt} \langle x_j^2 \rangle = 2kT + 2\langle x_j E(x_j) \rangle,$$

which together with $\langle (\delta x_j)^2 \rangle = \langle x_j^2 \rangle - \langle x_j \rangle^2$ yields (7.10).

Thus the descriptions of the motion by (7.9) and (7.10), and by (7.12) are equivalent. Also there exists a unique solution to (7.12) which is independent of time, and this time independent solution is given by (7.1) and (7.2).

A Brownian motion model can also be constructed for the matrix H, of which x_j are the eigenvalues. The independent real parameters $H_{ij}^{(\lambda)}$; $1 \leqslant i < j \leqslant N$, $0 \leqslant \lambda \leqslant \beta - 1$, which determine all the matrix elements of H, are $p = \frac{1}{2}N(N + 1) + \frac{1}{2}N(N - 1)(\beta - 1)$ in number. Let us denote them by H_μ, where μ is a single index that runs from 1 to p and replaces the three indices i, j, and λ. Suppose that the parameters H_μ have the values $H_1, H_2, ..., H_p$ at time t and $H_1 + \delta H_1, ..., H_p + \delta H_p$ at a later time $t + \delta t$. Brownian motion of H is defined by requiring that each δH_μ be a random variable with the ensemble averages

$$f\langle \delta H_\mu \rangle = -H_\mu \, \delta t, \tag{7.14}$$

$$f\langle (\delta H_\mu)^2 \rangle = g_\mu kT \, \delta t, \tag{7.15}$$

where

$$g_\mu = g_{ij}^{(\lambda)} = 1 + \delta_{ij} = \begin{cases} 2, & \text{if } i = j, \\ 1, & \text{if } i \neq j. \end{cases} \tag{7.16}$$

All other averages are of a higher order in δt. This is a Brownian motion of the simplest type, the various components H_μ being completely uncoupled and each being subject to a fixed simple harmonic force. The Smoluchowski equation which corresponds to (7.14) and (7.15) is

$$f \frac{\partial P}{\partial t} = \sum_\mu \left[\tfrac{1}{2} g_\mu kT \frac{\partial^2 P}{\partial H_\mu^{\,2}} + \frac{\partial}{\partial H} (H_\mu P) \right], \tag{7.17}$$

where $P(H_1, \ldots, H_p ; t)$ is the time-dependent joint probability density of H_μ. The solution to (7.17) which corresponds to a given initial condition $H = H'$ at $t = 0$, is known explicitly [Uhlenbeck and Ornstein, 1].

$$P(H, t) = C(1 - q^2)^{-(1/2)p} \exp\left[- \frac{\operatorname{tr}(H - qH')^2}{2kT(1 - q^2)} \right], \tag{7.18}$$

$$q = \exp\left[\frac{-t}{f} \right]. \tag{7.19}$$

The solution shows that the Brownian process is invariant under symmetry preserving unitary transformations of the matrix H; in fact, the awkward-looking factor g_μ in (7.15) is put in to ensure this invariance. When $t \to \infty$, $q \to 0$, and the probability density (7.18) tends to the stationary form,

$$P(H_1, \ldots, H_p) = (\text{constant}) \exp\left(\frac{-1}{2kT} \operatorname{tr} H^2 \right), \tag{7.20}$$

which is the unique time-independent solution of (7.17). Note that with the relation (7.3) between β and the temperature kT (7.20) is essentially the same as (2.72).

We are now in a position to state the main result of this chapter.

Theorem 7.1. *When the matrix H executes a Brownian motion according to the simple harmonic law* (7.14), (7.15), *starting from any initial conditions whatever, its eigenvalues $x_1, x_2, \ldots, x_N$ execute a Brownian motion that obeys the equations of motion* (7.9), (7.10), *and* (7.12) *of the time-dependent Coulomb gas.*

To prove the theorem we need only show that (7.9) and (7.10) follow from (7.14) and (7.15). Suppose, then, that (7.14) and (7.15) hold. We have seen that the process described by (7.14) and (7.15) is independent of the representation of H. Therefore we may choose the representation so that H is diagonal at time t. The instantaneous values of H_μ at time t are then

$$H_{jj}^{(0)} = x_j, \qquad j = 1, 2, \ldots, N, \tag{7.21}$$

and all other components are zero. At a later time $t + \delta t$ the matrix $H + \delta H$ is no longer diagonal and its eigenvalues $x_j + \delta x_j$ must be calculated by perturbation theory. We have to the second order in δH

$$\delta x_j = \delta H_{jj}^{(0)} + \sum_{\substack{i \\ (i \neq j)}} \sum_{\lambda=0}^{\beta-1} \frac{(\delta H_{ij}^{(\lambda)})^2}{x_j - x_i}. \tag{7.22}$$

Higher terms in the perturbation series will not contribute to the first order in δt. When we take the ensemble average on each side of (7.22) and use (7.14), (7.21), (7.15), (7.3), and (7.4), the result is (7.9). When we take the ensemble average of $(\delta x_j)^2$, only the first term on the right side of (7.22) contributes to the order δt, and this term gives (7.10) by virtue of (7.15) and (7.16). The theorem is thus proved.

When the limit $t \to \infty$ is taken, Theorem 7.1 reduces to Theorem 3.2. This new proof of Theorem 3.2 is in some respects more illuminating. It shows how the repulsive Coulomb potential (7.2), pushing apart each pair of eigenvalues, arises directly from the perturbation formula (7.22). It has long been known that perturbations generally split levels that are degenerate in an unperturbed system. We now see that this splitting effect of perturbations is quantitatively identical with the repulsive force of the Coulomb gas model.

Theorem 7.1 is a much stronger statement than Theorem 3.2. It shows that the electric force (7.4), acting on the eigenvalues x_j, has a concrete meaning for any matrix H whatever, not only for an ensemble of matrices in stationary thermal equilibrium. The force $E(x_j)$ is precisely proportional to the mean rate of drift of x_j which occurs when the matrix H is subjected to a random perturbation.

7.3. Some Ensemble Averages

We now describe a general property of the time-dependent Coulomb gas model which may be used to calculate a few ensemble averages. Dyson observed that if $G = G(x_1, \ldots, x_N)$ is any function of the positions of the charges, not depending explicitly on time, then the time variation of $\langle G \rangle$, the ensemble average of G, is governed by the equation

$$f \frac{d}{dt} \langle G \rangle = -\sum_j \left\langle \frac{\partial W}{\partial x_j} \frac{\partial G}{\partial x_j} \right\rangle + kT \sum_j \left\langle \frac{\partial^2 G}{\partial x_j^2} \right\rangle. \tag{7.23}$$

This equation is obtained by multiplying (7.12) throughout by G and partial integrations; W is given by (7.2).

As a first example, choose

$$R = \sum_j x_j^2 \tag{7.24}$$

for G so that

$$\frac{\partial W}{\partial x_j} \frac{\partial R}{\partial x_j} = -2 \sum_{\substack{i \\ (i \neq j)}} \frac{x_j}{x_j - x_i} + 2x_j^2,$$

$$\frac{\partial^2 R}{\partial x_j^2} = 2,$$

and (7.23) becomes

$$f \frac{\partial \langle R \rangle}{\partial t} = -2\langle R \rangle + N(N-1) + 2kTN$$

$$= 2(R_\infty - \langle R \rangle), \tag{7.25}$$

with

$$R_\infty = \tfrac{1}{2} N(N-1) + kTN. \tag{7.26}$$

The solution of (7.25) is

$$\langle R \rangle = R_0 q^2 + R_\infty (1 - q^2), \tag{7.27}$$

where q is given by (7.19) and R_0 is the value of $\langle R \rangle$ at $t = 0$. Equation 7.27 shows that the ensemble average $\langle R \rangle$ approaches its equilibrium value R_∞ with exponential speed as $t \to \infty$.

Next take $G = W$ in (7.23), so that

$$\left(\frac{\partial W}{\partial x_j}\right)^2 = \sum_{\substack{i \\ (i \neq j)}} \left[\left(\frac{1}{x_j - x_i}\right)^2 - \frac{2x_j}{x_j - x_i}\right]$$

$$+ \sum_{\substack{i,l \\ (i,j,l \text{ all} \\ \text{different})}} [(x_j - x_i)(x_j - x_l)]^{-1} + x_j^2 \tag{7.28}$$

and

$$\frac{\partial^2 W}{\partial x_j^2} = -1 + \sum_{\substack{i \\ (i \neq j)}} (x_j - x_i)^{-2}. \tag{7.29}$$

On performing a summation over j the second term in (7.28) drops out (cf. Appendix A.20), whereas the second term in the first bracket gives $-2N(N-1)$. Substituting in (7.23) and simplifying, we get

$$f\frac{\partial \langle W \rangle}{\partial t} = (kT - 1) \sum_{\substack{i,j \\ (i \neq j)}} \langle (x_j - x_i)^{-2} \rangle + (N^2 - N + NkT) - \sum_j \langle x_j^2 \rangle. \tag{7.30}$$

For the stationary Coulomb gas at temperature kT the left side of (7.30) vanishes and (7.26) may be used on the right. Thus we find a "virial theorem" for the stationary gas:

$$\sum_{\substack{i,j \\ (i \neq j)}} \langle (x_j - x_i)^{-2} \rangle = \frac{N(N-1)}{2(1 - kT)}. \tag{7.31}$$

The probability density of eigenvalues becomes proportional to $| x_i - x_j |^{\beta}$, when two eigenvalues x_i, x_j come close together. The ensemble average of $(x_j - x_i)^{-2}$ is therefore defined only for $\beta > 1$ and (7.30) and (7.31) hold only for $kT < 1$.

An especially interesting case, $\beta = 1$, requires a passage to the limit in (7.30). As $kT \to 1$, we have for any fixed value of Δ

$$\begin{aligned} \lim(kT - 1) \int_{-\Delta}^{\Delta} | y |^{\beta - 2}\, dy &= \lim(kT - 1)(\beta - 1)^{-1}\, 2\Delta^{\beta - 1} \\ &= -2. \end{aligned} \tag{7.32}$$

We obtain the correct limit in (7.30) if we replace

$$(kT - 1)(x_j - x_i)^{-2}$$

with

$$-2(x_j - x_i)^{-1}\, \delta(x_j - x_i), \tag{7.33}$$

which has a well-defined meaning as an ensemble average when $kT = 1$, for the probability density then contains a factor $| x_j - x_i |$; (7.30) thus becomes in the limit

$$f\frac{d \langle W \rangle}{dt} = -2 \sum_{\substack{i,j \\ (i \neq j)}} \langle | x_j - x_i |^{-1}\, \delta(x_j - x_i) \rangle + N^2 - \sum_j \langle x_j^2 \rangle, \qquad kT = 1. \tag{7.34}$$

The corresponding "virial theorem" is

$$\sum_{\substack{i,j \\ (i \neq j)}} \langle (x_j - x_i)^{-1} \, \delta(x_j - x_i) \rangle = \tfrac{1}{4} N(N-1), \qquad kT = 1 \qquad (7.35)$$

for the stationary gas.

8 / Circular Ensembles†

8.1. General Remarks

In the preceding chapters we presented a detailed study of the Gaussian ensembles. We pointed out at the end of Chapter 2 that the requirements of invariance and the statistical independence of various independent components seem to overrestrict the possible choices. In particular, the various values of the matrix elements are not equally weighted. Rosenzweig has tried to answer this question in part with a consideration of the "fixed-strength ensemble" in which the joint probability density function for the matrix elements is taken to be proportional to the Dirac delta function $\delta(\mathrm{tr}\, H^2 - r^2)$, where r is a fixed number. However, a uniform probability density cannot be defined on the infinite real line.

Because of this unsatisfactory feature Dyson [1–3] introduced his circular ensembles, as follows.

Suppose that the system is characterized not by its Hamiltonian H but by a unitary matrix S, whose elements give the transition probabilities between the various states. The matrix S is unitary; its eigenvalues are therefore of the form $e^{i\theta_j}$, where the angles θ_j are real and may be taken to lie between 0 and 2π. The matrix S is a function of the Hamiltonian H of the system. This functional dependence need not be specified. All that is needed is that for small ranges of variation the θ_j be linear functions of the eigenvalues x_j of H. To help the reader's imagination he may think of a relation such as

$$S = e^{i\tau H} \quad \text{or} \quad S = \frac{1 - i\tau H}{1 + i\tau H}. \tag{8.1}$$

However, such a definite relation between S and H cannot be correct in the large. We will deliberately leave this relation vague because we are going to restrict the order of our matrices to $N \times N$, where

† Dyson [1].

N is very large but finite. And this cannot represent, say, a nucleus, in the large, for the real nucleus has an infinite number of energy levels. Like the Gaussian ensembles, the circular ensembles are gross mutilations of the over-all actual situation. The most we can expect of such models is that in any energy region that is small compared to the total excitation energy the statistical distribution of the levels will be correctly reproduced. With no further apologies we make the following fundamental assumption:

The statistical behavior of n consecutive levels of an actual system, whenever n is small compared with the total number of levels, is the same as that of n consecutive angles $\theta_1, \theta_2, \ldots, \theta_n$, where n is small compared with N.

8.2. The Orthogonal Ensemble

According to the analysis of Chapter 2, a system having time-reversal invariance and rotational symmetry or having time-reversal invariance and integral spin will be characterized by a symmetric S. Following Dyson, we define the orthogonal circular ensemble E_{1c} of symmetric unitary matrices S by assigning the probabilities in the following way.

Every symmetric unitary S (cf. Appendix A.23) can be written as

$$S = U^T U, \tag{8.2}$$

where U is unitary. Define a small neighborhood of S by

$$S + dS = U^T(1 + i\,dM)U, \tag{8.3}$$

where dM is a real symmetric matrix with elements dM_{ij} and the elements dM_{ij} for $i \leqslant j$ vary independently in some small intervals of length $d\mu_{ij}$. The "volume" of this neighborhood is defined by

$$\mu_1(dS) = \prod_{i \leqslant j} d\mu_{ij}\,. \tag{8.4}$$

The ensemble E_{1c} is defined by the statement: *The probability that a matrix from the ensemble E_{1c} lies between S and $S + dS$ is proportional to*

$$P(S)\,dS = \frac{1}{V_1}\,\mu_1(dS), \tag{8.5}$$

where V_1 is the total volume of the space T_{1c} of unitary symmetric matrices of order $N \times N$ and therefore, a normalization constant.

For this definition to have a meaning one must be sure that $\mu_1(dS)$ does not depend on the choice of U in (8.2). The fact that it is so can be easily verified. Let

$$S = U^T U = V^T V, \tag{8.6}$$

where both U and V are unitary. The matrix

$$R = VU^{-1} \tag{8.7}$$

is unitary and also satisfies

$$R^T R = (U^T)^{-1} V^T V U^{-1} = (U^T)^{-1} U^T U U^{-1} = 1. \tag{8.8}$$

Therefore R is real and orthogonal. Let

$$\mu_1'(dS) = \prod_{i \leqslant j} d\mu_{ij} \tag{8.9}$$

be the volume derived from V as $\mu_1(dS)$ was derived from U. We now have

$$S + dS = V^T(1 + i\, dM')\, V \tag{8.10}$$

with

$$dM' = R\, dM R^{-1}. \tag{8.11}$$

To prove that $\mu_1(dS) = \mu_1'(dS)$ we need to show that the Jacobian

$$J = \det\left[\frac{\partial(dM_{ij}')}{\partial(dM_{ij})}\right] \tag{8.12}$$

has absolute value unity when dM, dM' are real symmetric matrices related by (8.11). A proof of this is given in Appendix A.17. Thus the volume $\mu_1(dS)$ is unique. Incidently, we have established that for a fixed S the unitary matrix U in (8.2) is undetermined precisely to the extent of a transformation

$$U \to RU, \tag{8.13}$$

where R is an arbitrary real orthogonal matrix.

The motivation for the choice of the ensemble E_{1c} will be made clearer by the following theorem.

Theorem 8.1. *The orthogonal ensemble E_{1c} is uniquely defined in the space of unitary symmetric matrices of order $N \times N$ by the property of being invariant under every automorphism*

$$S \to W^T S W \tag{8.14}$$

of T_{1c} into itself, where W is any $N \times N$ unitary matrix.

Theorem 8.1 comprises two statements: (a) *that E_{1c} is invariant under the automorphisms* (8.14) *and* (b) *that it is unique.* To prove (a) we suppose that a neighborhood $S + dS$ of S is transformed into a neighborhood $S' + dS'$ of S' by the automorphism (8.14). Equations 8.2 and 8.3 then hold and

$$S' = W^T S W = V^T V, \qquad V = UW \tag{8.15}$$

$$S' + dS' = V^T(1 + i\,dM)\,V. \tag{8.16}$$

The volumes $\mu_1(dS)$ and $\mu_1(dS')$ are then identical by definition as the same dM is occurring in (8.3) and (8.16). This proof of (a) is trivial, for we could choose a convenient unitary matrix V in (8.15), and it was already shown that the value of $\mu_1(dS)$ does not depend on the choice of V. To prove (b) let E_1' be any ensemble invariant under (8.14). The probability density $P'(S)\,dS$ associated with E_1' will define a certain volume $\mu_1'(dS)$ of the neighborhood of S in the space T_{1c}. The ratio

$$\varphi(S) = \frac{\mu_1'(dS)}{\mu_1(dS)} \tag{8.17}$$

is a function of S defined on T_{1c} and invariant under the transformations (8.14). If $S = U^T U$, we choose $W = U^{-1}$ in (8.14) so that S is transformed to unity and therefore $\varphi(S)$ is a constant. Thus the probability densities in E_{1c} and E_{1c}' are proportional. Also they are both normalized to unity and are therefore identical.

8.3. Symplectic Ensemble

Next we consider systems with half-integral spin and time-reversal invariance but no rotational symmetry. In this section we use the quaternion notation developed in Chapter 2. The systems are described by self-dual unitary quaternion matrices (cf. Chapter 2)

$$S^R \equiv -ZS^TZ = S, \qquad S^+ = S^{-1}. \tag{8.18}$$

Once again we have to assign a probability that a matrix chosen randomly from the space T_{4c} of self-dual unitary quaternion matrices of order $N \times N$ will lie between S and $S + dS$. This is done as follows:

Every matrix S in T_{4c} can be written as

$$S = U^R U, \tag{8.19}$$

where U is unitary. To see that this is possible, observe that in the ordinary language without quaternions SZ is an antisymmetric unitary matrix and can be reduced to the canonical form

$$SZ = VZV^T, \tag{8.20}$$

where V is unitary. Choosing $U = (ZV)^T$ then gives (8.19). (See also Appendix A.23.) For a given S the unitary matrix U in (8.19) is precisely undetermined to the extent of a transformation

$$U \to BU, \tag{8.21}$$

where B is an arbitrary symplectic matrix. For a proof it is sufficient to observe that the dual of a product of matrices is the product of their duals taken in the reverse order; and a symplectic matrix is, by definition, one that satisfies

$$B^R B = BB^R = 1. \tag{8.22}$$

A small neighborhood of S in T_{4c} is defined by

$$S + dS = U^R(1 + i\, dM)\, U \tag{8.23}$$

where dM is a self-dual quaternion real matrix with elements

$$dM_{ij} = dM_{ij}^{(0)} + \sum_{\alpha=1}^{3} dM_{ij}^{(\alpha)} e_\alpha . \tag{8.24}$$

The real coefficients $dM_{ij}^{(\alpha)}$ satisfy

$$dM_{ij}^{(0)} = dM_{ji}^{(0)}, \qquad dM_{ij}^{(\alpha)} = -dM_{ji}^{(\alpha)}, \qquad \alpha = 1, 2, 3. \tag{8.25}$$

There are $N(2N - 1)$ independent real variables $dM_{ij}^{(\alpha)}$ and they are allowed to vary over some small intervals of lengths $d\mu_{ij}^{(\alpha)}$. The neighborhood of S, thus defined, is assigned a volume

$$\mu_4(dS) = \prod_{i<j} \prod_{\alpha=1}^{3} d\mu_{ij}^{(\alpha)} \prod_{i \leqslant j} d\mu_{ij}^{(0)}. \tag{8.26}$$

In terms of this volume the symplectic ensemble E_{4c} is defined in exactly the same way as E_{1c} was defined in terms of the volume (8.4). The statistical weight of the neighborhood dS in T_{4c} is

$$P(S)\,dS = \frac{1}{V_4}\,\mu_4(dS), \tag{8.27}$$

where V_4 is the total volume of the space T_{4c} of self-dual unitary quaternion matrices of order $N \times N$.

We can now repeat almost without change the arguments in Section 8.2. We must first prove that the volume $\mu_4(dS)$ is independent of the choice of U in (8.19). This involves showing that the Jacobian J has absolute value unity, where

$$J = \det\left[\frac{\partial(dM'^{(\alpha)}_{ij})}{\partial(dM^{(\alpha)}_{ij})}\right], \tag{8.28}$$

$$dM' = B\,dMB^{-1}, \tag{8.29}$$

and B is symplectic. As before, Appendix A.17 contains a proof of this.

The analog of Theorem 8.1 is the following.

Theorem 8.2. *The symplectic ensemble E_{4c} is uniquely defined in the space T_{4c} of self-dual unitary quaternion matrices of order $N \times N$ by the property of being invariant under every automorphism*

$$S \to W^R S W \tag{8.30}$$

of T_{4c} into itself, where W is any $N \times N$ unitary quaternion matrix.

Theorem 8.2 can be proved by following word for word the proof of Theorem 8.2, the operation of transposition being replaced by that of taking the dual.

8.4. Unitary Ensemble

A system without time-reversal symmetry is associated with an arbitrary unitary matrix S not restricted to be symmetric or self-dual. A neighborhood of S in the space T_{2c} of all unitary $N \times N$ matrices is defined by

$$S + dS = U(1 + i\,dM)\,V, \tag{8.31}$$

where U and V are any unitary matrices that satisfy the equation $S = UV$ and dM is an infinitesimal Hermitian matrix with elements

$dM_{ij} = dM_{ij}^{(0)} + i\, dM_{ij}^{(1)}$. The real components $dM_{ij}^{(0)}$, $dM_{ij}^{(1)}$ are N^2 in number and are allowed to vary independently over small intervals of lengths $d\mu_{ij}^{(0)}$, $d\mu_{ij}^{(1)}$. The volume $\mu_2(dS)$ is defined by the equation

$$\mu_2(dS) = \prod_{i \leqslant j} d\mu_{ij}^{(0)} \prod_{i<j} d\mu_{ij}^{(1)} \tag{8.32}$$

and is independent of the choice of U and V. The ensemble E_{2c} gives to each neighborhood dS the statistical weight

$$P(S)\, dS = \frac{1}{V_2} \mu_2(dS), \tag{8.33}$$

where V_2 is the total volume of the space T_{2c}.

The invariance property of E_{2c}, analogous to Theorems 8.1 and 8.2, is stated in Theorem 8.3.

Theorem 8.3. *The unitary ensemble E_{2c} is uniquely defined in the space T_{2c} of all $N \times N$ unitary matrices by the property of being invariant under every automorphism*

$$S \to USV \tag{8.34}$$

of T_{2c} into itself, where U and V are any two $N \times N$ unitary matrices.

This theorem merely expresses the well-known result that $\mu_2(dS)$ is the invariant group-measure of the N-dimensional unitary group $U(N)$.

8.5. The Joint Probability Density Function for the Eigenvalues

We give below a few lemmas which will be used subsequently. A proof of the first two lemmas is given in Appendix A.23.

Lemma 8.1. *Let S be any unitary symmetric $N \times N$ matrix. Then there exists a real orthogonal matrix R which diagonalizes S; that is,*

$$S = R^{-1}ER, \tag{8.35}$$

where E is diagonal. The diagonal elements of E are complex numbers $e^{i\theta_j}$ lying on the unit circle.

Lemma 8.2. *Let S be a unitary self-dual quaternion matrix of order $N \times N$. Then there exists a symplectic matrix B such that*

$$S = BEB^{-1}, \tag{8.36}$$

where E is diagonal and scalar (cf. Section 2.4). The diagonal elements of E are N complex numbers $e^{i\theta_j}$ on the unit circle; each is repeated twice.

Lemma 8.3. *Let S be a unitary matrix. Then there exists a unitary matrix U such that*

$$S = UEU^{-1}, \tag{8.37}$$

where E is diagonal. The diagonal elements of E are N complex numbers $e^{i\theta_j}$ on the unit circle.

Though this result is well-known [Wigner, 2], we stated it here for completeness.

We are now in a position to prove the main result of this chapter.

Theorem 8.4. *In the ensemble $E_{\beta c}$ the probability of finding the eigenvalues $e^{i\phi_j}$ of S with an angle in each of the intervals $[\theta_j, \theta_j + d\theta_j]$; $j = 1,..., N$ is given by*

$$P_{N\beta}(\theta_1,..., \theta_N)\, d\theta_1 \cdots d\theta_N, \tag{8.38}$$

where

$$P_{N\beta}(\theta_1,..., \theta_N) = C'_{N\beta} \prod_{1 \leqslant l < j \leqslant N} | e^{i\theta_l} - e^{i\theta_j} |^{\beta}. \tag{8.39}$$

Here $\beta = 1$ for orthogonal, $\beta = 4$ for symplectic, and $\beta = 2$ for unitary circular ensembles. The constant $C'_{N\beta}$ is fixed by normalization.

Proof: 1. Let $\beta = 1$. By Lemma 8.1 every S in T_{1c} can be diagonalized in the form

$$S = R^{-1}ER, \tag{8.40}$$

with R orthogonal. We now wish to express the volume $\mu_1(dS)$ in terms of the volumes $\mu(dE)$ and $\mu(dR)$, defined for the neighborhoods of the matrices E and R, respectively. A small neighborhood of E is given by

$$dE = iE\, d\theta, \tag{8.41}$$

where $d\theta$ means a diagonal matrix with elements $d\theta_1, \ldots, d\theta_N$. To find the neighborhood of R we differentiate

$$RR^T = 1, \tag{8.42}$$

thus getting

$$R(dR)^T + (dR)\,R^T = 0 \tag{8.43}$$

showing that the infinitesimal matrix

$$dA \equiv (dR)R^T = -R(dR)^T \tag{8.44}$$

is a real antisymmetric matrix with elements dA_{ij}. The volumes $\mu(dE)$ and $\mu(dR)$ are given by

$$\mu(dE) = \prod_{j=1}^{N} d\theta_j\,, \tag{8.45}$$

$$\mu(dR) = \prod_{1 \leqslant i < j \leqslant N} dA_{ij}\,. \tag{8.46}$$

The volume $\mu(dS)$ is defined by (8.4), where dM is given by (8.3) and U is any unitary matrix satisfying (8.2). Differentiating (8.40) and using (8.2), (8.3), (8.41), (8.42), and (8.44), we obtain

$$iRU^T\,dMUR^{-1} = -dA\;E + iE\,d\theta + E\,dA, \tag{8.47}$$

which is the relation between dM, $d\theta$, and dA. Since E is a diagonal unitary matrix, it has a square root F which is also diagonal with elements $\pm e^{i\theta_j/2}$. There is an ambiguity in the sign of each element, but it does not matter how these signs are chosen. A convenient choice for U satisfying (8.2) is then

$$U = FR \tag{8.48}$$

by virtue of (8.40). With this choice of U, (8.47) reduces to

$$iF\,dM\,F = -dA\,F^2 + iF^2\,d\theta + F^2\,dA, \tag{8.49}$$

or

$$dM = d\theta + i(F^{-1}\,dA\,F - F\,dA\,F^{-1}). \tag{8.50}$$

Equation 8.50 gives dM_{ij} in terms of the quantities $d\theta_j$, dA_{ij}, and θ_j for each pair of indices i, j; namely

$$dM_{jj} = d\theta_j \tag{8.51}$$

$$dM_{ij} = 2 \sin[\tfrac{1}{2}(\theta_i - \theta_j)]\, dA_{ij}, \qquad i \neq j. \tag{8.52}$$

Assembling the definitions (8.4), (8.45), and (8.46), we deduce from (8.51) and (8.52)

$$\mu(dS) = \prod_{l<j} \left| 2 \sin\left(\frac{\theta_l - \theta_j}{2}\right) \right| \mu(dE)\, \mu(dR)$$

$$= \prod_{l<j} | e^{i\theta_l} - e^{i\theta_j} |\, \mu(dE)\, \mu(dR). \tag{8.53}$$

Now keep the angles $\theta_1, \ldots, \theta_N$ fixed and integrate (8.53) with respect to the parameters dA_{ij} over the entire allowed range. This will give (8.38) with P_{N1} given by (8.39). Thus the theorem is proved for the orthogonal case.

2. Next let $\beta = 4$. The matrix S is now diagonalized with the help of a symplectic matrix

$$S = B^{-1}EB \tag{8.54}$$

(Lemma 8.2). The infinitesimal matrix

$$dA = dB\, B^R \tag{8.55}$$

is quaternion real and anti-Hermitian. The components of dA are $dA_{ij}^{(\alpha)}$, which are real. They are antisymmetric in i, j for $\alpha = 0$ and symmetric in i, j for $\alpha = 1, 2$, and 3. The volume $\mu(dB)$ is now given by

$$\mu(dB) = \prod_{i,j,\alpha} dA_{ij}^{(\alpha)}. \tag{8.56}$$

The volume $\mu(dS)$ is given by (8.26) with dM given by (8.23). The matrix dM is Hermitian and quaternion real. The algebra leading up to (8.50) goes exactly as before. Equation 8.51 still holds, the diagonal elements dM_{jj} being real scalar quaternions with only one independent component. Equation 8.52 now holds separately for each of the four quaternion components $\alpha = 0, 1, 2, 3$. The equation analogous to (8.53) is

$$\mu(dS) = \left(\prod_{l<j} | e^{i\theta_l} - e^{i\theta_j} |^4 \right) \mu(dE)\, C, \tag{8.57}$$

where C does not depend on the θ_j. The power 4 in (8.57) arises from the fact that every nondiagonal element dM_{ij}, $i < j$, gives according to (8.52) four factors corresponding to the four components $\alpha = 0$, 1, 2, and 3. Note also that the C in (8.57) is not equal to $\mu(dB)$, for the diagonal components $dA_{jj}^{(\alpha)}$ with $\alpha = 1, 2, 3$ do not occur in (8.51), whereas $\mu(dB)$ contains their product. For our purposes it is sufficient to know that C does not depend on the θ_j.

The rest of the proof proceeds as in the case $\beta = 1$.

3. Lastly let $\beta = 2$. In this case (8.40) holds with a unitary R. The infinitesimal matrix dA is now anti-Hermitian, and the diagonal elements dA_{jj} are pure imaginary. The real part $dA_{ij}^{(0)}$ of the nondiagonal elements dA_{ij} is antisymmetric in i, j, whereas the imaginary part $dA_{ij}^{(1)}$ is symmetric in i, j. Equation (8.51) holds in this case as well; (8.52) holds separately for the real and imaginary components $dM_{ij}^{(0)}$ and $dM_{ij}^{(1)}$ of the nondiagonal elements dM_{ij}. The equation analogous to (8.53) is therefore

$$\mu(dS) = \left\{ \prod_{l<j} | e^{i\theta_l} - e^{i\theta_j} |^2 \right\} \mu(dE)\, C, \tag{8.58}$$

where C does not depend on the θ_j. As in the case $\beta = 4$, this is sufficient for our purposes.

The theorem is thus established for all three cases.

Counting the dimensions of T_{1c}, T'_{1c}, etc., proceeds as in the case of the Gaussian ensembles (cf. the end of Chapter 2 and Appendix A.2).

9 / Circular Ensembles. Correlation Functions, Spacing Distribution, etc.

9.1. Orthogonal Ensemble†

In this chapter we repeat briefly the considerations of Chapters 5, 6, and 7 for the case of circular ensembles.

First let us take the orthogonal ensemble, which is most important from a practical point of view. Here again the method of integration over alternate variables can be applied. For this purpose we use the identity

$$| e^{i\theta_j} - e^{i\theta_l} | = i^{-1}(e^{i\theta_j} - e^{i\theta_l}) \exp[-\tfrac{1}{2}i(\theta_j + \theta_l)], \qquad \text{if} \quad \theta_j \geqslant \theta_l\,, \tag{9.1}$$

to write

$$\prod_{1 \leqslant j < l \leqslant N} | e^{i\theta_j} - e^{i\theta_l} | = i^{-(1/2)N(N-1)} \exp\left[-\tfrac{1}{2}i(N-1)\sum_1^N \theta_j\right] \prod_{1 \leqslant j < l \leqslant N} (e^{i\theta_l} - e^{i\theta_j}), \tag{9.2}$$

where the θ_j are supposed to be ordered

$$-\pi \leqslant \theta_1 \leqslant \theta_2 \leqslant \cdots \leqslant \theta_N \leqslant \pi. \tag{9.3}$$

Writing the product of differences in (9.2) as a Vandermonde determinant and multiplying the column containing the powers of $e^{i\theta_j}$ by $\exp[-\frac{1}{2}i(N-1)\theta_j]$, we have

$$\prod_{1 \leqslant j < l \leqslant N} | e^{i\theta_j} - e^{i\theta_l} | = i^{-(1/2)N(N-1)} \det[\exp(ip\theta_j)]; \tag{9.4}$$

$$p = -\tfrac{1}{2}(N-1),\ -\tfrac{1}{2}(N-3),\ldots,\ \tfrac{1}{2}(N-1), \qquad j = 1, 2, \ldots, N, \tag{9.5}$$

if θ_j are ordered as in (9.3).

† Dyson [3].

Let us take N even, $N = 2m$. As in Chapter 5, let $u(\theta)$, $v(\theta)$ be functions defined over the range $(-\pi, \pi)$ and consider the average

$$H = \left\langle \prod_{\text{alt}} u(\theta_j) \prod'_{\text{alt}} v(\theta_l) \right\rangle, \tag{9.6}$$

taken with respect to the orthogonal ensemble $\beta = 1$, defined by (8.39). Here $\prod_{\text{alt}}$ denotes a product taken over a set of m alternate points θ_j as they lie on the unit circles and $\prod'_{\text{alt}}$ a product over the remaining m points. This average can again be calculated by integration over alternate variables, using (9.4). We define

$$f'_{pq} = \frac{1}{2} \int_{-\pi}^{\pi}\int u(\theta)\, v(\phi)\, \epsilon(\theta - \phi)(e^{ip\phi+iq\theta} - e^{ip\theta+iq\phi})\, d\theta\, d\phi, \tag{9.7}$$

with

$$\epsilon(\theta - \phi) = \begin{cases} 1, & \theta > \phi, \\ 0, & \theta = \phi, \\ -1, & \theta < \phi, \end{cases} \tag{9.8}$$

and do the integrations step by step as in Section 5.2. The result is

$$H^2 = C'^2_{N1}(-1)^m[(2m)!]^2 \det[f'_{pq}], \tag{9.9}$$

$$p, q = -m + \tfrac{1}{2}, -m + \tfrac{3}{2}, \ldots, m - \tfrac{1}{2}. \tag{9.10}$$

By reversing the order of the columns we can write (9.9) as

$$H^2 = C'^2_{N1}[(2m)!]^2 \det[f'_{p,-q}]\,. \tag{9.11}$$

To determine C'_{N1} put $u(\theta) = v(\theta) = 1$ in (9.6) so that by (9.7) we get

$$f'_{p,-q} = \frac{4\pi}{ip}\, \delta_{pq} \tag{9.12}$$

and (9.11) gives

$$1 = C'_{N1}(2m)!\, (4\pi)^m \frac{2^{2m} m!}{(2m)!}\,. \tag{9.13}$$

Inserting this value of C'_{N1} into (9.11), we finally get

$$H^2 = \det[f_{pq}], \tag{9.14}$$

$$p, q = -m + \tfrac{1}{2}, -m + \tfrac{3}{2}, \ldots, m - \tfrac{1}{2}, \tag{9.10}$$

where

$$f_{pq} = \frac{ip}{4\pi} f'_{p,-q}$$

$$= \frac{ip}{8\pi} \int\int_{-\pi}^{\pi} u(\theta)\, v(\phi)\, \epsilon(\theta - \phi)[e^{i(p\phi - q\theta)} - e^{i(p\theta - q\phi)}]\, d\theta\, d\phi. \quad (9.15)$$

If the functions $u(\theta)$ and $v(\theta)$ satisfy the relation

$$u(-\theta)\, v(-\phi) = u(\theta)\, v(\phi), \quad (9.16)$$

then

$$f_{-p,-q} = f_{p,q} \quad (9.17)$$

and there are further simplifications. We now have

$$\det[f_{p,q} + f_{-p,q}] = \det[f_{p,q} - f_{-p,q}], \qquad p, q = \tfrac{1}{2}, \tfrac{3}{2}, \ldots, m - \tfrac{1}{2},$$

and the first power of H may be written as a determinant:

$$H = \det[F_{pq}]_{p,q=\frac{1}{2},\frac{3}{2},\ldots,m-\frac{1}{2}}, \quad (9.18)$$

with

$$F_{pq} = f_{p,q} + f_{-p,q}$$

$$= \frac{p}{4\pi} \int\int_{-\pi}^{\pi} u(\theta)\, v(\phi)\, \epsilon(\theta - \phi)(\cos p\phi \sin q\theta - \cos p\theta \sin q\phi)\, d\theta\, d\phi. \quad (9.19)$$

9.1.1. Two Level Correlation Functions

Write $u(\theta) = v(\theta) = 1 + a(\theta)$ in (9.6). Then

$$H = \left\langle \prod_{1}^{2m} (1 + a(\theta_j)) \right\rangle. \quad (9.20)$$

Equations 9.14 and 9.15 now become

$$H^2 = \det[\delta_{pq} + r_{pq}], \quad (9.21)$$

where

$$r_{pq} = \frac{1}{2\pi} \left(1 + \frac{p}{q}\right) \int_{-\pi}^{\pi} a(\theta)\, e^{i(p-q)\theta}\, d\theta$$

$$+ \frac{p}{4\pi i} \int\int_{-\pi}^{\pi} a(\theta)\, a(\phi)\, \epsilon(\theta - \phi)\, e^{i(p\theta - q\phi)}\, d\theta\, d\phi. \quad (9.22)$$

Expanding the determinant (9.21) along its principal diagonal and keeping terms up to the second order in $a(\theta)$, we obtain

$$H^2 = 1 + \sum_p r_{pp} + \sum_{p<q} \begin{vmatrix} r_{pp} & r_{pq} \\ r_{qp} & r_{qq} \end{vmatrix} + \cdots \tag{9.23}$$

or, for the logarithm of H,

$$\ln H = \frac{m}{\pi} \int a(\theta)\, d\theta + \frac{1}{2} \sum_p \frac{p}{4\pi i} \iint a(\theta)\, a(\phi)\, \epsilon(\theta - \phi)\, e^{ip(\theta-\phi)}\, d\theta\, d\phi$$

$$- \frac{1}{16\pi^2} \sum_{p,q} \left(1 + \frac{p}{q}\right)\left(1 + \frac{q}{p}\right) \iint a(\theta)\, a(\phi)\, e^{i(p-q)(\theta-\phi)}\, d\theta\, d\phi. \tag{9.24}$$

The coefficient of the first-order term m/π is just the level density. The second-order term gives the two-level cluster function

$$\begin{aligned} T_2(\theta, \phi) &= -\sum_p \frac{p}{4\pi i} \epsilon(\theta - \phi)\, e^{ip(\theta-\phi)} \\ &\quad + \frac{1}{8\pi^2} \sum_{p,q} \left(1 + \frac{p}{q}\right)\left(1 + \frac{q}{p}\right) e^{i(p-q)(\theta-\phi)} \\ &= \tfrac{1}{2}\epsilon(\theta - \phi)\, Ds_N(\theta - \phi) + [s_N(\theta - \phi)]^2 \\ &\quad - [Is_N(\theta - \phi)][Ds_N(\theta - \phi)], \end{aligned} \tag{9.25}$$

where we have written

$$s_N(\alpha) = (2\pi)^{-1} \sum_p e^{ip\alpha} = \frac{\sin m\alpha}{2\pi \sin(\alpha/2)}, \tag{9.26}$$

$$Df(\alpha) = \frac{d}{d\alpha} f(\alpha), \tag{9.27}$$

and

$$If(\alpha) = \int_0^\alpha f(\theta)\, d\theta. \tag{9.28}$$

In the limit $m \to \infty$, while $(m/\pi)\,\theta = \xi$ and $(m/\pi)\,\phi = \eta$ remain finite, the function $(2\pi/N)\, s_N(\theta - \phi)$ becomes identical to the function $s(r)$ of (5.64) and (5.65). The two-level cluster function $Y_2(r)$ and the form-factor $b(k)$ are therefore identical to those of Section 5.4.2.

To obtain the cluster functions for the alternate series, put $u(\theta) = 1, v(\theta) = 1 + a(\theta)$ in (9.6). The result is

$$H^2 = \left\langle \prod_{\text{alt}} (1 + a(\theta_j)) \right\rangle^2 = \det[\delta_{pq} + r'_{pq}], \tag{9.29}$$

with

$$r'_{pq} = \frac{1}{4\pi}\left(1 + \frac{p}{q}\right)\int_{-\pi}^{\pi} a(\theta)\, e^{i(p-q)\theta}\, d\theta. \tag{9.30}$$

The analysis proceeds as before, only the term in $\epsilon(\theta - \phi)$ is now missing from (9.25). The results are identical to those of Section 5.4.3.

9.1.2. The Distribution of Spacings

To determine $E_m(\alpha)$, the probability that a randomly chosen interval of length 2α is empty of energy levels, put

$$\begin{aligned} u(\theta) = v(\theta) &= 1, \quad \text{if} \quad -\pi + \alpha < \theta < \pi - \alpha, \\ &= 0, \quad \text{otherwise,} \end{aligned} \tag{9.31}$$

in (9.6). We have chosen the center of the excluded interval to be at $\theta = \pi$ so that (9.16) will be satisfied. From (9.18) we then have

$$E_m(\alpha) = \det[F_{pq}]_{p,q=1/2,3/2,\ldots,m-1/2}, \tag{9.32}$$

where

$$F_{pq} = \delta_{pq} - \frac{\sin(p-q)\,\alpha}{(p-q)\,\pi} - \frac{\sin(p+q)\,\alpha}{(p+q)\,\pi} \tag{9.33}$$

$$= \delta_{pq} - \frac{1}{\pi}\int_{-\alpha}^{\alpha} \cos p\theta \cos q\theta\, d\theta. \tag{9.34}$$

In the limit $m \to \infty$, while $(m/\pi)\,\alpha = \frac{1}{2}t$ remains finite, the determinant (9.32) becomes $E(t)$, the Fredholm determinant of the integral equation (5.99) with the kernel (5.98). This can be seen by putting $\xi = (m/\pi)\,p\alpha$, $\eta = (m/\pi)\,q\alpha$, and taking the limit $m \to \infty$. Thus the final results are identical with those of Section 5.5.

The question of the spacing distribution for the alternate series was not raised in Chapter 5. This will be done in Chapter 16. Let $E'_m(\alpha)$ denote the expression H resulting from the choice

$$\begin{aligned} &u(\theta) = v(\theta) = 1, \quad \text{if} \quad -\pi + \alpha < \theta < \pi - \alpha, \\ &u(\theta) = 0, v(\theta) = 2, \quad \text{if} \quad \pi - \alpha < \theta < \pi + \alpha, \end{aligned} \tag{9.35}$$

in (9.6); $E'_m(\alpha)$ is then the probability that a randomly chosen interval of length 2α will contain *at most one eigenvalue*. The choice $v(\theta) = 2$, rather than 1, in the interval $(\pi - \alpha, \pi + \alpha)$ arises from the fact that while ordering $-\pi \leqslant \theta_1 \leqslant \cdots \leqslant \theta_{2m} \leqslant \pi$ half the interval $(\pi - \alpha, \pi + \alpha)$ becomes unattainable for levels of the alternate series. Equations 9.18 and 9.19 then give

$$E'_m(\alpha) = \det[F_{pq}]_{p,q=1/2,3/2,\ldots,m-1/2}, \tag{9.36}$$

$$F_{pq} = \delta_{pq} - \frac{\sin(p-q)\alpha}{(p-q)\pi} + \frac{\sin(p+q)\alpha}{(p+q)\pi} \tag{9.37}$$

$$= \delta_{pq} - \frac{1}{\pi}\int_{-\alpha}^{\alpha} \sin p\theta \sin q\theta \, d\theta. \tag{9.38}$$

When $m \to \infty$, while $(m/\pi)\alpha = \frac{1}{2}t$ remains finite, the limit of $E'_m(\alpha)$ is $E'(t)$, the probability that an interval t will contain not more than one eigenvalue in a series with mean spacing $D = 1$. This limit is obtained exactly as in the preceding paragraph, the sole difference being we are now concerned with the odd solutions of the integral equation (5.101) or (5.102). The $E'(t)$ thus coincides with that given by (6.21) and (6.22). Note that although $E'(t)$ is the probability that a randomly chosen interval of length t will not contain any of the eigenvalues belonging to the same alternate series its second derivative is not a probability. The probability density for spacings between pairs of *next nearest* neighbors is given instead (cf. Appendix A.11) by

$$p'(t) = \frac{d^2}{dt^2}[E(t) + E'(t)], \tag{9.39}$$

where $E(t)$, the probability that the interval t will contain none of the eigenvalues, is given by (5.105) and (5.106) and $E'(t)$ is given by (6.21) and (6.22).

9.2. Symplectic Ensemble. $\beta = 4$[†]

The joint probability density function for the eigenvalues of a unitary self-dual random matrix taken from the symplectic ensemble was derived in Chapter 8 as

$$P_{N4}(\theta_1, \ldots, \theta_N) = C'_{N4} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^4. \tag{9.40}$$

[†] Mehta and Dyson [1].

To deal with integrals containing such an expression we write (9.40) as a confluent alternant. As in (9.2), we have

$$\prod_{j<k} | e^{i\theta_j} - e^{i\theta_k} |^4 = \exp\left[-2i(N-1)\sum_1^N \theta_j\right] \prod_{j<k} (e^{i\theta_j} - e^{i\theta_k})^4. \tag{9.41}$$

Note that because the power index 4 is an even integer, ordering of the angles is no longer necessary. The fourth power of the product of the differences expressed as a determinant (cf. Appendix A.16) is

$$\prod_{1\leqslant j<k\leqslant N} (e^{i\theta_j} - e^{i\theta_k})^4 = \det[e^{il\theta_j}, le^{i(l-1)\theta_j}]_{\substack{l=0,1,\ldots,2N-1 \\ j=1,2,\ldots,N}}. \tag{9.42}$$

If we multiply $(2j-1)$th and the $(2j)$th columns by $e^{-(1/2)(2N-1)i\theta_j}$ and $e^{-(1/2)(2N-3)i\theta_j}$, respectively, we obtain

$$\begin{aligned}\exp\left[-(2N-1)\,i\sum_1^N \theta_j\right] \prod_{j<k} (e^{i\theta_j} - e^{i\theta_k})^4 &= \det[e^{ip\theta_j}, (p+N-\tfrac{1}{2})\, e^{ip\theta_j}] \\ &= \det[e^{ip\theta_j}, pe^{ip\theta_j}],\end{aligned} \tag{9.43}$$

where p varies over the half odd integers

$$-(N-\tfrac{1}{2}), -(N-\tfrac{3}{2}),\ldots, (N-\tfrac{1}{2}). \tag{9.44}$$

We are now in a position to prove the main theorem of this section.

Theorem 9.1. *The statistical properties of N alternate angles θ_j, where $e^{i\theta_j}$ are the eigenvalues of a symmetric unitary random matrix of order $(2N \times 2N)$ taken from the orthogonal ensemble, are identical to those of the N angles ϕ_j, where $e^{i\phi_j}$ are the eigenvalues of an $N \times N$ quaternion self-dual unitary random matrix taken from the symplectic ensemble.*

Proof: Suppose that $\theta_1 \leqslant \theta_2 \leqslant \cdots \leqslant \theta_{2N} \leqslant \theta_1 + 2\pi$. We write the joint probability density function

$$\begin{aligned}P_{2N,1}(\theta_1,\ldots,\theta_{2N}) &= C'_{2N,1} \prod_{1\leqslant j<k\leqslant 2N} | e^{i\theta_j} - e^{i\theta_k} | \\ &= C'_{2N,1} i^{-N} \det[e^{ip\theta_j}],\end{aligned} \tag{9.45}$$

$$\begin{aligned}p &= -(N-\tfrac{1}{2}), -(N-\tfrac{3}{2}),\ldots, (N-\tfrac{1}{2}), \\ j &= 1, 2,\ldots, 2N,\end{aligned} \tag{9.46}$$

as in (9.4), and integrate over the variables $\theta_1, \theta_3, \ldots, \theta_{2N-1}$, as in Section 5.2. The limits of integration for θ_{2j-1} are $(\theta_{2j-2}, \theta_{2j})$, except when $j = 1$. For θ_1 these limits are $(\theta_{2N} - 2\pi, \theta_2)$. Thus the integration over the odd-indexed variables replaces the θ_1-column with

$$\int_{\theta_{2N}-2\pi}^{\theta_2} d\theta_1 \, e^{ip\theta_1} = (ip)^{-1}(e^{ip\theta_2} + e^{ip\theta_{2N}}) \tag{9.47}$$

and the θ_{2j-1} column, for $j > 1$, with

$$\int_{\theta_{2j-2}}^{\theta_{2j}} d\theta \, e^{ip\theta} = (ip)^{-1}(e^{ip\theta_{2j}} - e^{ip\theta_{2j-2}}). \tag{9.48}$$

This later column can be changed, as in Section 5.2, to

$$(ip)^{-1}\,(e^{ip\theta_{2j}} + e^{ip\theta_{2N}}). \tag{9.49}$$

The $(2N - 1)$th column is now simply $(ip)^{-1}\, 2e^{ip\theta_{2N}}$, which allows us to drop the $e^{ip\theta_{2N}}$ term from every other column. The final result is

$$\begin{aligned}
&\int\cdots\int_{\theta_1 \leqslant \theta_2 \leqslant \ldots \leqslant \theta_{2N} \leqslant \theta_1 + 2\pi} d\theta_1 \, d\theta_3 \cdots d\theta_{2N-1} \, P_{2N,1}(\theta_1, \ldots, \theta_{2N}) \\
&\quad = C'_{2N,1}\, i^{-N} \prod_p (ip)^{-1} \, 2\det[e^{ip\theta_{2j}}, ipe^{i\theta_{2j}}] \\
&\quad = \frac{C'_{2N.1}}{C'_{N4}} \, 2 \left(\frac{2^{2N} N!}{(2N)!}\right)^2 P_{N4}(\theta_2, \theta_4, \ldots, \theta_{2N}),
\end{aligned} \tag{9.50}$$

which establishes the theorem.

As an important corollary of Theorem 9.1, we state the following result.

Theorem 9.2. *The probability density function for the spacings of the eigenvalues in a self-dual, quaternion, unitary random matrix taken from the symplectic ensemble is given by*

$$p(t) = \frac{1}{2} \frac{d^2}{dt^2} [E(2t) + E'(2t)], \tag{9.51}$$

where $E(t)$ and $E'(t)$ are given by (5.105) *and* (6.21). *Here t is the spacing measured in units of the mean spacing.*

9.3. Unitary Ensemble. $\beta = 2$

Systems with no time reversal invariance may be characterized by unitary random matrices. The joint probability density function for the eigenvalue angles of such matrices taken from the unitary ensemble (cf. Chapter 8, Theorem 8.4) is

$$P_{N2}(\theta_1, ..., \theta_N) = C'_{N2} \prod_{j<k} | e^{i\theta_j} - e^{i\theta_k} |^2. \tag{9.52}$$

Calculating averages with the probability density P_{N2} is mathematically the simplest. Writing

$$\begin{aligned} \prod_{j<k} | e^{i\theta_j} - e^{i\theta_k} |^2 &= \prod_{j<k} (e^{i\theta_j} - e^{i\theta_k}) \prod_{j<k} (e^{-i\theta_j} - e^{-i\theta_k}) \\ &= \det[e^{il\theta_j}] \det[e^{-il\theta_j}] \end{aligned} \tag{9.53}$$

and remembering that

$$\int_0^{2\pi} e^{il\theta} e^{-il'\theta} \, d\theta = 2\pi \, \delta_{ll'} = \begin{cases} 2\pi, & \text{if } l = l', \\ 0, & \text{if } l \neq l', \end{cases} \tag{9.54}$$

we get, as in Section 6.1.1, the n-level correlation function R_n :

$$\begin{aligned} R_n^{(u)}(\theta_1, ..., \theta_n) &\equiv \frac{N!}{(N-n)!} \int_0^{2\pi} \cdots \int P_{N2}(\theta_1, ..., \theta_N) \, d\theta_{n+1} \cdots d\theta_N \\ &= \frac{(2\pi)^{-n}}{n!} \sum_{k_1, k_2, ..., k_n} (\det[e^{ik_l\theta_j}]_{l,j=1,2,...,n})^2, \qquad (9.55) \\ &= \det[K_N(\theta_j, \theta_k)]_{j,k=1,2,...,n}, \qquad (9.56) \end{aligned}$$

with

$$K_N(\theta, \phi) = \frac{1}{2\pi} \sum_{l=0}^{N-1} e^{il\theta} e^{-il\phi} = \frac{\sin[\frac{1}{2}N(\theta - \phi)]}{2\pi \sin[\frac{1}{2}(\theta - \phi)]}. \tag{9.57}$$

We have used Gramm's result in getting from (9.55) to (9.56) (cf. Appendix A.13).

Thus the level density is

$$K_N(\theta, \theta) = \frac{N}{2\pi}, \tag{9.58}$$

as it should be. The two-level correlation function is

$$R_2^{(u)}(\theta,\phi) = K_N(\theta,\theta)\,K_N(\phi,\phi) - K_N{}^2(\theta,\phi)$$

$$= \left(\frac{N}{2\pi}\right)^2 - \left[\frac{\sin[\frac{1}{2}N(\theta-\phi)]}{2\pi\sin[\frac{1}{2}(\theta-\phi)]}\right]^2, \tag{9.59}$$

and the two level cluster function is

$$T_{2u}(\theta,\phi) = \left[\frac{\sin[\frac{1}{2}N(\theta-\phi)]}{2\pi\sin[\frac{1}{2}(\theta-\phi)]}\right]^2. \tag{9.60}$$

On taking the limit $N \to \infty$, while keeping $N\theta/2\pi = \xi$ and $N\phi/2\pi = \eta$ fixed, we have

$$Y_{2u}(\xi,\eta) = \lim\left(\frac{2\pi}{N}\right)^2 T_{2u}(\theta,\phi) = \left(\frac{\sin\pi r}{\pi r}\right)^2, \tag{9.61}$$

where

$$r = |\,\xi - \eta\,|. \tag{9.62}$$

This result is identical to the two-point correlation function for the case of the unitary Gaussian ensemble (6.13).

For the spacing distribution we first find the probability $E_{uN}(\alpha)$ that a randomly chosen interval of length 2α will contain none of the angles θ_j :

$$E_{uN}(\alpha) = \int_{\alpha}^{2\pi-\alpha}\cdots\int P_{N2}(\theta_1,\ldots,\theta_N)\,d\theta_1\cdots d\theta_N$$

$$= \det\left[\frac{1}{2\pi}\int_{\alpha}^{2\pi-\alpha} d\theta\, e^{i(j-k)\theta}\right]_{j,k=1,2,\ldots,N} \tag{9.63}$$

$$= \det\left[\delta_{jk} - \frac{1}{2\pi}\int_{-\alpha}^{\alpha} d\theta\, e^{i(j-k)\theta}\right]_{j,k=1,2,\ldots,N}$$

$$= \det\left[\delta_{pq} - \frac{1}{2\pi}\int_{-\alpha}^{\alpha} d\theta\, e^{i(p-q)\theta}\right]_{p,q=-(1/2)(N-1),\ldots,(1/2)(N-1)} \tag{9.64}$$

$$= \det\left[\delta_{pq} - \frac{1}{2\pi}\int_{-\alpha}^{\alpha} d\theta\cos(p-q)\,\theta\right]_{p,q=-(1/2)(N-1),\ldots,(1/2)(N-1)}, \tag{9.65}$$

where we have used the result given in Appendix A.13 to arrive at (9.63). Since

$$\sum_q (\delta_{pq} - \cos p\theta\cos q\theta)(\delta_{qr} - \sin q\theta\sin r\theta) = \delta_{pr} - \cos(p-r)\,\theta, \tag{9.66}$$

we may factorize the expression (9.65):

$$E_{uN}(\alpha) = \det\left[\delta_{pq} - \frac{1}{2\pi}\int_{-\alpha}^{\alpha} \cos(p-q)\,\theta\, d\theta\right]$$
$$= E_m(\alpha)\, E'_m(\alpha) \tag{9.67}$$

with

$$E_m(\alpha) = \det\left[\delta_{pq} - \frac{1}{2\pi}\int_{-\alpha}^{\alpha} \cos p\theta \cos q\theta\, d\theta\right], \tag{9.68}$$

$$E'_m(\alpha) = \det\left[\delta_{pr} - \frac{1}{2\pi}\int_{-\alpha}^{\alpha} \sin p\theta \sin q\theta\, d\theta\right], \tag{9.69}$$

$$p, q = -(m - \tfrac{1}{2}), -(m - \tfrac{3}{2}), \ldots, (m - \tfrac{1}{2}), \tag{9.70}$$

where we have put $N = 2m$.

Taking the limit $m \to \infty$, while $m\alpha/\pi = t$ is kept fixed, is exactly the same as in Section 9.1.2 and the results are identically given by (6.20), (5.105), (6.21), (5.106), and (6.22).

9.4. Browning Motion Model†

Just as in Chapter 7, we can construct a Brownian motion model for the elements of our unitary matrices. Every matrix U taken from the ensemble $E_{\beta c}$ can be written as

$$U = VV^D \tag{9.71}$$

where V is unitary and V^D is the transpose or the dual (cf. Chapter 2 and Appendix A.23) of V, according as β is 1 or 4. For the unitary ensemble, $\beta = 2$, V^D is unrelated to V, except that V^D is unitary and (9.71) holds. A permissible small change in U is then given by

$$\delta U = V(i\,\delta M)\, V^D, \tag{9.72}$$

where δM is an infinitesimal Hermitian matrix which is symmetric and hence real if $\beta = 1$ and self-dual if $\beta = 4$. Let us denote the independent real components of δM by δM_μ; $\mu = 1, 2, \ldots, p$; $p = N + \frac{1}{2}N(N-1)\beta$. The isotropic and representation-independent Brownian motion of U is defined by the statement that U at time t moves to $U + \delta U$ at time $t + \delta t$, where δU is given by

† Dyson [4].

(9.72) and the real parameters δM_μ are independent random variables with the moments

$$\langle \delta M_\mu \rangle = 0, \tag{9.73}$$

$$f\langle (\delta M_\mu)^2 \rangle = g_\mu kT\, \delta t, \tag{9.74}$$

where g_μ is given by (7.16).

The effect of the Brownian motion of U on its eigenvalues $e^{i\theta_j}$ may again be found by choosing U to be diagonal at time t and calculating $e^{i(\theta_j+\delta\theta_j)}$ at time $t+\delta t$ by perturbation theory:

$$\delta\theta_j = \delta M_{jj}^{(0)} + \sum_{\substack{l \\ (l\neq j)}} \sum_{\lambda=0}^{\beta-1} (\delta M_{lj}^{(\lambda)})^2 \left[\tfrac{1}{2}\cot\left(\frac{\theta_j-\theta_l}{2}\right)\right] + \cdots \tag{9.75}$$

Equations 9.73 and 9.74 imply that the angles θ_j execute a Brownian motion with

$$f\langle \delta\theta_j \rangle = E(\theta_j)\, \delta t, \tag{9.76}$$

$$f\langle (\delta\theta_j)^2 \rangle = 2kT\, \delta t, \tag{9.77}$$

where

$$E(\theta_j) = \sum_{\substack{l \\ (l\neq j)}} \frac{1}{2}\cot\left(\frac{\theta_j-\theta_l}{2}\right). \tag{9.78}$$

This force $E(\theta_j)$ is exactly the component, tangential to the circle, of the electric field produced at $e^{i\theta_j}$ by unit charges placed at all the other points $e^{i\theta_l}$ at which U has eigenvalues. Thus

$$E(\theta_j) = -\frac{\partial W}{\partial \theta_j}, \tag{9.79}$$

$$W = -\sum_{l<j} \ln |\, e^{i\theta_l} - e^{i\theta_j}\,|. \tag{9.80}$$

One may write the corresponding Focker-Planck equation and find that the unique stationary probability density for the eigenvalue angles which corresponds to the completely diffused probability density of U is

$$P(\theta_1, \ldots, \theta_N) = c' \prod_{l<j} |\, e^{i\theta_l} - e^{i\theta_j}\,|^\beta. \tag{9.81}$$

This is again a new proof of Theorem 8.4.

10 / Circular Ensembles. Thermodynamics†

10.1. General Remarks

In Chapter 9 we studied the correlation functions, cluster functions, and spacing distribution of the eigenvalues of a unitary matrix taken from Dyson's circular ensembles. In this chapter we calculate the partition function, the energy, the free energy, the entropy, and the specific heat of the energy levels defined in complete analogy with the classical mechanics. Later we apply thermodynamic ideas to "derive" the asymptotic behavior of the spacing distribution for large spacings. As Dyson [2] remarks, such a "derivation" is useful in two ways. As long as no direct analytical method is known to deal with this problem, the thermodynamic argument is the only one at our disposal. If, later, an analytical method is found, these results will serve, by way of comparison, the purpose of demarking the region in which such arguments are reliable.

10.2. The Partition Function

Analogous to Chapter 4, consider a thin circular conducting wire of radius unity; let N point charges be free to move on this wire. The universe is supposed to be two-dimensional. The charges repel one another according to the two-dimensional Coulomb law, so that the potential energy due to this electrostatic repulsion is

$$W = -\sum_{1 \leqslant l < j \leqslant N} \ln | e^{i\theta_l} - e^{i\theta_j} | . \tag{10.1}$$

As in Chapter 4, we discard the trivial velocity-dependent contributions. The positional partition function at a temperature kT is given by

$$\Psi_N(\beta) = (2\pi)^{-N} \int_0^{2\pi} \cdots \int e^{-\beta W} \, d\theta_1 \cdots d\theta_N , \tag{10.2}$$

$$\beta = (kT)^{-1}, \tag{10.3}$$

† Dyson [2].

where W is given by (10.1). The aim of this section is to prove the following theorem.

Theorem 10.1. *For any positive integer N and a real or complex β the $\Psi_N(\beta)$ is given by*

$$\Psi_N(\beta) = \Gamma(1 + \tfrac{1}{2}\beta N)[\Gamma(1 + \tfrac{1}{2}\beta)]^{-N}. \tag{10.4}$$

The proof may be divided into three parts:

1. If (10.4) holds for all even integers β, it will hold for any complex β.

The argument depends on the fact that W is bounded from below. It is intuitively clear (for a proof see Appendix A.21) that this minimum is attained when the N charges are situated at the corners of a regular polygon of N sides inscribed in the unit circle. The value of the minimum is easily calculated and gives

$$W \geqslant W_0 = -\tfrac{1}{2}N \ln N. \tag{10.5}$$

As in Chapter 4, we can now express the partition function as a moment function and deduce its analytic and asymptotic properties [Shohat and Tamarkin, 1]. Denoting the expression (10.4) by $\psi_N(\beta)$ we can then use Carlson's theorem [Tischmarsh, 1] for the function

$$\Delta(\beta) = e^{2\beta W_0}[\Psi_N(2\beta) - \psi_N(2\beta)] \tag{10.6}$$

to conclude that if the expressions (10.2) and (10.4) are equal for all even integers β they are identically equal.

2. The integral

$$(2\pi)^{-N} \int_0^{2\pi} \cdots \int \prod_{l<j} | e^{i\theta_l} - e^{i\theta_j} |^{2k} \, d\theta_1 \cdots d\theta_N \tag{10.7}$$

is equal to the constant term in the series expansion, involving positive as well as negative powers of z_j ; $j = 1, 2, \ldots, N$, of the product

$$\prod_{\substack{l,j \\ (l \neq j)}} \left(1 - \frac{z_l}{z_j}\right)^k. \tag{10.8}$$

To see this let $z_j = \exp(i\theta_j)$, so that

$$| z_j - z_l |^{2k} = (z_j - z_l)^k (z_j^{-1} - z_l^{-1})^k = \left(1 - \frac{z_l}{z_j}\right)^k \left(1 - \frac{z_j}{z_l}\right)^k \tag{10.9}$$

and note that any power other than zero of z_j ; $j = 1, 2,..., N$, vanishes on integration.

3. The constant term in the expansion of

$$\prod_{\substack{l,j \\ (l \neq j)}} \left(1 - \frac{z_j}{z_l}\right)^{a_j} \tag{10.10}$$

is given by

$$\frac{(a_1 + \cdots + a_N)!}{a_1! \cdots a_N!}. \tag{10.11}$$

A proof is reproduced in Appendix A.19 [Wilson, 1].

If we put $a_1 = a_2 = \cdots = a_N = k$ in (10.10) and (10.11), we get the constant term in the expansion of (10.8) as

$$(Nk)!\,(k!)^{-N} = \Gamma(1 + kN)\{\Gamma(1 + k)\}^{-N}. \tag{10.12}$$

The proof of Theorem 10.1 is now complete.

10.3. Thermodynamic Quantities

Theorem 10.1 specifies completely the thermodynamic properties of a finite Coulomb gas of N charges on the unit circle. For applications to the energy level series we are interested only in the special case of a very large N; $N \to \infty$. In this section we study the statistical mechanics of an infinite Coulomb gas or, equivalently, that of an infinitely long series of eigenvalues.

The partition function (10.2) is normalized in a way that the energy of the gas is zero at infinite temperature ($\beta = 0$). The potential energy at zero temperature is then the ground-state energy.

$$W_0 = -\tfrac{1}{2}N \ln N. \tag{10.13}$$

To obtain finite limits for the thermodynamic variables as $N \to \infty$ we must first change the zero of the energy to the position W_0. By definition the gas then has zero energy at zero temperature and positive energy at any positive temperature. The partition function defined on the new energy scale is

$$\Phi_N(\beta) = (2\pi)^{-N} \int_0^{2\pi} \cdots \int e^{-\beta(W - W_0)}\, d\theta_1 \cdots d\theta_N. \tag{10.14}$$

The free energy per particle $F_N(\beta)$ is

$$F_N(\beta) = -(\beta N)^{-1} \ln \Phi_N(\beta) \tag{10.15}$$

$$= \tfrac{1}{2} \ln N - (\beta N)^{-1} \ln \Gamma(1 + \tfrac{1}{2}\beta N) + \beta^{-1} \ln \Gamma(1 + \tfrac{1}{2}\beta), \tag{10.16}$$

where we have used Theorem 10.1. Taking the limit $N \to \infty$, we obtain the following theorem:

Theorem 10.2. *As $N \to \infty$ the free energy per particle of the Coulomb gas at temperature $kT = \beta^{-1}$ tends to the limiting value*

$$F(\beta) = \beta^{-1} L(\tfrac{1}{2}\beta) + \tfrac{1}{2}[1 - \ln(\tfrac{1}{2}\beta)], \tag{10.17}$$

$$L(z) = \ln \Gamma(1 + z). \tag{10.18}$$

The values of the other thermodynamic quantities follow from (10.17).

Energy per particle:

$$U(\beta) = F + \beta \frac{\partial F}{\partial \beta} = \tfrac{1}{2}[L'(\tfrac{1}{2}\beta) - \ln(\tfrac{1}{2}\beta)]. \tag{10.19}$$

Entropy per particle:

$$S(\beta) = \beta^2 \frac{\partial F}{\partial \beta} = \tfrac{1}{2}\beta[L'(\tfrac{1}{2}\beta) - 1] - L(\tfrac{1}{2}\beta). \tag{10.20}$$

Specific heat per particle:

$$C(\beta) = -\beta^2 \frac{\partial U}{\partial \beta} = -\tfrac{1}{4}\beta^2 L''(\tfrac{1}{2}\beta) + \tfrac{1}{2}\beta. \tag{10.21}$$

To calculate the values of these thermodynamic quantities for physically interesting values of β the following formulas [Bateman, 3] may be used:

$$L(z) = -\gamma z + \sum_2^\infty (-1)^n \frac{S_n}{n} z^n, \qquad |z| \leqslant 1, \tag{10.22}$$

$$L(z) = z \ln z - z + \tfrac{1}{2} \ln z + \tfrac{1}{2} \ln(2\pi) + \frac{1}{12z} + 0\left(\frac{1}{z^2}\right), \quad |z| \to \infty, \tag{10.23}$$

$$L(z + n) - L(z) = \sum_{r=1}^{n} \ln(z + r), \tag{10.24}$$

$$L''(z) = \sum_{n=1}^{\infty} (z + n)^{-2}, \tag{10.25}$$

and

$$L'(\tfrac{1}{2}) = 2 - \gamma - 2 \ln 2, \tag{10.26}$$

where γ is Euler's constant [Bateman, 4]

$$\gamma = 0.5772 \cdots \tag{10.27}$$

and S_k are the sums of the inverse powers of the integers

$$S_k = \sum_{n=1}^{\infty} n^{-k}. \tag{10.28}$$

In particular,

$$S_2 = \sum_{1}^{\infty} n^{-2} = \frac{\pi^2}{6} \tag{10.29}$$

and

$$\sum_{1}^{\infty} (2n - 1)^{-2} = \frac{\pi^2}{8}. \tag{10.30}$$

Table 10.1 summarizes this calculation.

10.4. Statistical Interpretation of U and C

If we denote the ensemble average by $\langle\ \rangle$,

$$\langle f \rangle = \frac{\int_0^{2\pi} \cdots \int f(\theta_1, \ldots, \theta_N)\, e^{-\beta(W - W_0)}\, d\theta_1 \cdots d\theta_N}{\int_0^{2\pi} \cdots \int e^{-\beta(W - W_0)}\, d\theta_1 \cdots d\theta_N}, \tag{10.31}$$

then from (10.14), (10.15), (10.19), and (10.21) we have

$$\langle W - W_0 \rangle = -[\Phi_N(\beta)]^{-1} \frac{\partial}{\partial \beta} \Phi_N(\beta) = NU \tag{10.32}$$

and

$$\begin{aligned} \langle (W - \langle W \rangle)^2 \rangle &= \langle (W - W_0)^2 \rangle - (\langle W - W_0 \rangle)^2 \\ &= [\Phi_N(\beta)]^{-1} \frac{\partial^2}{\partial \beta^2} \Phi_N(\beta) - (NU)^2 \\ &= \beta^{-2} NC, \end{aligned} \tag{10.33}$$

TABLE 10.1

	$\beta \to 0$	$\beta = 1$	$\beta = 2$	$\beta = 4$	$\beta \to \infty$
$L(\frac{1}{2}\beta)$	$-\frac{1}{2}\gamma\beta + \frac{\pi^2}{48}\beta^2 + \cdots$	$\frac{1}{2}\ln\left(\frac{\pi}{4}\right)$ ≈ -0.121	0	$\ln 2$ ≈ 0.693	$\frac{1}{2}\beta\ln(\frac{1}{2}\beta) - \frac{1}{2}\beta + \frac{1}{2}\ln(\pi\beta) + \frac{1}{6\beta} + \cdots$
$L'(\frac{1}{2}\beta)$	$-\gamma + \frac{\pi^2}{12}\beta + \cdots$	$2 - \gamma - \ln 4$ ≈ 0.035	$1 - \gamma$ ≈ 0.423	$\frac{3}{2} - \gamma$ ≈ 0.923	$\ln(\frac{1}{2}\beta) + \frac{1}{\beta} - \frac{1}{3\beta^2} + \cdots$
$L''(\frac{1}{2}\beta)$	$\frac{\pi^2}{6} + \cdots$	$\frac{\pi^2}{4} - 4$ ≈ -1.532	$\frac{\pi^2}{6} - 1$ ≈ 0.645	$\frac{\pi^2}{6} - \frac{5}{4}$ ≈ 0.395	$\frac{2}{\beta} - \frac{2}{\beta^2} + \frac{4}{3\beta^2} + \cdots$
F	$\frac{1}{2}(1 - \gamma - \ln(\frac{1}{2}\beta)) + \frac{\pi^2}{48}\beta + \cdots$	$\frac{1}{2}\left(1 + \ln\frac{\pi}{2}\right)$ ≈ 0.726	$\frac{1}{2}$ $= 0.500$	$\frac{1}{2} - \frac{1}{4}\ln 2$ ≈ 0.327	$\frac{1}{2\beta}\ln(\pi\beta) + \frac{1}{6\beta^2} + \cdots$
U	$\frac{1}{2}(-\gamma - \ln(\frac{1}{2}\beta)) + \frac{\pi^2}{24}\beta + \cdots$	$1 - \frac{1}{2}(\gamma + \ln 2)$ ≈ 0.365	$\frac{1}{2}(1 - \gamma)$ ≈ 0.211	$\frac{3}{4} - \frac{1}{2}(\gamma + \ln 2)$ ≈ 0.115	$\frac{1}{2\beta} - \frac{1}{6\beta^2} + \cdots$
S	$-\frac{1}{2}\beta + \frac{\pi^2}{48}\beta^2 + \cdots$	$\frac{1}{2}(1 - \gamma - \ln\pi)$ ≈ -0.361	$-\gamma$ ≈ -0.577	$1 - 2\gamma - \ln 2$ ≈ -0.848	$\frac{1}{2}(1 - \ln(\pi\beta)) - \frac{1}{3\beta} + \cdots$
C	$\frac{1}{2}\beta - \frac{\pi^2}{24}\beta^2 + \cdots$	$\frac{3}{2} - \frac{\pi^2}{8}$ ≈ 0.266	$2 - \frac{\pi^2}{6}$ ≈ 0.355	$7 - \frac{2}{3}\pi^2$ ≈ 0.420	$\frac{1}{2} - \frac{1}{3\beta} + \cdots$

where W is the electrostatic energy given by (10.1) and $W_0 = -\frac{1}{2}N \ln N$ is the minimum value of W when the charges are uniformly spaced. Thus U is, apart from normalization, the ensembles average of the logarithm of the geometric mean of all distances between pairs of eigenvalues, and $\beta^{-2}C$ is the statistical mean square fluctuation of the same quantity.

For analyzing the properties of observed eigenvalue series, W seems to be a good statistic. It has two great advantages over the other statistics such as F and S:

1. W can be computed from the eigenvalue pair-correlation function alone without analyzing higher order correlations.

2. The statistical uncertainty of W is known from the value of C.

We summarize the situation in the following theorem.

Theorem 10.3. *Let $z_1, z_2, \ldots, z_N$ be the eigenvalues of a random unitary matrix taken from one of the ensembles E_1, E_2, or E_4. The statistic*

$$W - W_0 = \tfrac{1}{2}N \ln N - \sum_{i<j} \ln | z_i - z_j | \tag{10.34}$$

has the average value NU and the root mean square deviation $\beta^{-1}(NC)^{1/2}$ with the values of U and C listed in Table 10.1.

10.5. Continuum Model for the Spacing Distribution

In this section we exploit an argument of classical statistical mechanics to arrive at the asymptotic form of spacing distribution for large spacings.

As in Section 10.2, we write the joint probability density function for the eigenvalues $e^{i\theta_j}$; $j = 1, 2, \ldots, N$, of a random unitary $N \times N$ matrix as

$$P_{N\beta}(\theta_1, \ldots, \theta_N) = C'_{N\beta} e^{-\beta W} \tag{10.35}$$

with

$$W = -\sum_{j<k} \ln | e^{i\theta_j} - e^{i\theta_k} |. \tag{10.36}$$

The probability that an arc of length 2α will contain none of the angles θ_j, that is, the $E_m(\alpha)$ of (9.32), is given by

$$E(\beta, \alpha) = \Psi_\beta(\alpha)/\Psi_\beta(0), \tag{10.37}$$

where

$$\Psi_\beta(\alpha) = \int_\alpha^{2\pi-\alpha} \cdots \int e^{-\beta W}\, d\theta_1 \cdots d\theta_N \tag{10.38}$$

is the partition function of the analogous Coulomb gas of N charges compressed in a circular arc of length $2\pi - 2\alpha$, whereas $\Psi_\beta(0)$ is the partition function of the same gas on the whole unit circle. We write

$$E(\beta, \alpha) = \exp\{-\beta[F_N(\beta, \alpha) - F_N(\beta, 0)]\}, \tag{10.39}$$

where $F_N(\beta, \alpha)$ is the free energy of the Coulomb gas on the arc $2\pi - 2\alpha$.

The hypothesis that for large N the Coulomb gas forms a continuous electric fluid obeying the laws of thermodynamics may be put in the form of the following three assumptions:

1. There is a macroscopic charge density; that is there exists a smooth function $\sigma_\alpha(\theta)$ such that the average number of charges on the arc $(\theta, \theta + d\theta)$ is $\sigma_\alpha(\theta)\, d\theta$.

2. For a given density $\sigma_\alpha(\theta)$ the free energy is the sum of the two terms

$$F = V_1 + V_2\,, \tag{10.40}$$

where V_1 is the macroscopic potential energy

$$V_1 = -\frac{1}{2}\int_\alpha^{2\pi-\alpha}\int \sigma_\alpha(\theta)\, \sigma_\alpha(\phi) \ln |\, e^{i\theta} - e^{i\phi}\,|\, d\theta\, d\phi \tag{10.41}$$

and V_2 is the contribution from the individual arcs, depending only on the local density

$$V_2 = \int_\alpha^{2\pi-\alpha} \sigma_\alpha(\theta) f_\beta[\sigma_\alpha(\theta)]\, d\theta, \tag{10.42}$$

$f_\beta[\sigma]$ being the free energy per particle of a Coulomb gas having uniform density σ on the whole unit circle. The factor $\frac{1}{2}$ in (10.41) is there because the interaction between two arc elements is counted twice.

3. In almost all cases the density σ_α adjusts itself in such a way that the free energy $F_N(\beta, \alpha)$ given by (10.40), (10.41), and (10.42) is a minimum, subject to the condition

$$\int_\alpha^{2\pi-\alpha} \sigma_\alpha(\theta)\, d\theta = N. \tag{10.43}$$

This last equation expresses the fact that the total number of charges is fixed to N.

There is no rigorous mathematical justification for the above assumptions. But they are so much accepted by tradition that we make no apologies for adopting them.

The functional $f_\beta(\sigma)$ remains to be specified, and we write

$$f_\beta(\sigma) = U_\beta(\sigma) - \beta^{-1} S_\beta(\sigma), \tag{10.44}$$

where $U_\beta(\sigma)$ is the energy and $S_\beta(\sigma)$ is the entropy per particle for a uniform gas of

$$N' = 2\pi\sigma \tag{10.45}$$

charges on the whole unit circle.

As in (10.19), the energy per particle is

$$U_\beta(\sigma) = -\tfrac{1}{2} \ln N' + U(\beta). \tag{10.46}$$

The term $-\frac{1}{2} \ln N'$ is included, for we now have to take the total energy, including the ground-state energy $-\frac{1}{2}N' \ln N'$.

The entropy $S_\beta(\sigma)$, if calculated as in Section 10.3, is independent of N' for large N'. However, one thing should be noted. The calculation of the entropy in Section 10.3 was made for a gas of N distinguishable particles. The entropy so defined is not an extensive quantity. To make it extensive we must subtract $N!$ from the classical entropy, which amounts to treating the particles as undistinguishable. As we need the $S_\beta(\sigma)$ in (10. 44) to be an extensive quantity, we write

$$S_\beta(\sigma) = \ln \left(\frac{N}{N'}\right) + S(\beta), \tag{10.47}$$

where $S(\beta)$ is given by (10.20).

Putting

$$\sigma_\alpha(\theta) = \frac{N}{2\pi} \rho_\alpha(\theta) \tag{10.48}$$

and collecting (10.40) to (10.47), we have

$$\beta F_N(\beta, \alpha) = G_2 + G_1 + G_0\,, \tag{10.49}$$

where

$$G_2 = -\tfrac{1}{2}\beta\left(\frac{N}{2\pi}\right)^2 \int_{\alpha}^{2\pi-\alpha}\int \rho_\alpha(\theta)\,\rho_\alpha(\phi)\ln|e^{i\theta} - e^{i\phi}|\,d\theta\,d\phi, \tag{10.50}$$

$$G_1 = (1 - \tfrac{1}{2}\beta)\left(\frac{N}{2\pi}\right)\int_{\alpha}^{2\pi-\alpha} \rho_\alpha(\theta)\ln\rho_\alpha(\theta)\,d\theta, \tag{10.51}$$

$$G_0 = \beta N[F(\beta) - \tfrac{1}{2}\ln N]; \tag{10.52}$$

and $F(\beta)$ given by (10.17). One has to minimize the quantity (10.49) under the restriction

$$\int_{\alpha}^{2\pi-\alpha} \rho_\alpha(\theta)\,d\theta = 2\pi. \tag{10.53}$$

When $\alpha = 0$, the equilibrium density $\rho_\alpha(\theta) = 1$, and $G_2 = G_1 = 0$, so that

$$\beta F_N(\beta, 0) = G_0 = \beta N[F(\beta) - \tfrac{1}{2}\ln N] \tag{10.54}$$

and from (10.39)

$$E(\beta, \alpha) = \exp[-\min_\rho (G_2 + G_1)]. \tag{10.55}$$

Using Lagrange's method to minimize $(G_2 + G_1)$ under the restriction (10.53), we get for $\beta \neq 2$

$$-\beta\left(\frac{N}{2\pi}\right)^2 \int_{\alpha}^{2\pi-\alpha} \rho_\alpha(\phi)\ln|e^{i\theta} - e^{i\phi}|\,d\phi + \left(\frac{N}{2\pi}\right)(1 - \tfrac{1}{2}\beta)\ln\rho_\alpha(\theta)$$

$$+ \left(\frac{N}{2\pi}\right)(1 - \tfrac{1}{2}\beta) - \lambda = 0, \qquad \alpha < \theta < 2\pi - \alpha, \tag{10.56}$$

where λ is the undetermined constant. Letting

$$V_\alpha(\theta) = -\int_{\alpha}^{2\pi-\alpha} \rho_\alpha(\phi)\ln|e^{i\theta} - e^{i\phi}|\,d\phi, \tag{10.57}$$

we have

$$\beta\left(\frac{N}{2\pi}\right)(1 - \tfrac{1}{2}\beta)^{-1}\,V_\alpha(\theta) + \ln\rho_\alpha(\theta) = \text{constant}$$

or

$$\rho_\alpha(\theta) = Ae^{-\gamma V_\alpha(\theta)}, \tag{10.58}$$

$$\gamma = \beta\left(\frac{N}{2\pi}\right)(1 - \tfrac{1}{2}\beta)^{-1}. \tag{10.59}$$

When $\beta = 2$, G_1 is zero, and the minimization leads to

$$V_\alpha(\theta) = -\int_\alpha^{2\pi-\alpha} \rho_\alpha(\phi) \ln | e^{i\theta} - e^{i\phi} | \, d\phi = \text{constant}. \tag{10.60}$$

The $V_\alpha(\theta)$ is the electrostatic potential at the angle θ produced by all the other charges. If $\beta \neq 2$, these charges are in thermal equilibrium at an effective temperature

$$kT_e = \frac{1}{\beta}(1 - \tfrac{1}{2}\beta) = \frac{1}{\beta} - \tfrac{1}{2} \tag{10.61}$$

under the potential $V_\alpha(\theta)$ generated by themselves. If $\beta = 2$, the potential $V_\alpha(\theta)$ is constant and the charges are in electrostatic equilibrium on a conducting circular arc of length $2\pi - 2\alpha$.

The problem in the case of unitary ensemble $\beta = 2$ is the easiest to handle. The classical problem of charge distribution on a slotted conducting cylinder is well known. Here we give only the result [Smythe, 1]. The solution of (10.53) and (10.60) is

$$\rho_\alpha(\theta) = \sin\frac{\theta}{2}\left(\sin^2\frac{\theta}{2} - \sin^2\frac{\alpha}{2}\right)^{-1/2} \tag{10.62}$$

$$V_\alpha = 2\pi \ln\left(\cos\frac{\alpha}{2}\right). \tag{10.63}$$

Equation 10.55 therefore gives

$$\ln E(2, \alpha) = -\min_\rho G_2 = N^2 \ln\left(\cos\frac{\alpha}{2}\right). \tag{10.64}$$

In the limit $N \to \infty$, $t = 2\alpha N/2\pi$ finite, we obtain

$$\begin{aligned} E(2, t) &= \lim \exp\left(N^2 \ln\cos\frac{\alpha}{2}\right) \\ &= \lim \exp(-\tfrac{1}{8}N^2\alpha^2) = \exp\left(-\frac{\pi^2}{8}t^2\right) \end{aligned} \tag{10.65}$$

and the probability density function of the spacings

$$p(t) = \frac{d^2E(2, t)}{dt^2} \approx \frac{\pi^4}{16}t^2 \exp\left(-\frac{\pi^2}{8}t^2\right), \qquad t \gg 1. \tag{10.66}$$

The case $\beta \neq 2$ is more difficult. Dyson has applied perturbation theory to expand F in inverse powers of αN. Since G_2 is of the

order $(\alpha N)^2$ and G_1 is of the order (αN), he treats G_1 as a "small perturbation" over G_2. The tedious details of this calculation may be found in the original paper [Dyson, 2]. The results are as follows:

The first- and second-order contributions to free energy in the limit of large N and large $t = \alpha N/\pi$ are

$$\beta F_1 = (1 - \tfrac{1}{2}\beta) N \ln \left(\sec \frac{\alpha}{2} + \tan \frac{\alpha}{2}\right) \tag{10.67}$$

$$\approx (1 - \tfrac{1}{2}\beta) \frac{\pi}{2} t \tag{10.68}$$

$$\beta F_2 \approx -\frac{1}{2\beta}(1 - \tfrac{1}{2}\beta)^2 \{\ln[\tfrac{1}{2}\beta t(1 - \tfrac{1}{2}\beta)^{-1}] + \gamma\}, \tag{10.69}$$

where γ is Euler's constant. Equations 10.55 and 5.84 then give the results

$$E(\beta, t) \simeq A t^{f(\beta)} \exp\left[-\frac{\pi^2}{16}\beta t^2 - (1 - \tfrac{1}{2}\beta)\frac{\pi}{2} t\right] \tag{10.70}$$

and the probability density for the spacings

$$p_\beta(t) \approx A t^{2+f(\beta)} \exp\left[-\frac{\pi^2}{16}\beta t^2 - (1 - \tfrac{1}{2}\beta)\frac{\pi}{2} t\right], \tag{10.71}$$

where

$$f(\beta) = \frac{1}{2\beta}(1 + \tfrac{1}{2}\beta)^2 \tag{10.72}$$

and A is a constant, which cannot, in principle, be determined from thermodynamic arguments.

11 / The Orthogonal Circular Ensemble. Wigner's Method

11.1. General Remarks

In this chapter we develop a method of calculating the spacing distribution in the orthogonal circular ensemble, due originally to Wigner. Though this method could not be adapted to numerical computation when the order of the matrices becomes large, it gave the normalization constant and the slope of the probability density of the spacings at zero spacing. The correct slope turned out to be $\pi^2/6$ rather than $\pi/2$, as given by the "Wigner surmise,"

$$p_W(t) = \frac{\pi}{2} t e^{-(\pi/4)t^2}, \qquad t = \frac{S}{D}. \tag{11.1}$$

This was how for the first time (11.1) was put in some "disrepute" by Wigner himself. The calculation was never published. Moreover, it was believed, incorrectly, to have been a calculation of the mean square spacing. This same slope was later calculated by Mehta [2, Section 7] and Dyson [6], using different methods (cf. Chapters 5 and 9).

11.2. The Method of Integration†

We start by proving a few lemmas. At first glance they may look unrelated, but their relevance to the problem will become clearer as we proceed.

Lemma 11.1. *Let* $\Delta(\nu_a, \nu_b, \nu_c, \ldots)$ *be defined by*

$$\Delta(\nu_a, \nu_b, \nu_c, \ldots) = [\nu_a(\nu_a + \nu_b)(\nu_a + \nu_b + \nu_c) \cdots]^{-1}. \tag{11.2}$$

† Wigner [7].

We then have the identity

$$\sum_{k=0}^{n} (-1)^k \Delta(\nu_k , \nu_{k-1} ,..., \nu_1) \Delta(\nu_{k+1} ,..., \nu_n) = 0. \tag{11.3}$$

It is assumed here and in the following that whenever the corresponding index does not exist the relevant factor is unity. For example, in (11.3) the $k = 0$ term is $\Delta(\nu_1 , \nu_2 ,..., \nu_n)$ and the $k = n$ term is $(-1)^n \Delta(\nu_n , \nu_{n-1} ,..., \nu_1)$.

Proof. Consider the left-hand side as a function of, say, ν_1. Its possible singularities are at the points $\nu_1 = -(\nu_2 + \cdots + \nu_l)$ and at $\nu_1 = 0$. At $\nu_1 = -(\nu_2 + \cdots + \nu_l)$ two terms in the sum become singular, $k = 0$ and $k = l$:

$$\begin{aligned}
&\Delta(\nu_1 , \nu_2 ,..., \nu_l ,..., \nu_n) + (-1)^l \Delta(\nu_l ,..., \nu_1) \Delta(\nu_{l+1} ,..., \nu_n) \\
&\quad = [\nu_1(\nu_1 + \nu_2) \cdots (\nu_1 + \cdots + \nu_{l-1})(s_l + \nu_{l+1}) \cdots (s_l + \nu_{l+1} + \cdots + \nu_n) s_l]^{-1} \\
&\qquad + (-1)^l [\nu_l(\nu_l + \nu_{l-1}) \cdots (\nu_l + \cdots + \nu_2) \nu_{l+1}(\nu_{l+1} + \nu_{l+2}) \cdots \\
&\qquad\quad (\nu_{l+1} + \cdots + \nu_n) s_l]^{-1},
\end{aligned} \tag{11.4}$$

where $s_l = \nu_1 + \nu_2 + \cdots + \nu_l$. The coefficients of s_l^{-1} in the two terms are equal and opposite at $s_l = 0$. To verify this, replace ν_l in the second term by $-(\nu_1 + \cdots + \nu_{l-1} - s_l)$, $(\nu_l + \nu_{l-1})$ by $-(\nu_1 + \cdots + \nu_{l-2} - s_l)$,..., and $(\nu_l + \nu_{l-1} + \cdots + \nu_2)$ by $-(\nu_1 - s_l)$; so that at $s_l = 0$ factors in the two terms may be compared.

This holds for all l, including $l = 1$; that is, the point $\nu_1 = 0$. The original function of ν_1 is therefore nowhere singular and so must be a polynomial in ν_1. Because every term in the sum (11.4) contains at least one linear factor in ν_1 in the denominator, the polynomial mentioned above must be identically zero.

Lemma 11.2. *Let a be the antisymmetrization operator, $a = \sum_P \epsilon_P P$, where P is a permutation of the variables $\nu_1 , \nu_2 ,..., \nu_N$ and ϵ_P is its parity. We have the identity*

$$\begin{aligned}
(-1)^N a \, \Delta(\nu_1 , \nu_2 ,..., \nu_N) &= (\nu_1 + \nu_2 + \cdots + \nu_N)^{-1} \\
&\quad \times \sum_{k=1}^{N} (-1)^k a \, \Delta(\nu_1 ,..., \nu_{k-1} , \nu_{k+1} ,..., \nu_N).
\end{aligned} \tag{11.5}$$

Proof: The right-hand side is antisymmetric in all of the variables $\nu_1, \nu_2, \ldots, \nu_N$ and contains $N(N-1)! = N!$ terms. It also contains

$$\Delta(\nu_1, \ldots, \nu_N) = (\nu_1 + \cdots + \nu_N)^{-1} \Delta(\nu_1, \ldots, \nu_{N-1})$$

with the coefficient $(-1)^N$. Therefore the two sides of (11.5) are equal.

Lemma 11.3. *The antisymmetrized* $\Delta(\nu_1, \nu_2, \ldots, \nu_N)$ *is*

$$a\,\Delta(\nu_1, \nu_2, \ldots, \nu_N) = \prod_{k=1}^{N} \nu_k^{-1} \prod_{1 \leqslant j < k \leqslant N} [(\nu_k - \nu_j)(\nu_k + \nu_j)^{-1}]. \tag{11.6}$$

We prove (11.6) by induction; let it then be valid for $N-1$. Using (11.6) on the right-hand side of the identity (11.5) and reducing to a common denominator, we note that the degree of this common denominator is $1 + N + \frac{1}{2}N(N-1)$, and by counting factors in each term of the sum we find that the degree of the common numerator is $1 + \frac{1}{2}N(N-1)$. Because the right-hand side of (11.5) is antisymmetric, it vanishes if $\nu_j = \nu_k$; the numerator therefore must contain a factor $(\nu_j - \nu_k)$. On taking out all such factors that are $\frac{1}{2}N(N-1)$ in number, we are left with a linear factor symmetric in all the $\nu_1, \nu_2, \ldots, \nu_N$. This linear factor can only be $(\nu_1 + \nu_2 + \cdots + \nu_N)$.

A constant multiplier is still undecided. For this we suppose that $\nu_1 \ll \nu_2 \ll \cdots \ll \nu_N$ and compare the dominant term on each side of (11.6). This term is $(\nu_1 \cdot \nu_2 \cdot \cdots \cdot \nu_N)^{-1}$. Finally, (11.6) is true for $N = 2$, as can be easily verified.

Lemma 11.4. *The integral*

$$\int_{\phi_1}^{\phi_2} d\theta_N \int_{\phi_1}^{\theta_N} d\theta_{N-1} \cdots \int_{\phi_1}^{\theta_2} d\theta_1 \exp[i(\nu_1\theta_1 + \cdots + \nu_N\theta_N)]$$

$$\equiv \int\cdots\int_{\substack{\phi_1 \leqslant \theta_1 \leqslant \theta_2 \leqslant \cdots \\ \leqslant \theta_N \leqslant \phi_2}} d\theta_1 \cdots d\theta_N \exp[i(\nu_1\theta_1 + \cdots + \nu_N\theta_N)] \tag{11.7}$$

is equal to

$$\sum_{k=0}^{N} i^{2k-N} \exp[i(\nu_1 + \cdots + \nu_k)\phi_1 + i(\nu_{k+1} + \cdots + \nu_N)\phi_2]$$

$$\times \Delta(\nu_k, \ldots, \nu_1)\,\Delta(\nu_{k+1}, \ldots, \nu_N). \tag{11.8}$$

The proof is by induction. Assume that

$$\int\cdots\int_{\phi_1\leqslant\theta_1\leqslant\theta_2\leqslant\cdots\leqslant\theta_N} d\theta_1\cdots d\theta_{N-1}\exp[i(\nu_1\theta_1+\cdots+\nu_{N-1}\theta_{N-1})]$$

$$=\sum_{k=0}^{N-1} i^{2k-(N-1)}\, e^{i(\nu_1+\cdots+\nu_k)\phi_1+i(\nu_{k+1}+\cdots+\nu_{N-1})\theta_N}\,\Delta(\nu_k,\ldots,\nu_1)\,\Delta(\nu_{k+1},\ldots,\nu_{N-1}).$$

Therefore the expression (11.7) is

$$\begin{aligned}&\sum_{k=0}^{N-1} i^{2k-N+1}\, e^{i(\nu_1+\cdots+\nu_k)\phi_1}\left[e^{i(\nu_{k+1}+\cdots+\nu_N)\phi_2}-e^{i(\nu_{k+1}+\cdots+\nu_N)\phi_1}\right]\\&\qquad\times i^{-1}(\nu_{k+1}+\cdots+\nu_N)^{-1}\,\Delta(\nu_k,\nu_{k-1},\ldots,\nu_1)\,\Delta(\nu_{k+1},\ldots,\nu_{N-1})\\&=\sum_{k=0}^{N-1} i^{2k-N}\exp[i(\nu_1+\cdots+\nu_k)\phi_1+i(\nu_{k+1}+\cdots+\nu_N)\phi_2]\\&\qquad\times\Delta(\nu_k,\nu_{k-1},\ldots,\nu_1)\,\Delta(\nu_{k+1},\ldots,\nu_N)\\&\qquad-\sum_{k=0}^{N-1} i^{2k-N}\, e^{i(\nu_1+\cdots+\nu_N)\phi_1}\,\Delta(\nu_k,\nu_{k-1},\ldots,\nu_1)\,\Delta(\nu_{k+1},\ldots,\nu_N),\end{aligned}\tag{11.9}$$

where we have used the identity

$$(\nu_{k+1}+\cdots+\nu_N)^{-1}\,\Delta(\nu_{k+1},\ldots,\nu_{N-1})=\Delta(\nu_{k+1},\ldots,\nu_N).$$

By Lemma 11.1, the second sum in (11.9) is the $k=N$ term in (11.8).

Finally, for $N=1$,

$$\begin{aligned}\int_{\phi_1}^{\phi_2} e^{i\nu_1\theta_1}\,d\theta_1&=\frac{1}{i\nu_1}(e^{i\nu_1\phi_2}-e^{i\nu_1\phi_1})\\&=i^{-1}\,e^{i\nu_1\phi_2}\,\Delta(\nu_1)+i^{2-1}\,e^{i\nu_1\phi_1}\,\Delta(\nu_1).\end{aligned}\tag{11.10}$$

Thus we have proved the lemma.

Theorem 11.1. *The integral*

$$I(\phi_1,\phi_2)=\int\cdots\int_{\substack{2\phi_1\leqslant\theta_1\leqslant\theta_2\leqslant\cdots\\ \leqslant\theta_{2m}\leqslant 2\phi_2}} d\theta_1\cdots d\theta_{2m}\det[\exp(i\nu_j\theta_k)]_{j,k=1,2,\ldots,2m}\tag{11.11}$$

is given by

$$I(\phi_1, \phi_2) = \frac{(-i)^m 2^{2m} \exp[i(\phi_1 + \phi_2) \sum_1^{2m} \nu_k]}{\nu_1 \cdots \nu_{2m} \prod_{i \leqslant j < k \leqslant 2m} (\nu_j + \nu_k)} \times \det[\nu_j^{2k-2} \sin \nu_j \alpha, \nu_j^{2k-1} \cos \nu_j \alpha]_{\substack{j=1,\ldots,2m \\ k=1,\ldots,m}}, \tag{11.12}$$

with $\alpha = \phi_2 - \phi_1$. *Note that the columns of the determinant in* (11.12) *are obtained, except for a possible change in sign, by successive differentiations of the first column, that is, of* $(\sin \nu_j \alpha)$.

Proof. From (11.11) and Lemma 11.4,

$$\begin{aligned} I &= a \int\cdots\int_{\substack{2\phi_1 \leqslant \theta_1 \leqslant \theta_2 \leqslant \cdots \\ \leqslant \theta_{2m} \leqslant 2\phi_2}} d\theta_1 \cdots d\theta_{2m}\, e^{i(\nu_1\theta_1 + \cdots + \nu_{2m}\theta_{2m})} \\ &= a \sum_{k=0}^{2m} i^{2k-2m} \exp[i(\nu_1 + \cdots + \nu_k) 2\phi_1 + i(\nu_{k+1} + \cdots + \nu_{2m}) 2\phi_2] \\ &\quad \times \Delta(\nu_k, \nu_{k-1}, \ldots, \nu_1)\, \Delta(\nu_{k+1}, \ldots, \nu_{2m}) \\ &= \sum_{k=0}^{2m} (-1)^{k-m} \sum_{(\alpha)} \exp[i(\nu_{\alpha_1} + \cdots + \nu_{\alpha_k}) 2\phi_1 + i(\nu_{\alpha_{k+1}} + \cdots + \nu_{\alpha_{2m}}) 2\phi_2] \\ &\quad \times (-1)^{\sum_1^k (\alpha_j - j)}\, a\, \Delta(\nu_{\alpha_k}, \nu_{\alpha_{k-1}}, \ldots, \nu_{\alpha_1})\, a\, \Delta(\nu_{\alpha_{k+1}}, \nu_{\alpha_{k+2}}, \ldots, \nu_{\alpha_{2m}}) \\ &= \sum_{k=0}^{2m} (-1)^{k-m} \sum_{(\alpha)} \exp\left(2i\phi_1 \sum_1^k \nu_{\alpha_j} + 2i\phi_2 \sum_{k+1}^{2m} \nu_{\alpha_j}\right) (-1)^{\sum_1^k (\alpha_j - j)} \\ &\quad \times (-1)^{\sum_1^{k-1} j}\, a\, \Delta(\nu_{\alpha_1}, \nu_{\alpha_2}, \ldots, \nu_{\alpha_k})\, a\, \Delta(\nu_{\alpha_{k+1}}, \nu_{\alpha_{k+2}}, \ldots, \nu_{\alpha_{2m}}), \end{aligned} \tag{11.13}$$

where $\sum_{(\alpha)}$ means a summation with the conditions $\alpha_1 < \alpha_2 < \cdots < \alpha_k$, $\alpha_{k+1} < \alpha_{k+2} < \cdots < \alpha_{2m}$ over all permutations $\alpha_1, \alpha_2, \ldots, \alpha_{2m}$ of the indices $1, 2, \ldots, 2m$.

Let us call $\alpha_1, \alpha_2, \ldots, \alpha_k$ the "occupied" states and $\alpha_{k+1}, \alpha_{k+2}, \ldots, \alpha_{2m}$, the "unoccupied" states. We define an "occupation number" ϵ_k, which is $+1$ for the "occupied" states, $\alpha_1, \alpha_2, \ldots, \alpha_k$, and zero for the "unoccupied" states, $\alpha_{k+1}, \alpha_{k+2}, \ldots, \alpha_{2m}$. With this device we can extend the summation over all indices $1, 2, \ldots, 2m$. For example,

$$\sum_{j=1}^{k} \alpha_j = \sum_{j=1}^{2m} j\epsilon_j, \quad \sum_{j=1}^{k} \nu_{\alpha_j} = \sum_{j=1}^{2m} \epsilon_j \nu_j, \text{ etc.} \tag{11.14}$$

Using (11.6), we obtain from (11.13)

$$I = \sum_{\epsilon_1,\ldots,\epsilon_{2m}} (-1)^{m+\sum j\epsilon_j} \exp\left[2i\sum \epsilon_j\nu_j\phi_1 + 2i\sum (1-\epsilon_j)\nu_j\phi_2\right] \times (\nu_1\cdots\nu_{2m})^{-1} \prod_{1\leqslant j<k\leqslant 2m} [(\nu_k-\nu_j)(\nu_k+\nu_j)^{-1}]^{(1-\epsilon_k)(1-\epsilon_j)+\epsilon_k\epsilon_j}, \tag{11.15}$$

for $\epsilon_k\epsilon_j + (1-\epsilon_k)(1-\epsilon_j)$ is unity when both states j, k are occupied, or when both are unoccupied, and $\epsilon_k\epsilon_j + (1-\epsilon_k)(1-\epsilon_j)$ is zero otherwise.

A more symmetrical form is obtained by introducing the "spin" that takes the value $+1$ for the occupied and -1 for the unoccupied states

$$s_k = 2(\epsilon_k - \tfrac{1}{2}). \tag{11.16}$$

Equation 11.15 gives then

$$I = \sum_{s_1,\ldots,s_{2m}} (-1)^{m+(1/2)\sum j(1+s_j)} \exp\left[i(\phi_1+\phi_2)\sum_j \nu_j + i(\phi_1-\phi_2)\sum_j s_j\nu_j\right] \times (\nu_1\cdots\nu_{2m})^{-1} \prod_{1\leqslant j<k\leqslant 2m} [(\nu_k-\nu_j)(\nu_k+\nu_j)^{-1}]^{(1/2)(1+s_js_k)}, \tag{11.17}$$

$$(-1)^{(1/2)\sum j(1+s_j)} = \prod_j (-s_j)^j, \tag{11.18}$$

and

$$\prod_{1\leqslant j<k\leqslant 2m} [\nu_k-\nu_j)(\nu_k+\nu_j)^{-1}]^{(1/2)(1+s_js_k)} = \prod_{1\leqslant j<k\leqslant 2m} [(\nu_k+\nu_j)^{-1}(\nu_k-\nu_js_js_k)]$$
$$= \prod_{1\leqslant j<k\leqslant 2m} [(\nu_k+\nu_j)^{-1}(\nu_ks_k-\nu_js_j)] \prod_k s_k^{k-1}. \tag{11.19}$$

Substituting (11.18) and (11.19) into (11.17), we get

$$I = (-1)^m \exp\left[i(\phi_1+\phi_2)\sum_j \nu_j\right] (-1)^{m(2m+1)} (\nu_1\cdots\nu_{2m})^{-1} \prod_{1\leqslant j<k\leqslant 2m} (\nu_k+\nu_j)^{-1}$$
$$\times \sum_{s_1,\ldots,s_{2m}} \prod_{1\leqslant j<k\leqslant 2m} (\nu_ks_k-\nu_js_j) \exp\left[i(\phi_1-\phi_2)\sum_j s_j\nu_j\right] \prod_j s_j$$
$$= \exp\left[i(\phi_1+\phi_2)\sum_j \nu_j\right] (\nu_1\cdots\nu_{2m})^{-1} \prod_{1\leqslant j<k\leqslant 2m} (\nu_k+\nu_j)^{-1}$$
$$\times \sum_{s_1,\ldots,s_{2m}} \exp\left[i(\phi_1-\phi_2)\sum_j s_j\nu_j\right] \det[s_j{}^k\nu_j^{k-1}]_{j,k=1,\ldots,2m}. \tag{11.20}$$

The summation over s_j in (11.20) is over all possible choices $s_j = \pm 1$. By absorbing the factor $e^{i(\phi_1-\phi_2)s_j\nu_j}$ in the jth row of the determinant in (11.20), we see that each row depends only on a single s_j. The summation over s_j can therefore be carried out in each row, independently of the others.

$$\sum_{s_j=\pm 1} \exp[i(\phi_1 - \phi_2)\, s_j\nu_j]\, s_j^{2k}\nu_j^{2k-1} = 2\nu_j^{2k-1}\cos\nu_j\alpha, \qquad \alpha = \phi_2 - \phi_1\,, \tag{11.21}$$

$$\sum_{s_j=\pm 1} \exp[i(\phi_1 - \phi_2)\, s_j\nu_j]\, s_j^{2k+1}\nu_j^{2k} = -2i\nu_j^{2k}\sin\nu_j\alpha, \qquad \alpha = \phi_2 - \phi_1\,. \tag{11.22}$$

Equation 11.12 is now evident.

Theorem 11.2. *This is really a corollary of Theorem 11.1, but because of its importance we state it as a theorem. The integral*

$$\int\cdots\int_{\substack{2\phi_1\le\theta_1\le\theta_2\le\cdots\\ \le\theta_{2m}\le 2\phi_2}} d\theta_1\cdots d\theta_{2m}\det[e^{ip\theta_j}]$$

$$= i^m 2^{6m}(m!)^2\left[\prod_{j=1}^{m}(2j)!\right]^{-2}\det[D(\alpha)] \tag{11.23}$$

where p varies over the half-odd integers $-m+\frac{1}{2}, -m+\frac{3}{2},\ldots, m-\frac{1}{2}$, whereas $j = 1, 2,\ldots, 2m$. On the right-hand side $\det D(\alpha)$ *is the determinant of*

$$D(\alpha) = \left[\begin{matrix} \dfrac{\partial}{\partial q}(q^{2l-2}\sin q\alpha) & q^{2l-2}\sin q\alpha \\ \dfrac{\partial}{\partial q}(q^{2l-1}\cos q\alpha) & q^{2l-1}\cos q\alpha \end{matrix}\right]_{\substack{l=1,2,\ldots,m\\ q=\frac{1}{2},\frac{3}{2},\ldots,m-\frac{1}{2}}}, \tag{11.24}$$

where

$$\alpha = \phi_2 - \phi_1\,. \tag{11.25}$$

Proof: In Theorem 11.1 take the limits $-\nu_{2j-1} = \nu_{2j} = j - \frac{1}{2}$. Because so many factors vanish in the numerator and the denominator, it is best to add the $(2j)$th row to the $(2j-1)$th, divide the sum by $\nu_{2j-1} + \nu_{2j}$, and then take the limits $\nu_{2j-1} + \nu_{2j} \to 0$. Finally put $\nu_{2j} = j - \frac{1}{2}$. This procedure gives

$$\lim\prod_{j=1}^{m}(\nu_{2j-1} + \nu_{2j})^{-1}\det[\nu_j^{2l-2}\sin\nu_j\alpha, \nu_j^{2l-1}\cos\nu_j\alpha] = D^T(\alpha), \tag{11.26}$$

where $D^T(\alpha)$ is the transpose of $D(\alpha)$. Also

$$\nu_1 \cdots \nu_{2m} = (-1)^m \prod_{j=1}^{m} (j - \tfrac{1}{2})^2 = (-1)^m \left[\frac{(2m)!}{2^{2m} m!}\right]^2, \tag{11.27}$$

$$\nu_1 + \cdots + \nu_{2m} = 0, \tag{11.28}$$

and a straightforward, though lengthy, calculation (cf. Appendix A.18) gives

$$\prod_{1 \leqslant j < k \leqslant 2m} (\nu_j + \nu_k) \prod_{j=1}^{m} (\nu_{2j-1} + \nu_{2j})^{-1} = \left[\prod_{j=1}^{m-1} (2j)!\right]^2. \tag{11.29}$$

Substituting from (11.26), (11.27), (11.28), and (11.29) into (11.12), we obtain (11.23).

11.3. Spacing Distribution

Now, putting $\alpha = \pi - x$ into (11.23), we obtain the probability that there are no levels in an interval of length $2x$.

$$\begin{aligned} E(x) &= \int_{-\pi+x}^{\pi-x} \cdots \int d\theta_1 \cdots d\theta_{2m} (4^{2m} \pi^m m!)^{-1} \prod_{1 \leqslant j < k \leqslant 2m} | e^{i\theta_j} - e^{i\theta_k} | \\ &= \frac{(2m)!}{4^m \pi^m m!} i^{-m} \int_{\substack{-\pi+x \leqslant \theta_1 \leqslant \cdots \\ \leqslant \theta_{2m} \leqslant \pi - x}} \cdots \int d\theta_1 \cdots d\theta_{2m} \det[e^{ip\theta_j}]_{\substack{j=1,\ldots,2m \\ p=-m+\frac{1}{2},\ldots,m-\frac{1}{2}}} \\ &= 2^{2m} \pi^{-m} \frac{m!}{(2m)!} \left[\prod_{j=1}^{m-1} (2j)!\right]^{-2} D(x), \end{aligned} \tag{11.30}$$

where $D(x)$ is the determinant

$$D(x) = \begin{bmatrix} \dfrac{\partial}{\partial q} (q^{2l-2} \cos qx) + \pi q^{2l-2} \sin qx & q^{2l-2} \cos qx \\ \dfrac{\partial}{\partial q} (q^{2l-1} \sin qx) - \pi q^{2l-1} \cos qx & q^{2l-1} \sin qx \end{bmatrix} \tag{11.31}$$

$$\begin{aligned} q &= \tfrac{1}{2}, \tfrac{3}{2}, \ldots, m - \tfrac{1}{2}, \\ l &= 1, 2, \ldots, m. \end{aligned} \tag{11.31'}$$

To expand $D(x)$ in powers of x it will be convenient to simplify it further so that for small x it looks like a checkerboard. We replace the $(2l-1)$th row by the linear combination

$$r_{2l-1} + \frac{2l-2}{\pi} r_{2l-2} + x r_{2l}, \tag{11.32}$$

where r_j denotes the jth row.

For small x, including powers up to, say, the second, (11.31) can therefore be written

$$D(x) = \begin{bmatrix} \frac{1}{\pi}(2l-2)^2 q^{2l-3}x + (2l-1)q^{2l-1}x^2, & q^{2l-2} + \frac{1}{\pi}(2l-2)q^{2l-2}x + \frac{1}{2}q^{2l}x^2 \\ -\pi q^{2l-1} + 2lq^{2l-1}x + \frac{1}{2}\pi q^{2l+1}x^2, & q^{2l}x \end{bmatrix}. \tag{11.33}$$

Following Wigner, one can compute the first few terms in the power-series expansion of $D(x)$. For instance, if we write

$$D(x) = D_0 + D_1 x + D_2 x^2 + \cdots, \tag{11.34}$$

then

$$\begin{aligned} D_0 = D(0) &= \det \begin{bmatrix} 0 & q^{2l-2} \\ -\pi q^{2l-1} & 0 \end{bmatrix}_{\substack{q=\frac{1}{2},\frac{3}{2},\ldots,m-\frac{1}{2} \\ l=1,2,\ldots,m}} \\ &= \pi^m \det[q^{2l-2}] \det[q^{2l-1}] \\ &= \pi^m \prod q \prod_{p<q} (p^2 - q^2)^2 \\ &= \pi^m \frac{(2m)!}{2^{2m} m!} \prod_{j=1}^{m-1} [(2j)!]^2 \end{aligned} \tag{11.35}$$

and

$$D_1 = \left[\sum_{l=1}^{m} \left(\frac{2l-2}{\pi}\right) + \sum_{l=1}^{m} \left(-\frac{2l}{\pi}\right)\right] D_0 = -\frac{2m}{\pi} D_0, \tag{11.36}$$

so that (11.30) gives

$$E(x) = 1 - \frac{2m}{\pi} x + \cdots. \tag{11.37}$$

To calculate higher terms in the series (11.37) is more laborious. The calculation should show that the term in x^2 is absent,

$$D_2 = 0, \tag{11.38}$$

for the probability density of the spacings at zero spacing is zero. If one has enough patience, one can go a step further and get the slope of the probability density of the spacings at zero spacing.

However, to make any further progress seems to be difficult.

12 / Matrices with Gaussian Element Densities But with No Unitary or Hermitian Condition Imposed†

An ensemble of matrices whose elements are complex, quaternion, or real numbers, but with no other restrictions as to their Hermitian or unitary character, is of no immediate physical interest, for their eigenvalues may lie anywhere on the complex plane. However, an effort [Ginibre, 1] has recently been made to investigate them and the results are interesting in their own right.

To define a matrix ensemble one has to specify two things: the space T on which the matrices vary and a probability density function over T. With a view to practical applications, one can take T as the set of all real symmetric matrices with a reasonable probability density. For example, one may assume that the matrix elements are independent and have Gaussian probability densities so that all the diagonal elements have the same variance σ_1 and the same mean value, whereas all the off-diagonal elements have the mean value zero and the same variance σ_2 ; the ratio of the variances σ_1 and σ_2 being arbitrary. This case has not yet been considered analytically.

12.1. Complex Matrices

The more tractable case is to take T as the set of all $N \times N$ complex matrices. The probability that a matrix from the set T will lie in $(S, S + dS)$ is $P(S)\,\mu(dS)$, where $\mu(dS)$ is the linear measure

$$\mu(dS) = \prod_{j,k} dS_{jk}^{(0)}\, dS_{jk}^{(1)} \tag{12.1}$$

and $S_{jk}^{(0)}$, $S_{jk}^{(1)}$ are the real and imaginary parts of the matrix element

$$S_{jk} = S_{jk}^{(0)} + iS_{jk}^{(1)}. \tag{12.2}$$

† This chapter is based on the articles by Ginibre [1] and Mehta and Srivastava [1].

For the function $P(S)$ we may choose, for example [Ginibre, 1],

$$P(S) = \exp[-\mathrm{tr}(S^\dagger S)]. \tag{12.3}$$

We denote the ensemble so defined as T_c. It is visibly invariant under all unitary transformations.

To get any information about the eigenvalues, we must first find their joint probability density. This can be done, as in Chapter 3, by changing the variables from S_{jk} to the (complex) eigenvalues z_j of S and the auxiliary variables p_j. Since $\mathrm{tr}(S^\dagger S)$ is not only a function of z_j but contains other variables p_j as well, these variables have to be chosen carefully to facilitate later integrations. Let the eigenvalues of S be distinct; the case when S has multiple eigenvalues need not be considered for the same reasons as in Chapter 3. Also, let X be the $N \times N$ matrix whose columns are the eigenvectors of S so that X is nonsingular and $X^{-1}SX = E$ is diagonal. From $S = XEX^{-1}$ we obtain by differentiation

$$dS = X(dE + dAE - E\,dA)\,X^{-1}, \tag{12.4}$$

with

$$dA = X^{-1}dX. \tag{12.5}$$

Equation 12.4 reads in terms of its components:

$$(X^{-1}\,dSX)^{(0)}_{jj} = dz^{(0)}_j = dx_j\,, \quad (X^{-1}\,dSX)^{(1)}_{jj} = dz^{(1)}_j = dy_j\,, \tag{12.6}$$

$$\begin{aligned}(X^{-1}\,dSX)^{(0)}_{jk} &= (x_k - x_j)\,dA^{(0)}_{jk} - (y_k - y_j)\,dA^{(1)}_{jk}\\ (X^{-1}\,dSX)^{(1)}_{jk} &= (y_k - y_j)\,dA^{(0)}_{jk} + (x_k - x_j)\,dA^{(1)}_{jk}\end{aligned} \qquad j \neq k, \tag{12.7}$$

where x_j, y_j are the real and imaginary parts of z_j, the diagonal elements of E, whereas $dA^{(0)}_{jk}$ and $dA^{(1)}_{jk}$ are the real and imaginary parts of dA_{jk}. Whenever any set of differentials is expressed linearly in terms of those of the others, the ratio of the volume elements is equal to the Jacobian. The volume element in (12.1) is therefore given by

$$\begin{aligned}\mu(dS) &= \mu(X^{-1}\,dSX)\\ &= \prod_{j \neq k} |\, z_k - z_j\,|^2\, dA^{(0)}_{jk}\, dA^{(1)}_{jk} \prod_i dx_i\, dy_i\,.\end{aligned} \tag{12.8}$$

One still has to evaluate the integral

$$J = \int \exp[-\mathrm{tr}(S^\dagger S)] \prod_{j \neq k} dA^{(0)}_{jk}\, dA^{(1)}_{jk}, \tag{12.9}$$

which can be done by a careful choice of the new variables of integration and by using properties of determinant expansions (see Appendix A.24). The result is as follows [Ginibre, 1].

The joint probability density for the eigenvalues of S belonging to the ensemble T_c of all complex matrices is given by

$$P_c(z_1, z_2, \ldots, z_N) = K_c \exp\left(-\sum_1^N |z_i|^2\right) \prod_{1 \leqslant i < j \leqslant N} |z_i - z_j|^2, \tag{12.10}$$

where K_c is the normalization constant given later by (12.17).

With this joint probability density function one can determine various quantities of interest as easily as in Section 9.3. For example, the probability that all the eigenvalues z_i will lie outside a circle of radius α centered at $z = 0$ is

$$E_{Nc}(\alpha) = \int\cdots\int_{|z_i| \geqslant \alpha} P_c(z_1, \ldots, z_N) \prod_i dx_i\, dy_i\,. \tag{12.11}$$

By writing

$$\prod_{i<j} |z_i - z_j|^2 = \prod_{i<j} (z_i - z_j)(z_i^* - z_j^*) = \begin{vmatrix} 1 & \cdots & 1 \\ z_1 & \cdots & z_N \\ \cdots & \cdots & \cdots \\ z_1^{N-1} & \cdots & z_N^{N-1} \end{vmatrix} \begin{vmatrix} 1 & \cdots & 1 \\ z_1^* & \cdots & z_N^* \\ \cdots & \cdots & \cdots \\ z_1^{*N-1} & \cdots & z_N^{*N-1} \end{vmatrix}$$

and multiplying the two determinants row by row we get

$$E_{Nc}(\alpha) = K_c \int\cdots\int_{|z_i| \geqslant \alpha} \left(\prod_i dx_i\, dy_i\right) \exp\left(-\sum_1^N |z_i|^2\right)$$

$$\times \begin{vmatrix} N & \sum_i z_i & \cdots & \sum_i z_i^{N-1} \\ \sum_i z_i^* & \sum_i z_i^* z_i & \cdots & \sum_i z_i^* z_i^{N-1} \\ \cdots & \cdots & \cdots & \cdots \\ \sum_i z_i^{*N-1} & \sum_i z_i^{*N-1} z_i & \cdots & \sum_i z_i^{*N-1} z_i^{N-1} \end{vmatrix}. \tag{12.12}$$

Since the integrand is symmetric in all the z_i, we can replace the first row with $1, z_1, z_1^2, \ldots, z_1^{N-1}$ and multiply the result by N; z_1 can

now be eliminated from the other rows by subtracting a suitable multiple of the first row. The resulting determinant is symmetric in the $N-1$ variables $z_2, z_3, \ldots, z_N$; therefore we replace the second row with $z_2^*, z_2^* z_2, \ldots, z_2^* z_2^{N-1}$ and multiply the result by $N-1$. The process can be repeated and we get

$$E_{Nc}(\alpha) = K_c N! \int_{|z_i| \geqslant \alpha} \cdots \int \left\{ \prod_i dx_i \, dy_i \right\} \exp\left(-\sum_1^N |z_i|^2 \right) \times \begin{vmatrix} 1 & z_1 & \cdots & z_1^{N-1} \\ z_2^* & z_2^* z_2 & \cdots & z_2^* z_2^{N-1} \\ \cdots & \cdots & \cdots & \cdots \\ z_N^{*N-1} & z_N^{*N-1} z_N & \cdots & z_N^{*N-1} z_N^{N-1} \end{vmatrix}. \tag{12.13}$$

Since the various rows now depend on distinct variables, we can integrate them separately with the exponential factor. By changing to polar coordinates and performing the angular integrations first we see that

$$\int_{|z| \geqslant \alpha} e^{-|z|^2} z^{*j} z^k \, dx \, dy = \pi \, \delta_{jk} \Gamma(j+1, \alpha^2) \tag{12.14}$$

so that

$$E_{Nc}(\alpha) = K_c N! \pi^N \prod_{j=1}^N \Gamma(j, \alpha^2), \tag{12.15}$$

where $\Gamma(j, \alpha^2)$ is the incomplete gamma function

$$\Gamma(j, \alpha^2) = \int_{\alpha^2}^{\infty} e^{-x} x^{j-1} \, dx = \Gamma(j) \, e^{-\alpha^2} \sum_{l=0}^{j-1} \frac{\alpha^{2l}}{l!}. \tag{12.16}$$

Since $E_{Nc}(0) = 1$, the constant K_c can be determined from (12.15) as

$$K_c^{-1} = \pi^N \prod_{j=1}^N j! \tag{12.17}$$

and therefore

$$E_{Nc}(\alpha) = \prod_{j=1}^N \left(e^{-\alpha^2} \sum_{l=0}^{j-1} \frac{\alpha^{2l}}{l!} \right). \tag{12.18}$$

It is easy to convince oneself that $E_{Nc}(\alpha)$ tends to a well-defined limit as $N \to \infty$. For small values of α one may expand $E_{Nc}(\alpha)$ in a power series:

$$E_{Nc}(\alpha) = 1 - \alpha^2 + \tfrac{1}{2}\alpha^6 - \tfrac{5}{12}\alpha^8 + \tfrac{7}{24}\alpha^{10} - \cdots. \tag{12.19}$$

To get the coefficient of α^{2i} in the above power series one may replace $e^{-\alpha^2}\sum_{l=0}^{j-1}\alpha^{2l}/l!$ by unity for all $j > i$. In fact, one can even get for $E_c(\alpha) = \lim_{N\to\infty} E_{Nc}(\alpha)$ a series of upper bounds and a series of lower bounds converging toward each other. We have the obvious inequality

$$0 < \prod_{j=r}^{N-1} [e^{-\alpha^2} a_j(\alpha^2)] \leqslant 1, \qquad r \geqslant 0, \tag{12.20}$$

where

$$a_j(x) = \sum_{l=0}^{j} \frac{x^l}{l!} \tag{12.21}$$

is the truncated exponential series. On the other hand, the identity

$$e^{-\alpha^2} a_j(\alpha^2) = \exp\left[-\int_0^{\alpha^2} \frac{a_j(x) - a_{j-1}(x)}{a_j(x)}\, dx\right] \tag{12.22}$$

and the inequality $a_j(x) \geqslant a_l(x)$ for $j \geqslant l$ give us

$$\begin{aligned}\prod_{j=r}^{N-1} [e^{-\alpha^2} a_j(\alpha^2)] &= \exp\left[-\sum_{j=r}^{N-1} \int_0^{\alpha^2} \frac{a_j(x) - a_{j-1}(x)}{a_j(x)}\, dx\right] \\ &\geqslant \exp\left[-\sum_{j=r}^{N-1} \int_0^{\alpha^2} \frac{a_j(x) - a_{j-1}(x)}{a_r(x)}\, dx\right] \\ &= \exp\left[-\int_0^{\alpha^2} \frac{a_{N-1}(x) - a_{r-1}(x)}{a_r(x)}\, dx\right].\end{aligned} \tag{12.23}$$

Taking the limit $N \to \infty$, the inequalities (12.20) and (12.23) give us finally

$$0 < F_s(\alpha^2) f_s(\alpha^2) \leqslant F_r(\alpha^2) f_r(\alpha^2) \leqslant E_c(\alpha) \leqslant F_{r'}(\alpha^2) \leqslant F_{s'}(\alpha^2) \leqslant 1, \quad \text{if } r > s > 0,\ r' > s' > 0, \tag{12.24}$$

where

$$F_r(\alpha^2) = \prod_{j=0}^{r-1} [e^{-\alpha^2} a_j(\alpha^2)] \equiv E_{rc}(\alpha) \tag{12.25}$$

and

$$f_r(\alpha^2) = \exp\left[-\int_0^{\alpha^2} \frac{e^x - a_{r-1}(x)}{a_r(x)}\, dx\right]. \tag{12.26}$$

Equation 12.24 gives us in particular

$$\exp\left(-\alpha^2 - \int_0^{\alpha^2} \frac{e^x - 1}{1 + x}\, dx\right) \leqslant E_c(\alpha) \leqslant (1 + \alpha^2)\, e^{-2\alpha^2}. \tag{12.27}$$

To get the n-point correlation function

$$R_n(z_1, \ldots, z_n) = \frac{N!}{(N-n)!} \int \cdots \int P_c(z_1, \ldots, z_N) \prod_{i=n+1}^{N} dx_i\, dy_i \tag{12.28}$$

we proceed exactly as in Section 6.1 or 9.3. Equation 12.14 corresponds to the orthogonality property of φ_k in Section 6.1. The final result is

$$R_n(z_1, \ldots, z_n) = \pi^{-n} \exp\left(-\sum_1^n |\, z_i\,|^2\right) \det[K_N(z_i, z_j)]_{i,j=1,\ldots,n}, \tag{12.29}$$

where

$$K_N(z_i, z_j) = \sum_{l=0}^{N-1} \frac{(z_i z_j^*)^l}{l!}. \tag{12.30}$$

As $N \to \infty$ the correlation functions tend to well-defined limits:

$$R_n(z_1, \ldots, z_n) \simeq \pi^{-n} \exp\left(-\sum_1^n |\, z_i\,|^2\right) \det[e^{z_i z_j^*}]_{i,j=1,2,\ldots,n}. \tag{12.31}$$

In particular, the density of the eigenvalues is

$$R_1(z) = \pi^{-1} e^{-|z|^2} \sum_{l=0}^{N-1} \frac{|\, z\,|^{2l}}{l!}. \tag{12.32}$$

This density is isotropic and depends only on $|\, z\,| = r$, which was to be expected. It is constant $R_1(z) \approx 1/\pi$ for $r^2 \ll N$ and $R_1(z) \approx 0$ for $r^2 \gg N$. The sum in (12.32) can be estimated in an elementary way. From the inequalities

$$\begin{aligned} e^{r^2} - \sum_0^{N-1} \frac{r^{2l}}{l!} &= \sum_N^{\infty} \frac{r^{2l}}{l!} \leqslant \frac{r^{2N}}{N!} \sum_0^{\infty} \frac{r^{2l}}{(N+1)^l} \\ &= \frac{r^{2N}}{N!} \frac{N+1}{N+1-r^2}, \quad \text{for} \quad r^2 \lesssim N \end{aligned} \tag{12.33}$$

and

$$\sum_0^{N-1} \frac{r^{2l}}{l!} \leqslant \frac{r^{2(N-1)}}{(N-1)!} \sum_0^{N-1} \left(\frac{N-1}{r^2}\right)^l$$

$$= \frac{r^{2(N-1)}}{(N-1)!} \frac{r^2}{r^2 - N + 1}, \quad \text{for} \quad r^2 \gtrsim N \tag{12.34}$$

we get

$$1 - \pi R_1(z) \leqslant e^{-r^2} \frac{r^{2N}}{N!} \frac{N+1}{N+1-r^2} \quad \text{for} \quad r^2 \lesssim N \tag{12.35}$$

and

$$\pi R_1(z) \leqslant e^{-r^2} \frac{r^{2N}}{N!} \frac{N}{r^2+1-N} \quad \text{for} \quad r^2 \gtrsim N. \tag{12.36}$$

One can also estimate how fast the eigenvalue density falls from $1/\pi$ to 0 around $r^2 = N$. Putting $r = N^{1/2} \pm u$, $0 \leqslant u \lesssim 1 \ll N$, the leading term in (12.35) and (12.36) is $e^{-u^2}/2u\sqrt{\pi}$.

The two-point correlation function in the limit $N \to \infty$ is

$$R_2(z_1, z_2) = \pi^{-2}[1 - \exp(-|z_1 - z_2|^2)] \tag{12.37}$$

and depends only on the distance between the eigenvalues.

12.2. Quaternion Matrices

In this section we consider matrices whose elements are real quaternions (cf. Chapter 2). All four quaternion components of each matrix element are random variables. To proceed any further one has to know about the diagonalization of these matrices. The eigenvalue equation may be written as

$$SY = Y\lambda, \tag{12.38}$$

where Y is a vector with N quaternion components (the eigenvector) and λ is a quaternion number (the eigenvalue). There is no reason a priori for (real quaternion) solutions to (12.38) to exist. Fortunately, they do and in sufficient number (Appendix A.25). Writing (12.38) as

$$SY\mu = Y\mu(\mu^{-1}\lambda\mu), \tag{12.39}$$

we see that if λ is an eigenvalue then so is $\mu^{-1}\lambda\mu$ for arbitrary μ. Thus the eigenvalues of a given matrix are not just discrete points

but describe closed curves, and one has to talk about the distribution of these eigencurves in the four-dimensional space. Even if one chooses to describe these curves by some fixed point or points on them, only one-sided linear independence of the corresponding eigenvector rays can be established by the usual methods. Although, for a given quaternion real matrix S another such X can be found (in the favorable circumstance of distinct eigencurves) which diagonalizes it,

$$S = XEX^{-1} \tag{12.40}$$

[E diagonal and real (Appendix A.25)], it seems difficult to establish it by purely quaternion means.

In view of these difficulties, from now on we shall employ the matrix representation of quaternions (cf. Chapter 2), thus doubling the size of the matrix S, and use well-known results on matrices with complex elements. Thus, in reality, this section does not deal with the quaternion matrices as such but with even-order complex matrices having a special structure; the elements of S satisfy the relations

$$S_{2i,2j} = S^*_{2i-1,2j-1}\,,\ S_{2i-1,2j} = -S^*_{2i,2j-1} \tag{12.41}$$

or, in the matrix notation,

$$SZ = ZS^*, \tag{12.42}$$

where Z is the antisymmetric, real, unitary matrix (2.21).

If X is the $2N \times 2N$ matrix whose columns are the eigenvectors of S, the eigenvalues being all distinct, then

$$S = XEX^{-1}, \tag{12.43}$$

where E is diagonal. The diagonal elements of E occur in complex conjugate pairs $z_j\,,\ z_j^*;\ j = 1, 2,..., N$. The linear measure is

$$\mu(dS) = \prod_{\substack{1 \le i < j \le N \\ \lambda = 0,1}} dS^{(\lambda)}_{2i,2j}\, dS^{(\lambda)}_{2i-1,2j} \tag{12.44}$$

where $dS^{(0)}_{ij}$ and $dS^{(1)}_{ij}$ are the real and imaginary parts of dS_{ij}. For $P(S)$ we take

$$P(S) = \exp[-\tfrac{1}{2}\ \mathrm{tr}(S^\dagger S)]; \tag{12.45}$$

the factor $\frac{1}{2}$ is there to compensate for the artificial doubling of the size of S. We denote the ensemble so defined by T_Q. Equations 12.4

and 12.5 are valid. If we write (12.4) in terms of the various components, the volume element of dS is

$$\mu(dS) = \prod_i | z_i - z_i^* |^2 \prod_{i \neq j} (| z_i - z_j |^2 | z_i - z_j^* |^2)$$

$$\times \prod_{\substack{i<j \\ \lambda=0,1}} dA^{(\lambda)}_{2i,2j}\, dA^{(\lambda)}_{2i-1,2j}\,. \tag{12.46}$$

The integration corresponding to (12.9) for this case is carried out in Appendix A.26. The result is as follows [Ginibre, 1].

The joint probability density function for the eigenvalues of S belonging to the ensemble T_Q of all complex matrices satisfying (12.41) is given by

$$P_Q(z_1, \ldots, z_N) = K_Q \exp\left(-\sum_1^N | z_i |^2\right) \prod_1 | z_i - z_i^* |^2$$

$$\times \prod_{i<j} (| z_i - z_j |^2 | z_i - z_j^* |^2), \tag{12.47}$$

where K_Q is the normalization constant given by (12.52).

With this joint probability density function one can determine the various quantities of interest with almost the same ease as in the unitary case. The method to be followed in all such calculations is to express $P_Q(z_1, \ldots, z_N)$ as a confluent alternant type determinant and use the integration method developed in Chapter 5 and Appendix A.7.

Let us write a $2N \times 2N$ Vandermonde determinant of the variables z_j, z_j^*, $j = 1, 2, \ldots, N$; that is, the determinant whose $(2j-1)$th column consists of the successive powers of z_j, $(1, z_j, z_j^2, \ldots, z_j^{2N-1})$, and whose $2j$th column consists of the successive powers of z_j^*, $(1, z_j^*, z_j^{*2}, \ldots, z_j^{*2N-1})$, for $j = 1, 2, \ldots, N$. We can clearly see that this determinant is nothing but

$$\prod_i (z_i^* - z_i) \prod_{i<j} (| z_i - z_j |^2 | z_i - z_j^* |^2). \tag{12.48}$$

Thus we are led to define

$$f_{ij}(u) = \iint e^{-|z|^2}(z - z^*)\, u(z)(z^i z^{*j} - z^j z^{*i})\, dx\, dy, \tag{12.49}$$

and the average value of $\prod_i u(z_i)$ (see Appendix A.7) is

$$\left\langle \prod_i u(z_i) \right\rangle = \int \cdots \int P_Q(z_1, \ldots, z_N) \prod_i u(z_i)\, dx_i\, dy_i$$

$$= K_Q N! \, (\det[f_{ij}]_{i,j=0,1,\ldots,2N-1})^{1/2}. \tag{12.50}$$

Putting $u(z) = 1$ and equating the average (12.50) to unity, we get the value of K_Q :

$$f_{ij}(1) = 2\pi(j!\, \delta_{i+1,j} - i!\, \delta_{j+1,i}) \tag{12.51}$$

$$K_Q^{-1} = N! \, (2\pi)^N \prod_1^N \Gamma(2j). \tag{12.52}$$

Next we put

$$u(z) = 0, \quad \text{if} \quad |z| < \alpha$$

$$= 1, \quad \text{if} \quad |z| \geqslant \alpha, \tag{12.53}$$

and obtain an expression for $E_{NQ}(\alpha)$, the probability that no eigenvalue will lie inside a circle of radius α centered at the origin

$$E_{NQ}(\alpha) = \prod_{j=1}^{N} \frac{\Gamma(2j, \alpha^2)}{\Gamma(2j, 0)}, \tag{12.54}$$

where the incomplete gamma functions $\Gamma(j, \alpha^2)$ are defined by (12.16). Corresponding to (12.19), (12.20), and (12.26), we now have

$$E_Q(\alpha) = \lim_{N\to\infty} E_{NQ}(\alpha) = 1 - \tfrac{1}{2}\alpha^4 + \tfrac{1}{3}\alpha^6 - \tfrac{1}{6}\alpha^8 + \tfrac{1}{15}\alpha^{10} - \cdots, \tag{12.55}$$

$$f_s(\alpha) \leqslant f_r(\alpha^2) \leqslant E_Q(\alpha) \leqslant F_{r'}(\alpha^2) \leqslant F_{s'}(\alpha^2), \quad \text{for} \quad r > s, r' > s' > 0, \tag{12.56}$$

where now

$$F_r(\alpha^2) = \prod_{j=0}^{r-1} (e^{-\alpha^2} a_{2j+1}(\alpha^2)) \tag{12.57}$$

and

$$f_r(\alpha^2) = F_r(\alpha^2) \exp\left[-\int_0^{\alpha^2} \frac{\frac{1}{2}(e^x - e^{-x}) - a_{2r-1}(x)}{a_{2r+1}}\, dx\right], \tag{12.58}$$

with $a_j(x)$ given by (12.21).

To get the correlation functions $R_n(z_1, z_2, \ldots, z_N)$, equation (5.2), we substitute $u(z) = 1 + a(z)$ and expand in powers of $a(z)$. To avoid manipulations analogous to those leading from (6.34) to (6.45) (or to get them built in from the very beginning) we replace z^{2i} by $\sum_{k=0}^{i} 2^i i!\, z^{2k}/2^k k!$ and z^{2i+1} by $z^{2i+1}/(2i+1)!$ in the determinantal form of (12.47). These replacements amount to multiplying (12.47) by a known constant. We are thus led to introduce the quantities

$$\lambda_{ij} \equiv (2\pi)^{-1} f_{2i,2j} = (2\pi)^{-1} \sum_{k=0}^{i} \sum_{l=0}^{j} \frac{2^i i!}{2^k k!} \frac{2^j j!}{2^l l!} \times \iint (z^{2k} z^{*2l} - z^{2l} z^{*2k})(z - z^*)\, e^{-|z|^2}\, a(z)\, dx\, dy, \tag{12.59}$$

$$\mu_{ij} \equiv (2\pi)^{-1} f_{2i+1,2j+1} = (2\pi)^{-1} [(2i+1)!\,(2j+1)!]^{-1} \times \iint (z^{2i+1} z^{*2j+1} - z^{2j+1} z^{*2i+1})(z - z^*)\, e^{-|z|^2}\, a(z)\, dx\, dy, \tag{12.60}$$

and

$$\delta_{ij} + \nu_{ij} = (2\pi)^{-1} f_{2i,2j+1}, \tag{12.61}$$

with

$$\nu_{ij} = (2\pi)^{-1} [(2j+1)!]^{-1} \sum_{k=0}^{i} \frac{2^i i!}{2^k k!} \times \iint (z^{2k} z^{*2j+1} - z^{2j+1} z^{*2k})(z - z^*)\, e^{-|z|^2}\, a(z)\, dx\, dy. \tag{12.62}$$

We thus get the average value of $\prod_i (1 + a_i(z_i))$ in terms of ν_{ij}, λ_{ij}, and μ_{ij} (cf. Appendix A.7):

$$\left\langle \prod_i (1 + a(z_i)) \right\rangle = \int \cdots \int P_Q(z_1, z_2, \ldots, z_N) \prod_i (1 + a(z_i))\, dx_i\, dy_i = \left(\det \begin{bmatrix} \lambda_{ij} & \delta_{ij} + \nu_{ij} \\ -\delta_{ij} - \nu_{ji} & \mu_{ij} \end{bmatrix}_{i,j=0,1,\ldots,2N-1} \right)^{1/2} \tag{12.63}$$

$$= 1 + \sum_i \nu_{ii} + \frac{1}{2!} \sum_{i,j} (\nu_{ii}\nu_{jj} - \nu_{ij}\nu_{ji} - \lambda_{ij}\mu_{ij}) + \cdots, \tag{12.64}$$

where the summations are taken from 0 to $N-1$ over the indices independently. The various correlation functions can now be obtained by functional differentiations of (12.64) at $a(z) = 0$.

For example, the eigenvalue density is

$$R_1(z) = \left[\frac{\delta}{\delta a(z)}\left\langle\prod_i (1 + a(z_i))\right\rangle\right]_{a(z)=0} = \frac{\delta}{\delta a(z)}\sum_i \nu_{ii}$$

$$= (2\pi)^{-1} e^{-|z|^2}(z - z^*) \sum_{i=0}^{N-1}\sum_{k=0}^{i} \frac{2^i i!}{2^k k!}\frac{1}{(2i+1)!}(z^{2k}z^{*2i+1} - c.c.). \tag{12.65}$$

To get the limiting value when $N \to \infty$ it is more convenient to take k and $i - k = k_1$ as independent summation indices. This gives us

$$R_1(z) = e^{-|z|^2}(z - z^*)\,\phi(z, z^*), \tag{12.66}$$

where

$$\phi(z, z^*) = (2\pi)^{-1/2} \sum_{k,k_1=0}^{\infty} [k!\Gamma(k + k_1 + \tfrac{3}{2})]^{-1}$$

$$\times \left[\left(\frac{z^2}{2}\right)^k \left(\frac{z^{*2}}{2}\right)^{k+k_1+1/2} - \left(\frac{z^2}{2}\right)^{k+k_1+1/2}\left(\frac{z^{*2}}{2}\right)^k\right] \tag{12.67}$$

$$= (2\pi)^{-1/2} \sum_{k_1=0}^{\infty} I_{k_1+1/2}(zz^*)\left[\left(\frac{z^*}{z}\right)^{k_1+1/2} - \left(\frac{z}{z^*}\right)^{k_1+1/2}\right] \tag{12.68}$$

and $I_{k_1+1/2}$ is the Bessel function

$$I_{k_1+1/2}(x) = \sum_{k=0}^{\infty} [k!\,\Gamma(k + k_1 + \tfrac{3}{2})]^{-1}\,(\tfrac{1}{2}x)^{2k+k_1+1/2}. \tag{12.69}$$

By using the recurrence relation for Bessel functions, we obtain the following simpler equation (cf. Appendix A.27)

$$\phi(z, z^*) = \pi^{-1}(z^* - z)\,e^{zz^*}\int_0^1 e^{(1/2)(z-z^*)^2 x}\frac{dx}{\sqrt{1-x}}. \tag{12.70}$$

Equations 12.66 and 12.70 give the eigenvalue density. The two-point and three-point correlation functions $R_2(z_1, z_2)$ and $R_3(z_1, z_2, z_3)$ can be computed similarly. The same function ϕ with different

arguments makes it appearance, and after a little manipulation the final result can be written as

$$R_2(z_1, z_2) = \prod_{i=1}^{2} [e^{-|z_i|^2}(z_i - z_i^*)] \times \left\{ \det \begin{bmatrix} \phi(z_i, z_j) & \phi(z_i, z_j^*) \\ \phi(z_i^*, z_j) & \phi(z_i^*, z_j^*) \end{bmatrix}_{i,j=1,2} \right\}^{1/2} \tag{12.71}$$

and

$$R_3(z_1, z_2, z_3) = \prod_{i=1}^{3} [e^{-|z_i|^2}(z_i - z_i^*)] \times \left\{ \det \begin{bmatrix} \phi(z_i, z_j) & \phi(z_i, z_j^*) \\ \phi(z_i^*, z_j) & \phi(z_i^*, z_j^*) \end{bmatrix}_{i,j=1,2,3} \right\}^{1/2}, \tag{12.72}$$

where

$$\phi(\alpha, \beta) = \pi^{-1}(\beta - \alpha)\, e^{\alpha\beta} \int_0^1 \exp[\tfrac{1}{2}(\alpha - \beta)^2 x] \frac{dx}{\sqrt{1 - x}}. \tag{12.73}$$

Although we have no proof, it appears that (12.71) and (12.72) have an obvious generalization.

12.3. Real Matrices

A matrix with real elements does not necessarily possess a sufficient number of real solutions to the eigenvalue equation (12.38). This is perhaps the reason for the great difficulties experienced in the investigation of random matrices with real elements. If all the eigenvalues are real, then, taking

$$P(S)\, \mu(dS = e^{-\mathrm{tr}(S^\dagger S)} \prod_{i,j} dS_{ij}, \tag{12.74}$$

Ginibre has shown that the probability density function for the eigenvalues is identical to that of Gaussian orthogonal ensembles (3.17) (cf. Appendix A.28). If some of the eigenvalues are not real, one has only a complicated integral expression for the joint probability density function of the eigenvalues, which has not yet been simplified. We hope that this will be done some day.

13 / Gaussian Ensembles. Level Density in the Tail of the Semicircle

The density of nuclear levels increases steeply almost like an exponential in the experimentally observed energy range. On the other hand, the eigenvalue density for the Gaussian ensembles is a semicircle in the first approximation:

$$\sigma(x) \approx \pi^{-1}(A - x^2)^{1/2}, \qquad A = 2N. \tag{13.1}$$

Therefore one might think that near the lower end, $x = -A^{1/2}$, this density looks like the actual rise in nuclear level density. Although the deviations must be small compared with the dominant behavior (13.1), the tail might still contain an infinite number of eigenvalues. For example, $\sigma(x)$ may be proportional to $N^{1/3}$ in a region extending to $N^{1/6}$, so that the number of eigenvalues not accounted for by (13.1) will be proportional to $N^{1/2}$, increasing rapidly with the dimension number N of the matrices of the set. If this were the case, we should expect that the correlation functions and the spacing distribution in the tail part would be nearer the actual situation. However, the following calculation [Bronk, 2] shows that this is not true. Near $x = \pm A^{1/2}$ the eigenvalue density $\sigma_N(x)$ is $\sim N^{1/6}$ in a region of extent $\sim N^{-1/6}$, so that the total number of eigenvalues in the tail part remains finite and amounts to only a few, even when $N \to \infty$.

For the Gaussian unitary ensemble the eigenvalue density $\sigma_N(x)$ is the sum (cf. (6.9) and (A9.1))

$$\sigma_N(x) = \sum_{j=0}^{N-1} \varphi_j^2(x) \equiv \sum_{j=0}^{N-1} \{2^j j! \pi^{1/2}\}^{-1/2} e^{-x^2} \{H_j(x)\}^2 \tag{13.2}$$

$$= (\tfrac{1}{2}N)^{1/2} [\varphi_N'(x)\, \varphi_{N-1}(x) - \varphi_N(x)\, \varphi_{N-1}'(x)], \tag{13.3}$$

where φ' is the derivative of φ. To have any similarity with the exponential the function $\sigma_N(x)$ must be convex from below, and we

will be interested only in that region. Let us therefore determine the inflection point x_0 such that $\sigma_N(x)$ is convex from below for all $x > x_0$. Differentiating (13.3) and substituting from the differential equation

$$\varphi_j''(x) + (2j + 1 - x^2)\,\varphi_j(x) = 0, \tag{13.4}$$

satisfied by the harmonic oscillator function $\varphi_j(x)$, we get

$$\sigma_N'(x) = -(2N)^{1/2}\,\varphi_N(x)\,\varphi_{N-1}(x).$$

Differentiating once more, we obtain

$$\sigma_N''(x) = -(2N)^{1/2}\,[\varphi_N'(x)\,\varphi_{N-1}(x) + \varphi_N(x)\,\varphi_{N-1}'(x)]. \tag{13.5}$$

We are interested in the location of the largest zero of $\sigma_N''(x)$. For $x \geqslant (2N)^{1/2}$, $\varphi_N(x)$ and $\varphi_{N-1}(x)$ are both positive and decreasing so that $\varphi_N'(x)$ and $\varphi_{N-1}'(x)$ are both negative. Because the outermost maxima of $\varphi_j(x)$ move out with the increase of j, $\varphi_{N-1}'(x)$ is negative and $\varphi_N(x)$ is positive when $\varphi_N'(x)$ first becomes zero. As we decrease x across the value $(2N)^{1/2}$, $\varphi_N(x)$ will attain its maximum value and then decrease to zero, whereas $\varphi_{N-1}(x)$ will always remain positive. Thus $\sigma_N''(x)$ changes sign as x varies from the largest zero of $\varphi_N'(x)$ to the largest zero of $\varphi_N(x)$ and therefore must vanish somewhere in between. These largest zeros lie very near each other and their location is known [Szegö, 1]:

$$x_0 \approx (2N)^{1/2} - 1.856(2N)^{-1/6}. \tag{13.6}$$

We are interested in estimating the number of eigenvalues larger than x_0:

$$\int_{x_0}^{\infty} \sigma_N(x)\,dx. \tag{13.7}$$

To estimate $\varphi_j(x)$ near the transition point put

$$x = (2j + 1)^{1/2} - 2^{-1/2}3^{-1/3}j^{-1/6}t \tag{13.8}$$

so that the differential equation (13.4) for $\varphi_j(x)$ is transformed to

$$\frac{d^2}{dt^2}\tilde{\phi}_j(t) + \tfrac{1}{3}t\tilde{\phi}_j(t) = 0, \qquad t \cdot j^{-2/3} \ll 1 \ll j; \tag{13.9}$$

$$\varphi_j(x) = \tilde{\phi}_j(t). \tag{13.10}$$

This is Airy's equation. The solution that goes to zero for $t \to -\infty$ is given by

$$\tilde{\phi}_j(t) = i\, a\, j^{-1/12}(-t)^{1/2}\, e^{i(\pi/6)} H^{(1)}_{1/3}\left(2i\left(\frac{-t}{3}\right)^{3/2}\right), \qquad t \leqslant 0, \tag{13.11a}$$

$$= 2(3)^{-1/2} a\, j^{-1/12} t^{1/2}\left\{J_{-1/3}\left[2\left(\frac{t}{3}\right)^{3/2}\right] + J_{1/3}\left[2\left(\frac{t}{3}\right)^{3/2}\right]\right\}, \quad t \geqslant 0, \tag{13.11b}$$

The ratio of the normalizations in (13.11a) and (13.11b) is such that the two forms of $\tilde{\phi}_j(t)$ join smoothly at $t = 0$. The constant a may be determined from the condition that the average value of $|\tilde{\phi}_j(t)|^2$ over an interval for $t > 0$, $t \gg 1$ coincides with the classical approximation of the quantum mechanical probability density for a harmonic oscillator in that region. The asymptotic form of (13.11b) to be used for this purpose is

$$\tilde{\phi}_j(t) \propto t^{-1/4} \cos(\beta t^{3/2} - \tfrac{1}{4}\pi), \qquad t > 0, \tag{13.12}$$

$$\beta = 2(3)^{-3/2}. \tag{13.13}$$

For our purposes it is sufficient to note that a is a small constant of the order of 0.3 and it does not depend on j. Using the power series for $J_{-1/3}$ and $J_{1/3}$ in (13.11b), we obtain an approximate expression for $\tilde{\phi}_j(t)$ for $|t| < 1$:

$$\tilde{\phi}_j(t) \approx \bar{\phi}_j(t) = c_1 j^{-1/12}\, e^{c_2 t}, \tag{13.14}$$

$$c_1 \approx 1.477a, \qquad c_2 \approx 0.506. \tag{13.15}$$

As $t \to 0$, $\bar{\phi}_j(t) \to \tilde{\phi}_j(t)$. We can actually prove that, for $t < 1.856.6^{1/3} j^{-1/6}$, $\tilde{\phi}_j(t) \leqslant \bar{\phi}_j(t)$. However, we shall evade this issue and be satisfied with the approximation of replacing $\tilde{\phi}_j(t)$ by $\bar{\phi}_j(t)$ in the entire region $x > x_0$. Realizing that the only terms contributing appreciably to the summation (13.2) are those with a large j, we set

$$j = N - 1 - \mu, \tag{13.16}$$

$$(2N)^{1/2} + y = (2j+1)^{1/2} - 2^{-1/2} 3^{-1/3} j^{-1/6} t, \tag{13.17}$$

expand in powers of μ/N, and keep only the dominant terms to get

$$\{\bar{\phi}_{N-\mu-1}(t)\}^2 \approx c_1^2 N^{-1/6} \exp\left[-2^{3/2} 3^{1/3} c_2 N^{1/6}\left(y + \frac{\mu}{(2N)^{1/2}}\right)\right]$$

$$\approx 2.18 a^2 N^{-1/6} \exp\{-2.06 N^{1/6}[y + \mu(2N)^{-1/2}]\}. \tag{13.18}$$

Putting this in equation (13.2),

$$\sigma_N(x) \approx 2.18a^2N^{-1/6}\exp\{-2N^{1/6}y\}\int_0^N \exp(-2^{1/2}N^{-1/3}\mu)\,d\mu$$
$$\approx 1.54a^2N^{1/6}\exp(-2N^{1/6}y). \tag{13.19}$$

Thus the eigenvalue density in the tail varies as $N^{1/6}$. The total number of eigenvalues in the tail is

$$\int_{x_0}^{\infty}\sigma_N(x)\,dx \approx 1.54a^2N^{1/6}\int_{-1.856(2N)^{-1/6}}^{\infty}\exp\{-2N^{1/6}y)\,dy$$
$$\approx 0.77a^2\exp(2^{5/6\times 1.856}) \approx 14.3a^2 \lesssim 2. \tag{13.20}$$

To get the eigenvalue density (5.56) for the Gaussian orthogonal ensemble in the tail of the semicircle we need an estimation of

$$I = m^{1/2}\int_{x_0}^{\infty} dx\left[\varphi_{2m-1}(x)\int_0^x \varphi_{2m}(y)\,dy\right]. \tag{13.21}$$

where $N = 2m$. We transform the second integral on the right as follows:

$$\int_0^x \varphi_{2m}(y)\,dy = \int_0^{\infty}\varphi_{2m}(y)\,dy - \int_x^{\infty}\varphi_{2m}(y)\,dy$$
$$= \frac{1}{2}\int_{-\infty}^{\infty}\varphi_{2m}(y)\,dy - \int_x^{\infty}\varphi_{2m}(y)\,dy = -\int_x^{\infty}\varphi_{2m}(y)\,dy, \quad m \neq 0$$

so that

$$I = -m^{1/2}\int_{x_0}^{\infty} dx\left[\varphi_{2m-1}(x)\int_x^{\infty}\varphi_{2m}(y)\,dy\right]. \tag{13.22}$$

Now one can use the approximation (13.14) to see that the right-hand side of (13.22) is a small constant independent of m.

14 / Bordered Matrices

In some physical situations it is instructive to consider an ensemble of matrices that is slightly more difficult than the diagonal ones. Matrices that have their elements in the principal diagonal and a few neighboring super- or underdiagonals distributed at random, while all other matrix elements are zero, are of importance, for example, in the theory where glass is represented as a collection of random nets. It is required to determine the distribution of the characteristic frequencies and of the modulus square of the characteristic amplitudes of such random nets. The problem of such a linear chain with nearest neighbor interactions was first solved by Dyson [6]. The problem is simple enough to be treated analytically. It corresponds to an ensemble of Hermitian matrices whose only nonzero elements are those that lie in the layers immediately above and immediately below the principal diagonal. Later Wigner [5] treated the case of real symmetric matrices whose diagonal elements were equispaced. Elements in a few layers on each side of this principal diagonal had the same nonzero magnitude with a random sign, whereas all other elements were zero. In this chapter we present briefly some important features of these investigations.

In the physical situation of a disordered linear chain in which each atom interacts with many of its neighbors one has to deal with matrices with many layers of nonzero random elements on both sides of the principal diagonal [Englman, 1]. Attempts have also been made to treat the two- and three-dimensional lattices of random oscillators. A nice review containing references to the earlier work is due to Maradudin et al. [1].

14.1. Random Linear Chain†

Consider a chain of N masses, each connected to its immediate neighbors by springs that obey Hooke's law. The masses and the

† Dyson [6].

spring constants are random variables with known average characteristics. The problem is to determine the probability density function of the normal frequencies of this chain. This theory applies equally well to an electric transmission line composed of alternating capacitances and inductances with random characteristics. We will be interested in the limit $N \to \infty$.

By simple algebraic manipulations it can be shown [Dyson, 6] that the normal frequencies of such a chain are the eigenvalues of the $(2N - 1) \times (2N - 1)$ Hermitian matrix with elements

$$H_{j+1,j} = -H_{j,j+1} = i\lambda_j^{1/2}, \tag{14.1}$$

where the λ_j are given in terms of the masses and the spring constants K_j (connecting m_j and m_{j+1}) by

$$\lambda_{2j-1} = \frac{K_j}{m_j}, \qquad \lambda_{2j} = \frac{K_j}{m_{j+1}}. \tag{14.2}$$

Because H is antisymmetric having an odd order its determinant is zero and hence one of its eigenvalues vanishes; this corresponds to the degenerate motion in which all the masses have exactly the same displacement. All other eigenvalues occur in pairs ω_j, $-\omega_j$. Let $M(\mu)$ be the distribution function defined as the proportion of the eigenvalues ω_j for which $\omega^2 \leqslant \mu$, so that a probability density function can be defined as

$$D(\mu) = \frac{dM(\mu)}{d\mu}. \tag{14.3}$$

It is required to find either $M(\mu)$ or $D(\mu)$ in the limit $N \to \infty$ when the distribution of the λ_j is given.

Dyson, instead, considers

$$\begin{aligned} \Omega(z) &= \lim_{N\to\infty} (2N - 1)^{-1} \sum_j \ln(1 + z\omega_j^2) \\ &= \int_0^\infty \ln(1 + z\mu)\, D(\mu)\, d\mu \end{aligned} \tag{14.4}$$

as a function of the complex variable z. That branch of the logarithm is taken which is real for real positive z. Then the integral (14.4) is convergent and defines an analytical function of z for the whole z-plane, except for the negative real axis. As z tends from above to

a point $-x$ on the negative real axis, the imaginary part of $\ln(1 + z\mu)$ tends to zero if $x\mu < 1$ and to π if $x\mu > 1$. Hence (14.4) gives

$$\operatorname{Re}\left[\frac{1}{i\pi}\lim_{\epsilon\to 0}\Omega(-x + i\epsilon)\right] = \int_{1/x}^{\infty} D(\mu)\,d\mu$$
$$= 1 - M\left(\frac{1}{x}\right), \tag{14.5}$$

or on differentiating,

$$D\left(\frac{1}{x}\right) = -x^2 \operatorname{Re}\left[\frac{1}{i\pi}\lim_{\epsilon\to 0}\Omega'(-x + i\epsilon)\right]. \tag{14.6}$$

Therefore, once $\Omega'(z)$ is known, $D(\mu)$ is determined by its limiting values on the negative real axis. Most of the time, however, it is not possible to express $\Omega'(z)$ as a closed analytical expression and direct analytical continuation becomes impossible. For use in such cases Dyson derived the formula

$$D(\mu) = (2\pi^2\mu)^{-1}\int_{-\infty}^{\infty} d\alpha(\cosh \pi\alpha)\left\{\int_0^{\infty} dx(x\mu)^{-1}\cos[\alpha \ln(x\mu)]\,\Omega'(x)\right\} \tag{14.7}$$

to express $D(\mu)$ in terms of the values of $\Omega'(z)$ for the real positive values of z. The details of the derivation will not be given. Thus, even if $\Omega'(z)$ is known only numerically or approximately on the positive real axis, $D(\mu)$ can be evaluated by numerical integration.

Dyson derives an explicit formula for $\Omega(z)$ in terms of the λ_j. The derivation is based on expanding $\ln(1 + z\omega_j^2)$ in powers of z, replacing the sums of powers of ω_j^2 by the traces of the even powers of H and counting the terms that give a nonzero contribution to such traces. We will content ourselves by giving only the results and refer the interested reader to the original paper [Dyson, 6].

For an arbitrary chain with given coefficients λ_j, $\Omega(z)$ is given by

$$\Omega(z) = \lim_{N\to\infty}\frac{1}{N}\sum_{a=1}^{2N-1}\ln(1 + \xi(a)), \tag{14.8}$$

where $\xi(a)$ is the continued fraction

$$\xi(a) = \cfrac{z\lambda_a}{1 + \cfrac{z\lambda_{a+1}}{1 + \cfrac{z\lambda_{a+2}}{1 + \cdots}}}. \tag{14.9}$$

Various assumptions can now be made as to what extent the λ_j are random and the consequences for the function $\Omega(z)$ can be derived.

For example, if all the λ_j were the same, then all the ξ would be the same and

$$\xi = \frac{z\lambda}{1+\xi}; \qquad \xi \to 0, \quad \text{as} \quad z \to 0.$$

Therefore

$$\xi = \tfrac{1}{2}[(1+4z\lambda)^{1/2} - 1] \tag{14.10}$$

and by (14.8)

$$\Omega(z) = 2 \ln [\tfrac{1}{2}(1+4z\lambda)^{1/2} + \tfrac{1}{2}]. \tag{14.11}$$

On differentiation we get

$$\Omega'(z) = z^{-1}[1 - (1+4z\lambda)^{-1/2}], \qquad z > 0. \tag{14.12}$$

Continuing through the upper half plane to real negative values of $z < -(4\lambda)^{-1}$ we find

$$\Omega'(z) = z^{-1}[1 + i(1+4z\lambda)^{-1/2}]. \tag{14.13}$$

Hence (14.6) gives

$$\begin{aligned} D(\mu) &= \frac{1}{\pi}(4\lambda\mu - \mu^2)^{-1/2}, \qquad \mu < 4\lambda, \\ &= 0, \qquad\qquad\qquad\quad \mu > 4\lambda, \end{aligned} \tag{14.14}$$

a result that in this simple case can be checked directly.

Two more special cases have been analytically treated by Dyson:

1. All the λ_j are independent random variables with a given probability density function $G(\lambda)$. In this case the $\xi(a)$ also have a probability density function $F(\xi)$, the same for all a. An integral equation for $F(\xi)$ can be derived by equating the probabilities on both sides of the equality

$$\xi(a) = \frac{z\lambda_a}{1+\xi(a+1)}.$$

The kernel of the integral equation contains $G(\lambda)$. Once this integral equation is solved by iteration or otherwise and $F(\xi)$ so obtained is normalized,

$$\int_0^\infty F(\xi)\, d\xi = 1, \tag{14.15}$$

$\Omega(z)$ is given by (14.8):

$$\Omega(z) = 2\int_0^\infty F(\xi)\ln(1+\xi)\,d\xi. \tag{14.16}$$

$D(\mu)$ can then be calculated.

2. The λ_j are correlated in the following manner. The spring constants K_j are fixed and equal and the masses are independent random variables with a given probability density function $G(m)$. In this case the variables $\eta_j = [\xi(2j)]^{-1}$ are uncorrelated and an integral equation for their probability density function can be derived from the recurrence formula

$$\eta_j = \frac{m_j + 1}{zk} + \frac{\eta_{j+1}}{1+\eta_{j+1}}, \qquad (k_j = k). \tag{14.17}$$

Equation 14.17 follows from (14.9) and (14.2). Also from (14.9) we have

$$[1+\xi(2j)][1+\xi(2j-1)] = 1 + \xi(2j) + \frac{zk}{m_j} \tag{14.18}$$

relating the probability density functions for $\xi(2j)$ and $\xi(2j+1)$, which will in general be different. Once these probability density functions are known as solutions of the above equations, the function $\Omega(z)$, and hence $D(\mu)$, can be calculated.

14.2. Bordered Matrices†

Let us consider only the real symmetric matrices. The diagonal elements are integers ..., $-2, -1, 0, 1, 2, \ldots$. The elements H_{jk} for which $|j-k| > m$ are zero, whereas the elements H_{jk} with $|j-k| \leqslant m$ all have the same magnitude h:

$$\begin{aligned} H_{jk} &= \pm h, \quad \text{if} \quad |j-k| \leqslant m, \\ &= 0, \qquad \text{if} \quad |j-k| > m. \end{aligned} \tag{14.19}$$

Subject to the symmetry condition $H_{jk} = H_{kj}$, the signs of the H_{jk} are random. Let us denote the eigenvalues and the normalized eigenvectors of H by λ and $\psi^{(\lambda)}$, respectively:

$$H\psi^{(\lambda)} = \lambda\psi^{(\lambda)} \quad \text{or} \quad \sum_k H_{jk}\psi_k^{(\lambda)} = \lambda\psi_j^{(\lambda)}; \tag{14.20}$$

† Wigner [5].

and for a λ lying between x and $x + \delta x$ find the expectation value of

$$(\psi_0^{(\lambda)})^2, \tag{14.21}$$

where $\psi_0^{(\lambda)}$ is a particular component of $\psi^{(\lambda)}$. This expectation value will be written as $\sigma(x)\,\delta x$, where $\sigma(x)$ is named "the strength function." As the absorption of an energy-level depends, under certain conditions, only on the square of a definite component of the corresponding eigenstate, the function $\sigma(x)$ represents the strength of absorption around the energy value x. The problem is to find $\sigma(x)$ and the distribution of the eigenvalues λ.

When there is a single border, $m = 1$ (i.e., when H_{jk} are zero for $|j - k| > 1$), the problem is simple enough for an explicit evaluation of the eigenvalues and the eigenvectors. Such an H can be transformed by a diagonal matrix S with diagonal elements ± 1 to a matrix H_1 and the signs of the diagonal elements of S can be so chosen that all the off-diagonal elements of H_1 will have the negative sign, whereas the diagonal elements are the same as in H. In so doing neither the eigenvalues nor the squares of the components of the eigenvectors undergo any change. The resulting matrix H_1 can be transformed to $-H_1$ by interchanging jth and $-j$th row and column and by transforming with an S whose diagonal elements are alternately $+1$ and -1. Furthermore, H_1 can be changed to $H_1 + 1$ by renumbering the rows and columns (the dimension of H should be infinite for the argument to apply here). Thus, along with λ_k, $-\lambda_k$ and $\lambda_k + 1$ are also eigenvalues. By the continuity in h and the condition that for $h = 0$ the eigenvalues of H are all integers we see that the eigenvalues of H always consist of all the integers:

$$\lambda_k = k, \qquad k = 0, \pm 1, \pm 2, \ldots . \tag{14.22}$$

Denoting the corresponding eigenvector by $\psi^{(k)}$, one sees from the above remarks about changing H_1 to $H_1 + 1$ that $\psi_l^{(k)} = \psi_{l+1}^{(k+1)}$ that is, that the lth component of $\psi^{(k)}$ depends only on the difference $(l - k)$:

$$\psi_l^{(k)} = \psi_{l-k}^{(0)}. \tag{14.23}$$

Again from the remarks about changing H_1 to $-H_1$, it follows that

$$\psi_{-l}^{(0)} = c(-1)^l \psi_l^{(0)},$$

where $c = \pm 1$. The continuity from $h = 0$ again gives $c = 1$. The equation $H_1\psi = \lambda\psi$ can now be written as

$$-h\psi_{l-1}^{(0)} + l\psi_l^{(0)} - h\psi_{l+1}^{(0)} = 0, \tag{14.24}$$

and compared with the recursion formula for the Bessel functions [Bateman, 2]

$$-J_{l-1}(z) + \frac{2l}{z} J_l(z) - J_{l+1}(z) = 0. \tag{14.25}$$

We infer that

$$\psi_l^{(0)} = J_l(2h) = \psi_{l+k}^{(k)}. \tag{14.26}$$

The irregular Bessel functions also satisfy (14.25), but if the eigenvector components are taken as a linear combination of the regular and irregular Bessel functions then one cannot normalize.

The case of thick borders, $m \gg 1$, and large off-diagonal elements, $h \gg 1$, has been treated by Wigner under the condition that $h^2/m = q$ remains constant. Some of the remarks that apply for a singly bordered matrix also apply here, and from such considerations it can be deduced that $\sigma(x)$ is an even function. The average number of eigenvalues per unit interval at x is a periodic function of x with the period 1. Wigner then calculates the moments for the strength function

$$M_{2\nu} = \int_{-\infty}^{\infty} x^{2\nu}\, \sigma(x)\, dx \tag{14.27}$$

and derives an integral equation for $\sigma(x)$. Because the calculations are long, they are not reproduced here. Putting

$$\rho(\xi) = \lim_{m\to\infty} m\, \sigma(m\xi),$$

the final result is the integral equation

$$\rho(\xi) = q\, R_1(\xi) \int_{-\infty}^{\infty} \frac{\rho(x) - \rho(\xi)}{\xi^2 - x^2}\, dx + q\, \rho(\xi) \int_{-\infty}^{\infty} \frac{R_1(x) - R_1(\xi)}{\xi^2 - x^2}\, dx, \tag{14.28}$$

where

$$R_1(x) = \int_{x-1}^{x+1} \rho(\xi)\, d\xi, \qquad q = \frac{h^2}{m}, \tag{14.29}$$

$$R_1(0)\, \rho(0) = (\pi^2 q)^{-1}. \tag{14.30}$$

In the limiting case $q \to \infty$ we have

$$\begin{aligned} \rho(\xi) &\simeq (4\pi q)^{-1}(8q - \xi^2)^{1/2}, \qquad \xi^2 < 8q \\ &\simeq 0, \qquad\qquad\qquad\quad \xi^2 > 8q; \end{aligned} \tag{14.31}$$

whereas in the opposite limiting case $q \to 0$ we have

$$\rho(\xi) \simeq \frac{q}{\pi^2 q^2 + \xi^2}, \qquad \xi \ll 1. \tag{14.32}$$

However, for $\xi \to \infty$ we have

$$\rho(\xi) \simeq (\text{constant})(2q\xi^2 \ln \xi)^{-\xi}\, e^{2\xi}. \tag{14.33}$$

15 / Invariance Hypothesis and Matrix Element Correlations

The entire theory of Gaussian ensembles is based on the two assumptions put forward in Chapter 2:

1. The ensemble is statistically invariant under a change of basis.

2. The matrix elements are statistically independent, hence uncorrelated.

As mentioned in that chapter, Assumption 1 is quite natural, whereas Assumption 2 is somewhat artificially introduced to simplify the calculations.

There have been efforts [Ullah, 1; Ullah and Porter, 1] to determine what kind of correlations among the various matrix elements are implied by Assumption 1. We shall see here that it leads to the vanishing of the ensemble averages of the following quantities:

1. Any odd power of an off-diagonal element.

2. The product of an odd power of an off-diagonal element and any power of a diagonal or another off-diagonal element.

We illustrate the method in relation to the orthogonal ensemble, though it is equally applicable to the unitary or the symplectic ensemble. Thus we consider a set of real symmetric matrices that is statistically invariant under real orthogonal transformations

$$H = H^* = H^T,$$

$$P(H)\,dH = P(H')\,dH', \quad \text{if} \quad H' = RHR^{-1},\ RR^T = R^TR = 1.$$

Choosing some complete set of basic functions, the eigenvalue equation can be written as

$$\sum_\nu H_{\mu\nu}a_{\lambda\nu} = \theta_\lambda a_{\lambda\mu}\,, \tag{15.1}$$

where θ_λ is the eigenvalue and $a_{\lambda\mu}$ is the μth component of the corresponding eigenvector. The eigenvectors are orthogonal and form a complete set:

$$\sum_\mu a_{\lambda\mu}a_{\xi\mu} = \sum_\mu a_{\mu\lambda}a_{\mu\xi} = \delta_{\lambda\xi}\,. \tag{3.6}$$

Using (3.6) and (15.1), we can express the matrix elements as

$$H_{\mu\nu} = \sum_\lambda \theta_\lambda a_{\lambda\mu}a_{\lambda\nu}\,. \tag{3.5}$$

Now because the joint probability density function $P(H)$ is invariant under orthogonal transformations of the basis, it must depend essentially only on the eigenvalues θ_λ and that, too, in a symmetric manner. The discussion leading to (3.16) is valid, and the joint probability density function can be written as a product of functions depending on mutually exclusive sets of variables:

$$P(\theta_\lambda\,, p) = P(\theta_\lambda)f(p). \tag{15.2}$$

We are interested in the averages of products of the matrix elements given by (3.5). Because of the separable nature of (15.2), this averaging can be done separately over the θ_λ and the parameters p to get the result

$$\langle H_{\mu\nu}H_{\xi\eta}\cdots\rangle = \sum_{\lambda_1,\lambda_2,\cdots=1}^{N} \langle\theta_{\lambda_1}\theta_{\lambda_2}\cdots\rangle\langle a_{\lambda_1\mu}a_{\lambda_1\nu}a_{\lambda_2\xi}a_{\lambda_2\eta}\cdots\rangle. \tag{15.3}$$

Thus we have to find the averages of the products of components of a set of orthogonal unit vectors randomly oriented in the N-dimensional space. These averages may be written as the ratio of two integrals:

$$\langle Q(\mathbf{a}_{\lambda_1}\,, \mathbf{a}_{\lambda_2}\cdots)\rangle = \frac{\mathscr{N}}{\mathscr{D}}\,, \tag{15.4}$$

where

$$\mathscr{N} = \int_{-\infty}^{\infty}\cdots\int Q\,\delta\Big(\sum_{\mu=1}^{N} a_{\lambda_1\mu}^2 - 1\Big)\,\delta\Big(\sum_{\mu=1}^{N} a_{\lambda_1\mu}a_{\lambda_2\mu}\Big)$$

$$\times\,\delta\Big(\sum_{\mu=1}^{N} a_{\lambda_2\mu}^2 - 1\Big)\cdots\prod_{\mu=1}^{N} da_{\lambda_1\mu}\,da_{\lambda_2\mu}\cdots \tag{15.5}$$

and $\mathscr{D}$ is the same integral without the Q inside.

The integral $\mathcal{N}$ can be evaluated when it involves only one or two vectors, whereas its evaluation in general is not at all easy. However, from the symmetry arguments we can conclude that

$$\langle a_{\lambda_1\mu} a_{\lambda_1\nu} \cdots a_{\lambda_{2m+1}\mu} a_{\lambda_{2m+1}\nu} \rangle = 0, \qquad \mu \neq \nu, \tag{15.6}$$

implying that $\langle H_{\mu\nu}^{2m+1} \rangle = 0$. By similar reasonings we can convince ourselves that

$$\langle H_{\mu\nu}^{2m+1} H_{\xi\eta}^{r} \rangle = 0 \tag{15.7}$$

if $\mu \neq \nu$ and the pair $(\mu\nu)$ is distinct from $(\xi\eta)$ or $(\eta\xi)$.

We can actually evaluate the ratio (15.4) for a simple expression Q, the average of the product of two matrix elements, for example [Ullah, 1]. Let us first take only one N-dimensional unit vector with random components $u_1, u_2, \ldots, u_N$. Equation 15.4 then gives

$$\langle Q(\mathbf{u}) \rangle = \frac{\int_{-\infty}^{\infty} \cdots \int Q(\mathbf{u})\, \delta\left(\sum_1^N u_i^2 - 1\right) \prod_1^N du_i}{\int_{-\infty}^{\infty} \cdots \int \delta\left(\sum_1^N u_i^2 - 1\right) \prod_1^N du_i}. \tag{15.8}$$

For Q we substitute $\prod_{i=1}^N u_i^{2m_i}$ and evaluate the integral

$$I = \int_{-\infty}^{\infty} \cdots \int \delta\left(\sum_1^N u_i^2 - 1\right) \prod_1^N u_i^{2m_i}\, du_i. \tag{15.9}$$

Replacement of u_i with $u_i/\sqrt{r}$ gives

$$I.r^{1/2N+\Sigma_i m_i - 1} = \int_{-\infty}^{\infty} \cdots \int \delta\left(\sum_1^N u_i^2 - r\right) \prod_1^N u_i^{2m_i}\, du_i. \tag{15.10}$$

Multiplying on both sides with e^{-r} and integrating (first!) over r from 0 to ∞, and then over the u_i, we get

$$\begin{aligned} I.\Gamma\left(\tfrac{1}{2}N + \sum_i m_i\right) &= \int_{-\infty}^{\infty} \cdots \int \exp\left(-\sum_i u_i^2\right) \prod_i u_i^{2m_i}\, du_i \\ &= \prod_{i=1}^N \left(\int_{-\infty}^{\infty} e^{-u^2} u^{2m_i}\, du\right) \\ &= \prod_{i=1}^N \Gamma(m_i + \tfrac{1}{2}). \end{aligned} \tag{15.11}$$

To get the denominator we put $m_i = 0$, $i = 1, 2, \ldots, N$. Thus the average of $\prod_{i=1}^{N} u_i^{2m_i}$ is

$$\left\langle \prod_{i=1}^{N} u^{2m_i} \right\rangle = \frac{\prod_{i=1}^{N} \Gamma(m_i + \frac{1}{2})}{\Gamma\left(\frac{1}{2}N + \sum_{i=1}^{N} m_i\right)} \frac{\Gamma(\frac{1}{2}N)}{[\Gamma(\frac{1}{2})]^N}. \tag{15.12}$$

Extending this method to two vectors $\mathbf{u} = (u_1, \ldots, u_N)$ and $\mathbf{v} = (v_1, \ldots, v_N)$, we can write

$$Q(\mathbf{u}, \mathbf{v}) = \frac{\int_{-\infty}^{\infty} \cdots \int Q\, \delta\left(\sum_i u_i^2 - 1\right) \delta\left(\sum_i u_i v_i\right) \delta\left(\sum_i v_i^2 - 1\right) \prod_i du_i\, dv_i}{\int_{-\infty}^{\infty} \cdots \int \delta\left(\sum_i u_i^2 - 1\right) \delta(u_i v_i)\, \delta\left(\sum_i v_i^2 - 1\right) \prod_i du_i\, dv_i}$$

$$= \frac{\mathcal{N}}{\mathcal{D}}. \tag{15.13}$$

Let us calculate the denominator. The evaluation of the numerator is similar, through somewhat lengthy. Replacing u_i with $u_i/\sqrt{r_1}$ and v_i with $v_i/\sqrt{r_2}$, we obtain

$$\mathcal{D}.(r_1 r_2)^{(1/2)(N-3)} = \int_{-\infty}^{\infty} \cdots \int \delta\left(\sum_i u_i^2 - r_1\right) \delta\left(\sum_i u_i v_i\right) \delta\left(\sum_i v_i^2 - r_2\right) \prod_i du_i\, dv_i. \tag{15.14}$$

Multiplying by $e^{-r_1} e^{-r_2}$ and integrating on r_1, r_2 from 0 to ∞ as before, we now have

$$\mathcal{D}.[\Gamma(\tfrac{1}{2}N - \tfrac{1}{2})]^2 = \int_{-\infty}^{\infty} \cdots \int \exp\left(-\sum_i (u_i^2 + v_i^2)\right) \delta\left(\sum_i u_i v_i\right) \prod_i du_i\, dv_i. \tag{15.15}$$

We substitute

$$u_i = \frac{1}{\sqrt{2}}(p_i + q_i), \qquad v_i = \frac{1}{\sqrt{2}}(p_i - q_i)$$

to get

$$\mathcal{D}.[\Gamma(\tfrac{1}{2}N - \tfrac{1}{2})]^2 = 2 \int_{-\infty}^{\infty} \cdots \int \exp\left[-\sum_i (p_i^2 + q_i^2)\right] \delta\left[\sum_i (p_i^2 - q_i^2)\right] \prod_i dp_i\, dq_i. \tag{15.16}$$

Introducting the spherical polar coordinates and integrating over the angles the right hand side gives

$$2(NV_N)^2 \int_0^\infty \cdots \int e^{-(p^2+q^2)}\, \delta(p^2 - q^2)(pq)^{N-1}\, dp\, dq, \tag{15.17}$$

where

$$V_N = \pi^{(1/2)N}[\Gamma(\tfrac{1}{2}N + 1)]^{-1} = \frac{2}{N}\pi^{(1/2)N}[\Gamma(\tfrac{1}{2}N)]^{-1} \tag{15.18}$$

is the volume of the N-dimensional unit sphere. The remaining integrals are elementary. We finally get

$$\begin{aligned}\mathscr{D} &= (NV_N)^2\, \Gamma(N-1)(\tfrac{1}{2})^N[\Gamma(\tfrac{1}{2}N - \tfrac{1}{2})]^{-2} \\ &= 2^{N-2}\pi^{N-1}[\Gamma(N-1)]^2.\end{aligned} \tag{15.19}$$

After lengthy algebra, Ullah obtained a complicated expression for the average of $\prod_i u_i^{2m_i} v_i^{2n_i}$. To get further results by this method for three or more vectors seems to be extremely difficult.

Such calculations may be used to compute some simple correlations; for example, the correlation coefficient of two diagonal elements [Ullah and Porter, 1]

$$C_{\mu\nu} = \frac{\langle \delta H_\mu\, \delta H_\nu \rangle}{[\langle (\delta H_\mu)^2 \rangle \langle (\delta H_\nu)^2 \rangle]^{1/2}}, \qquad \delta H_\mu = H_{\mu\mu} - \langle H_{\mu\mu} \rangle \tag{15.20}$$

can be obtained from the following averages

$$\langle H_{\mu\mu} \rangle = \langle \theta_1 \rangle, \qquad \langle H_{\mu\mu}^2 \rangle = \frac{1}{N+2}[3\langle \theta_1^2 \rangle + (N-1)\langle \theta_1 \theta_2 \rangle],$$

$$\langle H_{\mu\mu} H_{\nu\nu} \rangle = \frac{1}{N+2}[\langle \theta_1 \rangle^2 + (N+1)\langle \theta_1 \theta_2 \rangle], \qquad \mu \neq \nu, \tag{15.21}$$

so that

$$C_{\mu\nu} = \frac{1 + (N+1)\, C}{3 + (N-1)\, C}, \tag{15.22}$$

where C is the correlation coefficient of two eigenvalues:

$$C = \frac{\langle \delta\theta_1\, \delta\theta_2 \rangle}{[\langle (\delta\theta_1)^2 \rangle \langle (\delta\theta_2)^2 \rangle]^{1/2}}, \qquad \delta\theta_1 = \theta_1 - \langle \theta_1 \rangle. \tag{15.23}$$

16 / The Joint Probability Density Functions for Two Nearby Spacings

Integrating an expression like

$$P_{N1}(x_1, \ldots, x_N) = C_{N1} \exp\left[-\frac{1}{2}\sum_1^N x_j^2\right] \prod_{j<k} |x_j - x_k| \tag{3.18}$$

over all the variables except a few, though usually very tedious, is sometimes possible. A case worthy of demonstration is the following calculation. Fix one of the eigenvalues at α, and integrate all other eigenvalues outside the interval $(-\theta, \theta)$, where $|\alpha| < \theta$. Putting $\theta + \alpha = S_1$ and $\theta - \alpha = S_2$ will give us the probability that there are no eigenvalues for a distance S_1 on one side and a distance S_2 on the other side of a given eigenvalue. Differentiating partially with respect to S_1 and S_2 will then give us the joint probability density function for two nearby spacings.

16.1. Integrations

As usual, the unitary ensemble is mathematically the simplest to work with. However, we will consider the orthogonal (Gaussian) ensemble, which is more useful in applications. As in Chapter 5 we express $P_{N1}(x_1, \ldots, x_N)$, equation (3.18), as a determinant of normalized oscillator wave functions:

$$P_{N1}(x_1, \ldots, x_N) = c\{\det[\varphi_{j-1}(x_j)]_{j=1,2,\ldots,N}\},$$

$$\varphi_j(x) = (2^{-j} j!\, \sqrt{\pi})^{-1/2} e^{1/2x^2} \left(-\frac{d}{dx}\right)^j e^{-x^2}, \tag{5.15}$$

where c is a known constant and the ordering $x_1 \leqslant x_2 \leqslant \cdots \leqslant x_N$ is understood.

Because the variables are ordered, to fix one of them at α we should write

$$A_N(\theta, \alpha) = \int \cdots \int P_{N1}(\alpha, x_2, \ldots, x_N)\, dx_2 \cdots dx_N$$

$$= c \left\{ \int \cdots \int \{\det[\varphi_{j-1}(x_j)]_{j=1,\ldots,N}\}_{x_1=\alpha}\, dx_2 \cdots dx_N \right.$$

$$\left. + \int \cdots \int \{\det[\varphi_{j-1}(x_j)]_{j=1,\ldots,N}\}_{x_2=\alpha}\, dx_1\, dx_3 \cdots dx_N + \cdots + \cdots \right\}$$

$$= \mathcal{O}(\theta, \alpha) + \mathcal{E}(\theta, \alpha), \tag{16.1}$$

where $\mathcal{O}(\theta, \alpha)$ includes the integrals in which an x with an odd index is put equal to α and $\mathcal{E}(\theta, \alpha)$ includes the other terms. The integrations are supposed to be carried out on the ordered variables $-\infty < x_1 \leqslant x_2 \leqslant \cdots \leqslant x_N < \infty$ and to exclude the interval $(-\theta, \theta)$. To avoid minor complications we will take N even, $N = 2m$. Introducing the functions

$$u(x) = \begin{cases} 1, & \text{if } |x| > \theta, \\ 0, & \text{if } -\theta < x < \theta, \end{cases} \tag{16.2}$$

$$F_i(x) = \int_{-\infty}^{x} \varphi_i(y)\, u(y)\, dy \tag{16.3}$$

and

$$\epsilon(x) = \begin{cases} 1, & \text{if } x > \theta, \\ 0, & \text{if } x < \theta, \end{cases} \tag{16.4}$$

an integration over the alternate variables (cf. Chapter 5) readily gives

$$\mathcal{O}(\theta, \alpha) = -2^{-(3/2)m}(2m)^{-1} \prod_{j=1}^{m} (2j-1)^{1/2} \cdot \int_{-\infty}^{\infty} \cdots \int \prod_{j=1}^{m} u(x_{2j})\, dx_{2j}$$

$$\cdot \det \begin{bmatrix} 0 & \epsilon(x_{2j}) & 0 \\ \varphi_i(\alpha) & F_i(x_{2j}) & \varphi_i(x_{2j}) \end{bmatrix}_{\substack{j=1,2,\ldots,m \\ i=0,1,\ldots,2m-1}} \tag{16.5}$$

We can now drop the ordering and integrate over the remaining variables from $-\infty$ to ∞. This gives, as in Chapter 5,

$$\mathcal{O}(\theta, \alpha) = \mathcal{O}_{\text{even}}(\theta, \alpha) + \mathcal{O}_{\text{odd}}(\theta, \alpha) \tag{16.6}$$

where

$$\mathcal{O}_{\text{even}}(\theta, \alpha) = \mathcal{O}_{\text{even}}(\theta, -\alpha)$$
$$= -2^{-(3/2)m}(2m)^{-1} \prod_{j=1}^{m} (2j-1)^{1/2} \det\left[\begin{matrix} 0 & \rho_{2j+1}(\theta) \\ \varphi_{2i}(\alpha) & f'_{2i,2i+1}(\theta) \end{matrix}\right]_{i,j=0,1,\ldots,m-1} \tag{16.7}$$

and

$$\mathcal{O}_{\text{odd}}(\theta, \alpha) = -\mathcal{O}_{\text{odd}}(\theta, -\alpha)$$
$$= 2^{-(3/2)m}(2m)^{-1} \prod_{j=1}^{m} (2j-1)^{1/2} \det\left[\begin{matrix} 0 & \rho_{2j}(\theta) \\ \varphi_{2i+1}(\alpha) & f'_{2j,2i+1}(\theta) \end{matrix}\right]_{i,j=0,1,\ldots,m-1} \tag{16.8}$$

where we have introduced the functions

$$\rho_j(\theta) = \int_{-\infty}^{\infty} \epsilon(x)\, u(x)\, \varphi_j(x)\, dx = \int_{\theta}^{\infty} \varphi_j(x)\, dx \tag{16.9}$$

and

$$f'_{2i,2j+1}(\theta) = \int_{-\infty < y \leqslant x < \infty} dx\, dy\, u(x)\, u(y)\{\varphi_{2i}(y)\, \varphi_{2j+1}(x) - \varphi_{2i}(x)\, \varphi_{2j+1}(y)\}. \tag{16.10}$$

As in Chapter 5, we can use (5.25)

$$f'_{2i,2j+1} - \left(\frac{2j}{2j+1}\right)^{1/2} f'_{2i,2j-1} = \left(\frac{8}{2j+1}\right)^{1/2} g_{ij} \tag{16.11}$$

and

$$\rho_{2j+1} - \left(\frac{2j}{2j+1}\right)^{1/2} \rho_{2j-1} = \left(\frac{2}{2j+1}\right)^{1/2} \varphi_{2j}(\theta) \tag{16.12}$$

to reduce (16.7) to the form

$$\mathcal{O}_{\text{even}}(\theta, \alpha) = -\frac{1}{4m} \det\left[\begin{matrix} 0 & \varphi_{2j}(\theta) \\ \varphi_{2i}(\alpha) & g_{ij}(\theta) \end{matrix}\right]_{i,j=0,1,\ldots,m-1}, \tag{16.13}$$

where

$$g_{ij} = \delta_{ij} - \int_{-\theta}^{\theta} \varphi_{2i}(x)\, \varphi_{2j}(x)\, dx. \tag{16.14}$$

As for $\mathscr{E}(\theta, \alpha)$, introducing the functions

$$F_i'(x) = \int_x^\infty \varphi_i(y)\, u(y)\, dy, \tag{16.15}$$

$$\epsilon'(x) = \begin{cases} 0, & \text{if } x > -\theta, \\ 1, & \text{if } x < -\theta, \end{cases} \tag{16.16}$$

one can integrate over the variables with even suffixes. A change of variables to their negatives gives then

$$\mathscr{E}(\theta, \alpha) = \mathscr{O}(\theta, -\alpha), \tag{16.17}$$

so that

$$\mathscr{O}(\theta, \alpha) + \mathscr{E}(\theta, \alpha) = 2\mathscr{O}_{\text{even}}(\theta, \alpha). \tag{16.18}$$

Putting together (16.1), (16.18) and (16.13) we get

$$A_{2m}(\theta, \alpha) = -\frac{1}{2m} \det \begin{bmatrix} 0 & \varphi_{2j}(\theta) \\ \varphi_{2i}(\alpha) & g_{ij}(\theta) \end{bmatrix}_{i,j=0,1,\ldots,m-1}, \tag{16.19}$$

with $g_{ij}(\theta)$ given by (16.14).

16.2. An Integral Equation with a Boundary Condition and $A_{2m}(\theta, \alpha)$

Let us write (16.19) in a slightly different form:

$$2m \frac{\pi}{2\sqrt{m}} A_{2m}(\theta, \alpha) = E_m(\theta) - A_m^{(1)}(\theta, \alpha) \tag{16.20}$$

where

$$E_m(\theta) = \det[g_{ij}(\theta)]_{i,j=0,1,\ldots,m-1} \tag{16.21}$$

and

$$A_m^{(1)}(\theta, \alpha) = \frac{\pi}{2\sqrt{m}} \det \begin{bmatrix} \frac{2\sqrt{m}}{\pi} & \varphi_{2j}(\theta) \\ \varphi_{2i}(\alpha) & g_{ij}(\theta) \end{bmatrix}_{i,j=0,1,\ldots,m-1} \tag{16.22}$$

In Chapter 5 we expressed $E_m(\theta)$ as the product

$$E_m(\theta) = \prod_{i=0}^{m-1} (1 - \lambda_{2i}) \tag{16.23}$$

where λ_{2i} are the eigenvalues of the homogeneous integral equation (5.88),

$$\lambda f(x) - \int_{-\theta}^{\theta} K_m(x, y) f(y)\, dy = 0, \tag{16.24}$$

whose solutions are even functions

$$f_{2j}(-x) = f_{2j}(x), \qquad j = 0, 1, \ldots, m-1 \tag{16.25}$$

and the kernel $K_m(x, y)$ is given by (5.89)

$$K_m(x, y) = \sum_{i=0}^{m-1} \varphi_{2i}(x)\, \varphi_{2i}(y). \tag{16.26}$$

One can find a similar expression for $A_m^{(1)}(\theta, \alpha)$ by considering this time the inhomogeneous equation

$$\mu g(x) - \int_{-\theta}^{\theta} K_m(x, y)\, g(y)\, dy = \frac{\pi}{2\sqrt{m}} K_m(x, \alpha). \tag{16.27}$$

Given μ, not equal to one of the discrete values λ_i for which the homogeneous equation (16.24) has a nonzero solution, the integral equation (16.27) has a solution and this solution is unique. It is even and has the form

$$g(x) = \sum_{i=0}^{m-1} c_i \varphi_{2i}(x) \tag{16.28}$$

where the constants c_i are determined by the simultaneous linear equations

$$\sum_{j=0}^{m-1} \left\{ \mu\, \delta_{ij} - \int_{-\theta}^{\theta} \varphi_{2i}(y)\, \varphi_{2j}(y)\, dy \right\} c_j = \frac{\pi}{2\sqrt{m}} \varphi_{2i}(\alpha), \qquad i = 0, 1, \ldots, m-1. \tag{16.29}$$

Let us impose the extra restriction

$$g(\theta) \equiv \sum_{i=0}^{m-1} c_i \varphi_{2i}(\theta) = 1, \tag{16.30}$$

and eliminate the c_i between (16.29) and (16.30). Thus (16.30) will be satisfied if and only if

$$\det \begin{bmatrix} 1 & \varphi_{2j}(\theta) \\ \dfrac{\pi}{2\sqrt{m}} \varphi_{2i}(\alpha) & \mu\, \delta_{ij} - \displaystyle\int_{-\theta}^{\theta} \varphi_{2i}(y)\, \varphi_{2j}(y)\, dy \end{bmatrix} = 0. \tag{16.31}$$

The algebraic equation (16.31) has m roots; let them be

$$\mu_0, \mu_2, \cdots, \mu_{2m-2}.$$

Therefore we have the identity

$$\frac{\pi}{2\sqrt{m}} \det \begin{bmatrix} \frac{2\sqrt{m}}{\pi} & \varphi_{2j}(\theta) \\ \varphi_{2i}(\alpha) & \mu\,\delta_{ij} - \int_{-\theta}^{\theta} \varphi_{2i}(y)\,\varphi_{2j}(y)\,dy \end{bmatrix} = \prod_{i=0}^{m-1} (\mu - \mu_{2i}). \quad (16.32)$$

From (16.20), (16.23), (16.22) and (16.32) we thus have

$$\pi\sqrt{m}\, A_{2m}(\theta, \alpha) = \prod_{i=0}^{m-1} (1 - \lambda_{2i}) - \prod_{i=0}^{m-1} (1 - \mu_{2i}) \quad (16.33)$$

where λ_{2i} are the eigenvalues of (16.24) and the μ_{2i} are the "eigenvalues" of (16.27) with the boundary condition (16.30).

16.3. The Limit of $A_{2m}(\theta, \alpha)$

Let us define

$$t = \frac{4\theta\sqrt{m}}{\pi}, \qquad \tau = \frac{4\alpha\sqrt{m}}{\pi}, \quad (16.34)$$

and take the limit as $m \to \infty$, $\theta \to 0$ and $\alpha \to 0$ in such a way that t and τ are finite. The limit of (16.26) is (cf. Appendix A.9)

$$K_m(x, y) = \frac{2\sqrt{m}}{\pi}\, Q(\xi, \eta), \quad (16.35)$$

$$\xi = \frac{2\sqrt{m}}{\pi}\, x, \qquad y = \frac{2\sqrt{m}}{\pi}\, y \quad (16.36)$$

where

$$Q(\xi, \eta) = \frac{1}{2}\left\{\frac{\sin(\xi + \eta)\,\pi}{(\xi + \eta)\,\pi} + \frac{\sin(\xi - \eta)\,\pi}{(\xi - \eta)\,\pi}\right\}; \quad (16.37)$$

and the equations (16.24), (16.27) and (16.30) take the form

$$\lambda f_1(\xi) - \int_{-\frac{1}{2}t}^{\frac{1}{2}t} Q(\xi, \eta)\, f_1(\eta)\, d\eta = 0, \quad (16.38)$$

$$\mu g_1(\xi) - \int_{-\frac{1}{2}t}^{\frac{1}{2}t} Q(\xi, \eta)\, g_1(\eta)\, d\eta = Q(\xi, \tfrac{1}{2}\tau), \quad (16.39)$$

and

$$g_1(\tfrac{1}{2}t) = 1 \tag{16.40}$$

where

$$f_1(\xi) = f\left(\frac{\pi\xi}{2\sqrt{m}}\right), \qquad g_1(\xi) = g\left(\frac{\pi\xi}{2\sqrt{m}}\right). \tag{16.41}$$

By a change of scale, $\xi = \frac{1}{2}t\xi'$, $\eta = \frac{1}{2}t\eta'$, the limits of integration in (16.38), (16.39) and (16.40), can be brought to $(-1, 1)$:

$$\lambda\tilde{f}(x) - \int_{-1}^{1} \tilde{Q}(x, y)\tilde{f}(y)\, dy = 0 \tag{16.42}$$

$$\mu\tilde{g}(x) - \int_{-1}^{1} \tilde{Q}(x, y)\,\tilde{g}(y)\, dy = \frac{2}{t}\tilde{Q}\left(x, \frac{\tau}{t}\right) \tag{16.43}$$

and

$$\tilde{g}(1) = 1 \tag{16.44}$$

where

$$\tilde{Q}(x, y) = \frac{1}{2}\left\{\frac{\sin(x + y)\frac{1}{2}t\pi}{(x + y)\pi} + \frac{\sin(x - y)\frac{1}{2}t\pi}{(x - y)\pi}\right\}, \tag{16.45}$$

$$\tilde{f}(x) = f_1(\tfrac{1}{2}tx), \qquad \tilde{g}(x) = g_1(\tfrac{1}{2}tx). \tag{16.46}$$

As in Chapter 5, the eigenfunctions of (16.42) are even prolate spheroidal functions, the solutions of the differential equation

$$\left\{(x^2 - 1)\frac{d^2}{dx^2} + 2x\frac{d}{dx} + \frac{\pi^2}{4}t^2x^2\right\}\tilde{f}(x) = l\tilde{f}(x) \tag{5.107}$$

and are extensively tabulated [Stratton *et al.*, 1]. From a knowledge of the eigenfunctions one can calculate the eigenvalues λ as in Chapter 5. Let us assume then that the normalized spheroidal functions $\tilde{f}_{2j}(x)$ and the corresponding eigenvalues λ_{2j} are known. The functions $\tilde{f}_{2j}$ form a complete set for even functions integrable over $(-1, 1)$. An expansion in terms of them gives

$$\tilde{Q}(x, y) = \sum_{j=0}^{\infty} \lambda_{2j}\tilde{f}_{2j}(x)\tilde{f}_{2j}(y), \tag{16.47}$$

$$\tilde{g}_{2i}(x) = \sum_{i=0}^{\infty} c_{ji}\tilde{f}_{2i}(x). \tag{16.48}$$

To determine the c_{ji} we substitute (16.47) and (16.48) in (16.43) and use the orthonormal properties of the $\tilde{f}_{2i}(x)$:

$$\int_{-1}^{1} \tilde{f}_{2i}(x)\, \tilde{f}_{2j}(x)\, dx = \delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j \end{cases} \tag{16.49}$$

to get

$$(\mu_{2j} - \lambda_{2i})\, c_{ji} = \frac{2}{t}\, \lambda_{2i}\, \tilde{f}_{2i}\left(\frac{\tau}{t}\right) \tag{16.50}$$

or in view of (16.48),

$$\tilde{g}_{2j}(x) = \sum_{i=0}^{\infty} \frac{2}{t} \frac{\lambda_{2i}}{\mu_{2j} - \lambda_{2i}} \tilde{f}_{2i}\left(\frac{\tau}{t}\right) \tilde{f}_{2i}(x). \tag{16.51}$$

Substituting (16.51) in (16.44) we get an equation whose roots are the eigenvalues μ_{2j} :

$$\frac{2}{t} \sum_{i=0}^{\infty} \frac{\lambda_{2i}}{\mu - \lambda_{2i}} \tilde{f}_{2i}\left(\frac{\tau}{t}\right) \tilde{f}_{2i}(1) = 1. \tag{16.52}$$

Multiplying by $\prod_i (\mu - \lambda_{2i})$ to rationalize the denominators in (16.52), one gets

$$\prod_i (\mu - \mu_{2i}) \equiv \prod_i (\mu - \lambda_{2i}) - \left\{\prod_i (\mu - \lambda_{2i})\right\} \left\{\sum_i \frac{(2/t)\, \lambda_{2i}}{\mu - \lambda_{2i}}\right\} \tilde{f}_{2i}\left(\frac{\tau}{t}\right) \tilde{f}_{2i}(1). \tag{16.53}$$

To get (16.33) we put $\mu = 1$ in the identity (16.53)

$$\begin{aligned} A(t, \tau) &= \lim_{m \to \infty} \pi \sqrt{m}\, A_{2m}\left(\frac{\pi t}{4\sqrt{m}}, \frac{\pi \tau}{4\sqrt{m}}\right) \\ &= \left\{\prod_i (1 - \lambda_{2i})\right\} \cdot \left\{\sum_i \frac{(2/t)\, \lambda_{2i}}{1 - \lambda_{2i}} \tilde{f}_{2i}\left(\frac{\tau}{t}\right) \tilde{f}_{2i}(1)\right\} \end{aligned} \tag{16.54}$$

Substituting

$$x_1 = \tfrac{1}{2}(t + \tau), \qquad x_2 = \tfrac{1}{2}(t - \tau) \tag{16.55}$$

we get the probability that there are no eigenvalues for a distance x_1 on one side and a distance x_2 on the other side of a given eigenvalue

$$\mathscr{B}(x_1, x_2) = A((x_1 + x_2), |(x_1 - x_2)|) \tag{16.56}$$

where x_1 and x_2 are measured in units of the mean spacing at the origin. The joint probability density function $\mathscr{P}(x_1, x_2)$ for two adjacent spacings x_1 and x_2 is obtained by differentiating $\mathscr{B}(x_1, x_2)$ twice (cf. Appendix A.11):

$$\mathscr{P}(x_1, x_2) = \frac{\partial^2}{\partial x_1 \partial x_2} \mathscr{B}(x_1, x_2). \tag{16.57}$$

16.4. Power Series Expansion and Numerical Results

For small t the spheroidal functions can be expanded in terms of the Legendre polynomials. Thus the first few terms in the expansion of λ_{2i}, f_{2i}, $A(t, \tau)$ and $\mathscr{B}(x_1, x_2)$ as a power series are

$$\lambda_0 = t - \frac{\pi^2}{36} t^3 + \frac{23\pi^4}{32400} t^5 + 0(t^7), \tag{16.58}$$

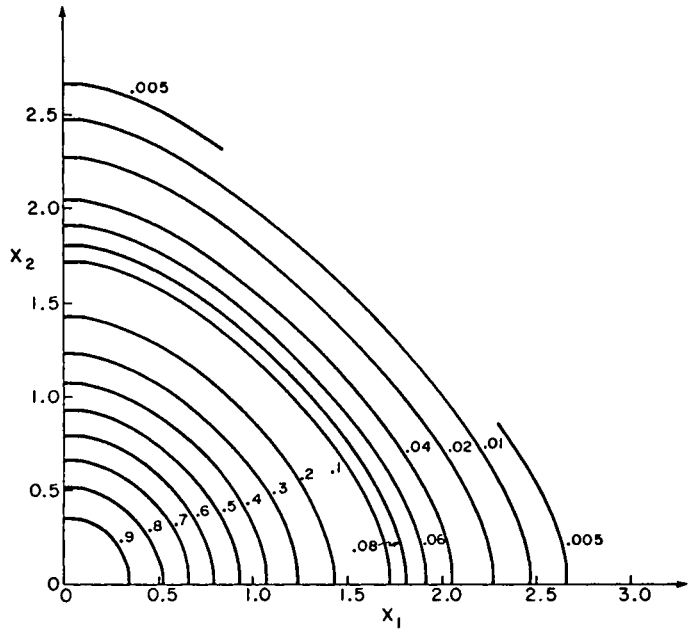


FIG. 16.1. Contour map of the function $\mathscr{B}(x_1, x_2)$, the probability that no eigenvalues lie for a distance x_1 on one side and x_2 on the other side of a given eigenvalue, the distances being measured in units of the mean spacing.

$$\lambda_2 = \frac{\pi^4}{8100} t^5 + 0(t^7), \tag{16.59}$$

$$\lambda_{2i} = 0(t^7), \qquad i \geqslant 2, \tag{16.60}$$

$$\tilde{f}_0(x) = \frac{1}{\sqrt{2}} \left(\left[1 - \frac{\pi^4}{12960} t^4 \right) + \left(- \frac{\pi^2}{36} t^2 + \frac{\pi^4}{4536} t^4 \right) P_2(x) + \frac{\pi^4}{8400} t^4 P_4(x) + 0(t^6) \right], \tag{16.61}$$

$$\tilde{f}_2(x) = (\tfrac{5}{2})^{1/2} P_2(x) + 0(t^2), \tag{16.62}$$

$$\tilde{f}_{2i}(x) = 0(1), \qquad i \geqslant 2, \tag{16.63}$$

$$A(t, \tau) = \left\{ \prod_i (1 - \lambda_{2i}) \right\} \left\{ \sum_i \frac{2}{t} \frac{\lambda_{2i}}{1 - \lambda_{2i}} \tilde{f}_{2i} \left(\frac{\tau}{t} \right) \tilde{f}_{2i}(1) \right\}$$
$$= 1 - \frac{\pi^4}{24} (t^2 + \tau^2) + \frac{\pi^4}{1920} (t^4 + 6t^2\tau^2 + \tau^4)$$
$$+ \frac{\pi^4}{5400} (t^5 - 5t^3\tau^2) + 0(t^r\tau^s), \qquad r + s = 6, \tag{16.64}$$

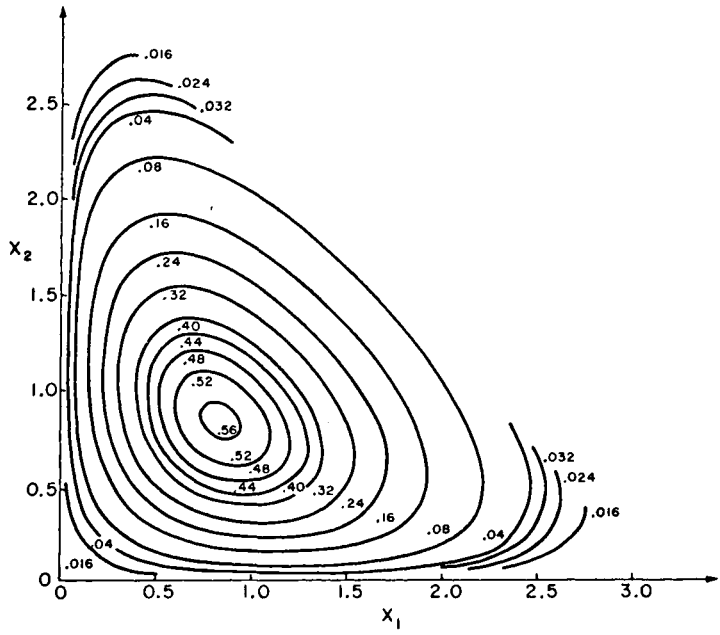


FIG. 16.2. Contour map of the function $\mathscr{P}(x_1, x_2)$, the joint probability density function for two adjacent spacings x_1 and x_2 measured in units of the mean spacing.

$$\mathscr{B}(x_1, x_2) = 1 - \frac{\pi^2}{12}(x_1^2 + x_2^2) + \frac{\pi^4}{240}(x_1^4 + x_2^4) - \frac{\pi^4}{1350}\{x_1^5 - 5x_1^2x_2^2(x_1 + x_2) + x_2^5\} + \cdots, \quad (16.65)$$

$$\mathscr{P}(x_1, x_2) = \frac{\pi^4}{45} x_1x_2(x_1 + x_2) + \cdots. \quad (16.66)$$

Functions $\mathscr{B}(x_1, x_2)$ and $\mathscr{P}(x_1, x_2)$ are represented graphically as contour maps on Fig. 16.1 and 16.2, while their numerical values are given in Appendix A.30.

16.5. The Distribution of Spacings between Next-Nearest Neighbors

The probability of having exactly one level anywhere in an interval of length t can be immediately calculated by integrating $A(t, \tau)$, equation (16.54), over $\frac{1}{2}\tau$ from $-\frac{1}{2}t$ to $\frac{1}{2}t$. The result is

$$E'(t) - E(t) = \left\{\prod_{i=0}^{\infty}(1 - \lambda_{2i})\right\}\left\{\sum_{i=0}^{\infty}\frac{\lambda_{2i}}{1 - \lambda_{2i}}\tilde{f}_{2i}(1)\int_{-1}^{1}\tilde{f}_{2i}(x)\,dx\right\}, \quad (16.67)$$

where $E'(t)$, as in Chapter 9, is the probability that an interval of length t (measured in units of the mean spacing) will contain at most one eigenvalue, and

$$E(t) = \prod_{i=0}^{\infty}(1 - \lambda_{2i}) \quad (5.105)$$

is the probability that an interval t will contain none of the eigenvalues. The results are identical to those for the orthogonal circular ensemble, as a comparison with the discussion following equation (9.35) will show, provided we have the identity

$$\prod_{i=0}^{\infty}\left(\frac{1 - \lambda_{2i+1}}{1 - \lambda_{2i}}\right) = 1 + \sum_{i=0}^{\infty}\left\{\frac{\lambda_{2i}}{1 - \lambda_{2i}}\tilde{f}_{2i}(1)\int_{-1}^{1}\tilde{f}_{2i}(x)\,dx\right\} \quad (16.68)$$

(for a proof see Appendix A.31), where the λ are the eigenvalues and the $\tilde{f}$ the normalized eigenfunctions of the integral equation

$$\lambda\tilde{f}(x) = \int_{-1}^{1}\frac{\sin\{(x + y)\frac{1}{2}t\pi\}}{(x + y)\pi}\tilde{f}(y)\,dy. \quad (16.69)$$

The even functions and the corresponding eigenvalues are labeled by an even index

$$\tilde{f}_{2i}(-x) = \tilde{f}_{2i}(x), \tag{16.70}$$

whereas the odd functions and the corresponding eigenvalues are labeled by an odd index

$$\tilde{f}_{2i+1}(-x) = -\tilde{f}_{2i+1}(x). \tag{16.71}$$

In fact the prolate spheroidal functions $\tilde{f}_j(x)$ also satisfy the differential equation

$$\left\{(x^2 - 1)\frac{d^2}{dx} + 2x\frac{d}{dx} + \frac{\pi^2}{4}t^2x^2 - l\right\}\tilde{f}(x) = 0, \tag{5.107}$$

and for $t \to 0$, the $\tilde{f}_j(x)$ becomes proportional to the Legendre polynomial $P_j(x)$.

17 / Restricted Trace Ensembles. Ensembles Related to the Classical Orthogonal Polynomials

As mentioned toward the end of Chapter 2, Gaussian ensembles are unsatisfactory because the various matrix elements $H_{ij}^{(\lambda)}$ are not equally weighted. Apart from Dyson's method, efforts have been made to equalize this weighting in a straightforward manner. For example, by diagonalizing on a computer a large number of random matrices [Porter and Rosenzweig, 1], the elements of which can be made to conform to a given probability law, we can learn a lot about their eigenvalue distributions. Such knowledge, although useful, is purely empirical, and we restrict ourselves to only those cases in which these empirical findings can be put on a firmer footing.

17.1. Fixed Trace Ensemble†

When working with large but finite dimensional Hermitian matrices, we cannot allow the elements to grow indefinitely because then one would be unable to normalize. Gaussian ensembles overcome this difficulty by giving exponentially vanishing weights to large values of matrix elements. Another method will be to apply a cut-off. Proceeding from the analogy of a fixed energy in classical statistical mechanics, Rosenzweig defines his "fixed trace" ensemble by the requirement that the trace of H^2 be fixed to a number r^2 with no other constraint. The number r is called the strength of the ensemble. The joint probability density function for the matrix elements of H is therefore given by

$$P_r(H) = K_r^{-1}\,\delta\left[\frac{1}{r^2}\,\mathrm{tr}\,H^2 - 1\right],$$

† Rosenzweig [1].

with

$$K_r = \int_{-\infty}^{\infty} \cdots \int \delta\left(\frac{1}{r^2} \operatorname{tr} H^2 - 1\right) \prod_{\lambda} \prod_{i \leqslant j} dH_{ij}^{(\lambda)}.$$

This probability density function is invariant under a change of basis

$$H' = W^R H W,$$

where W is an orthogonal, unitary, or symplectic matrix according to the three possibilities noted in Chapter 2: $W^R W = 1$. This is evident from the fact that under such a transformation the volume element $dH = \prod_\lambda \prod_{i \leqslant j} dH_{ij}^{(\lambda)}$ and the quantity $\operatorname{tr} H^2$ are invariant.

The important thing to be noted about these ensembles is their moment equivalence with Gaussian ensembles of large dimensions. More precisely, if we choose the constant a in (2.72) to give

$$\langle \operatorname{tr} H^2 \rangle_G \approx K_G^{-1} \int_{-\infty}^{\infty} \cdots \int \operatorname{tr} H^2 \, e^{-a \operatorname{tr} H^2} \, dH = r^2,$$

then for any fixed value of the sum

$$s = \sum_{\lambda} \sum_{i \leqslant j} \eta_{ij}^{(\lambda)}, \qquad \eta_{ij}^{(\lambda)} \geqslant 0,$$

the ratio of

$$M_r(N, \eta) = \left\langle \prod_{\lambda} \prod_{i \leqslant j} (H_{ij}^{(\lambda)})^{\eta_{ij}^{(\lambda)}} \right\rangle_r$$

and

$$M_G(N, \eta) = \left\langle \prod_{\lambda} \prod_{i \leqslant j} (H_{ij}^{(\lambda)})^{\eta_{ij}^{(\lambda)}} \right\rangle_G$$

tends to unity as the number of dimensions N tends to infinity. The subscripts r and G denote that the average is taken in the fixed trace and Gaussian ensembles, respectively.

Notice the analogy with the assumption

$$\langle E \rangle \text{ grand canonical} = E \text{ canonical},$$

made in classical statistical mechanics to prove the equivalence there.

From (3.42) we get, with a little manipulation,

$$K_G = a^{-(1/2)N - (1/4)\beta N(N-1)} C_{N\beta},$$

so that a partial differentiation with respect to a gives

$$\langle \operatorname{tr} H^2 \rangle_G = (\tfrac{1}{2}N + \tfrac{1}{4}\beta N(N-1))\, a^{-1}.$$

Therefore we make the choice

$$a = (2r^2)^{-1} N[1 + \tfrac{1}{2}\beta(N-1)].$$

Next, to calculate $M_r(N, \eta)$, put

$$H_{ij}^{(\lambda)} = a^{1/2} r h_{ij}^{(\lambda)} \xi^{-1/2},$$

where ξ is a parameter. This gives

$$M_r(N, \eta) \left(\frac{\xi}{ar^2}\right)^{(1/2)N+(1/4)\beta N(N-1)+(1/2)s}$$

$$= K_r^{-1} \int_{-\infty}^{\infty} \cdots \int \delta\left(\frac{a \operatorname{tr} h^2}{\xi} - 1\right) \prod_{\lambda} \prod_{i \leqslant j} [(h_{ij}^{(\lambda)})^{\eta_{ij}^{(\lambda)}}\, dh_{ij}^{(\lambda)}]$$

Multiplying both sides by $e^{-\xi}$ and integrating (first!) on ξ from 0 to ∞, we get

$$M_r(N, \eta)\, \Gamma(L + \tfrac{1}{2}s)\, L^{-L-(1/2)s}$$

$$= K_r^{-1} \int_{-\infty}^{\infty} \cdots \int e^{-a \operatorname{tr} h^2} \prod_{\lambda} \prod_{i \leqslant j} [(h_{ij}^{(\lambda)})^{\eta_{ij}^{(\lambda)}}\, dh_{ij}^{(\lambda)}],$$

where we have put

$$L = ar^2 = \tfrac{1}{2}N + \tfrac{1}{4}\beta N(N-1).$$

Or

$$M_r(N, \eta) = \frac{L^{L+(1/2)s}}{\Gamma(L + \frac{1}{2}s)} \frac{K_G}{K_r} M_G(N, \eta).$$

Setting $\eta_{ij}^{(\lambda)} = 0$ in the above and using the normalization condition $M_r(N, 0) = M_G(N, 0) = 1$, we get the ratio of the constants K_G and K_r. Substituting this ratio we then obtain

$$M_r(N, \eta) = \frac{L^{(1/2)s}\Gamma(L)}{\Gamma(L + \frac{1}{2}s)} M_G(N, \eta).$$

As $N \to \infty, L \to \infty$, and we can use Stirling's formula for the gamma functions

$$\Gamma(x) = x^{x-1/2} e^{-x}(2\pi)^{1/2} \left[1 + 0\left(\frac{1}{x}\right)\right]$$

to prove the asymptotic equality of all the finite moments $s \ll N$.

17.2. Bounded Trace Ensembles

Instead of keeping the trace constant, we might require it to be bounded [Bronk, 1]. We would then obtain a bounded trace ensemble defined by the joint probability density function

$$P_B(H) = \begin{cases} \text{constant}, & \text{if } \operatorname{tr} H^2 \leqslant r^2, \\ 0, & \text{if } \operatorname{tr} H^2 > r^2. \end{cases}$$

The joint probability density function for the eigenvalues is

$$\begin{aligned} P_B(x_1, \ldots, x_N) &= K^{-1} \prod_{i<j} |x_i - x_j|^\beta, \quad \text{if } \sum_i x_i^2 \leqslant r^2, \\ &= 0, \qquad\qquad\qquad\quad \text{if } \sum_i x_i^2 > r^2. \end{aligned}$$

The density of eigenvalues for this ensemble can easily be found in the existing literature. A theorem of Stieltjes [1] states that if there are N unit masses located at the variable points $x_1, x_2, \ldots, x_N$ in the interval $[-\infty, \infty]$ such that their moment of inertia is bounded by

$$\sum_{i=1}^{N} x_i^2 \leqslant \tfrac{1}{2}N(N-1),$$

the unique maximum of the function

$$V(x_1, \ldots, x_N) = \prod_{1 \leqslant i < j \leqslant N} |x_i - x_j|^\beta$$

will be obtained when the x_i are the zeros of the Hermite polynomial

$$H_N(x) = e^{x^2}\left(-\frac{d}{dx}\right)^N e^{-x^2}.$$

Thus, making the usual assumption of classical statistical mechanics that the actual eigenvalue density makes the logarithm of $V(x_1, \ldots, x_N)$ a maximum, we get this result:

The eigenvalue density for the bounded trace ensembles is identical to the density of zeros of Hermite-like polynomial

$$e^{N(N-1)x^2/2r^2}\left(-\frac{d}{dx}\right)^N e^{-N(N-1)x^2/2r^2}$$

and for large N is given by

$$\begin{aligned}\sigma(x) &\cong \frac{N^2}{2\pi r^2}\left(\frac{4r^2}{N}-x^2\right)^{1/2}, \quad \text{if} \quad |x|<2rN^{-1/2}\\ &\cong 0, \qquad\qquad\qquad \text{if} \quad |x|>2rN^{-1/2}.\end{aligned}$$

To work out the eigenvalue spacing distribution is much more difficult.

17.3. Matrix Ensembles and Classical Orthogonal Polynomials

One cannot but notice in Chapter 5, for example, that the Gaussian ensembles are closely related to the Hermite polynomials. Orthogonal polynomials other than the Hermite have been extensively investigated [Bateman, 1], and some authors [Fox and Kahn, 1; Leff, 1] have tried to take advantage of this fact. We can define a matrix ensemble by giving the joint probability density function for its eigenvalues arbitrarily:

$$P(x_1,\ldots,x_N)=\prod_{i=1}^{N} f(x_i)\prod_{i<j}|x_i-x_j|^{\beta},$$

where the function $f(x)$ can be chosen to suit the needs.

A series of orthogonal polynomials is uniquely defined, apart from a phase factor, by the range (a, b) of the variable and the weight function $f(x)\geqslant 0$. The construction of these polynomials amounts to an application of Schmidt's orthonormalization procedure to the series of powers $1, x, x^2,\ldots$ with the scalar product

$$(\varphi_1,\varphi_2)=\int_a^b f(x)\varphi_1(x)\varphi_2(x)\,dx$$

and gives the following set of polynomials [Bateman, 1]:

$$P_r(x)=(A_r A_{r-1})^{-1/2}\begin{vmatrix} c_0 & c_1 & \cdots & c_r \\ c_1 & c_2 & \cdots & c_{r+1} \\ \cdots & \cdots & \cdots & \cdots \\ c_{r-1} & c_r & \cdots & c_{2r-1} \\ 1 & x & \cdots & x^r \end{vmatrix},$$

with

$$A_r = \det[c_{i+j}]_{i,j=0,1,\ldots,r},$$

and

$$c_i = \int_a^b f(x)\, x^i\, dx.$$

The choices $f(x) = \exp(-\frac{1}{2}\beta x^2)$, $-\infty < x < \infty$, and $f(x) = 1$, $x = e^{i\theta}$, $0 \leqslant \theta \leqslant 2\pi$ correspond to the Gaussian and the circular ensembles, respectively. The other typical choices

$$f(x) = (1 - x)^{\mu}(1 + x)^{\nu}; \qquad \mu, \nu > -1; \qquad -1 \leqslant x \leqslant 1,$$

and

$$f(x) = x^{\alpha}x^{-x}, \qquad \alpha > -1, \qquad 0 \leqslant x < \infty$$

give the Jacobi and the Laguerre ensembles, respectively. Though from the point of view of applications $\beta = 2$ is the least interesting, it is mathematically the easiest to handle. The eigenvalue density $\sigma_N(x)$ can be expressed in terms of the related orthonormal polynomials. Thus

$$\sigma_N(x) = \sum_{j=0}^{N-1} p_j^2(x),$$

where $p_j(x)$ is the normalized polynomial corresponding to the weight function $f(x)$ and the interval $[a, b]$. The n-point correlation function

$$R_n(x_1, \ldots, x_N) = \frac{N!}{(N-n)!} \int \cdots \int P(x_1, \ldots, x_N)\, dx_{n+1} \cdots dx_N$$

is given by

$$R_n(x_1, \ldots, x_N) = \det[K_N(x_j, x_k)]_{j,k=1,2,\ldots,n},$$

where

$$K_N(x, y) = \sum_{j=0}^{N-1} p_j(x)\, p_j(y).$$

The sum $K_N(x, y)$ can be expressed in closed form by using the Christoffel–Darboux formula [Bateman, 1], but it is not very useful in asymptotic evaluations.

Appendices

A.1. Proof of Equation (2.52)

The proof is by induction. Let us assume that (2.52) has been verified for $r \leqslant n - 1$. For $r = n$ we expand the determinant by its last row and last column and replace the lower order determinants with the corresponding σ_r ; therefore we have to prove that

$$n!\sigma_n = (n-1)!\,[t_1\sigma_{n-1} - t_2\sigma_{n-2} + \cdots + (-1)^{n-1}t_n],$$

or

$$\sigma_n = \frac{1}{n}\sum_{r=1}^{n}(-1)^{r-1}\,t_r\,\sigma_{n-r}\,, \tag{A1.1}$$

with the convention $\sigma_0 \equiv 1$.

Now $t_1\sigma_{n-1}$ contains $n\sigma_n$. It also contains terms of the form $x_1^2x_2\cdots x_{n-1}$. To remove them we subtract $t_2\sigma_{n-2}$; but then we shall also have subtracted terms of the form $x_1^3x_2\cdots x_{n-2}$ which were not present in $t_1\sigma_{n-1}$. To compensate, we add $t_3\sigma_{n-3}$. However, we shall then have added something more, so we subtract $t_4\sigma_{n-4}$, and so on.

Thus we have proved that (2.52) is valid for $r = n$ if it is true for every $r \leqslant n - 1$. Moreover, (2.52) is trivially valid for $n = 1$. The proof is thus complete.

If $r > N$, no σ_r exists and

$$\sum_{r=n-N}^{n}(-1)^r\,t_r\,\sigma_{n-r} \equiv 0, \qquad n > N. \tag{A1.2}$$

A.2. Counting the Dimensions of $T_{\beta G}$ and $T'_{\beta G}$ (Chapter 3) and of $T_{\beta C}$ and $T'_{\beta C}$ (Chapter 8)

When we require that two of the eigenvalues be equal, we drop a number of parameters needed to specify a certain two-dimensional

subspace, the subspace of these two equal eigenvalues. However, this degenerate eigenvalue is itself one real parameter. Thus, if $f(N, \beta)$ is the number of independent real parameters needed to specify a particular matrix from the ensemble $E_{\beta G}$, the number needed to specify a matrix from the ensemble $E_{\beta G}$ with two equal eigenvalues is

$$f(N, \beta) - f(2, \beta) + 1. \tag{A2.1}$$

In other words if the number of dimensions of the space $T_{\beta G}$ is $f(N, \beta)$, that of the space $T'_{\beta G}$ is $f(N, \beta) - f(2, \beta) + 1$.

Now to specify a matrix from any of the ensembles $E_{\beta G}$ we need specify only the matrix elements H_{ij} with $i \leqslant j$. The diagonal elements are real and therefore require N real parameters for their specification. The off-diagonal elements H_{ij} with $i < j$ are $\frac{1}{2}N(N-1)$ in number and they need $\frac{1}{2}N(N-1)\beta$ real parameters. Thus

$$f(N, \beta) = N + \tfrac{1}{2}N(N-1)\beta. \tag{A2.2}$$

By inserting $\beta = 1$, 2, or 4 into (A2.2) and (A2.1) we get the dimensions of $T_{\beta G}$ and $T'_{\beta G}$.

To count the dimensions of the spaces $T_{\beta C}$ and $T'_{\beta C}$ we must find the corresponding numbers $f(N, \beta)$, and for this purpose it is sufficient to consider matrices in the neighborhood of unity. Let us then have

$$S = 1 + iA,$$

where A is infinitesimal. Since S is unitary,

$$S^\dagger S \equiv (1 - iA^\dagger)(1 + iA) = 1$$

or, up to terms linear in A,

$$A = A^\dagger; \tag{A2.3}$$

A is then Hermitian. If, in addition, S is symmetric (self-dual), then A is symmetric (self-dual). Thus the number of independent real parameters needed to specify a symmetric unitary, self-dual unitary, or unitary matrix S is the same as that needed to specify a symmetric Hermitian, self-dual Hermitian, or Hermitian matrix A, respectively. Thus the dimensions of $T_{\beta C}$ and $T_{\beta G}$ are equal, and hence also those of $T'_{\beta C}$ and $T'_{\beta G}$.

A.3. Two Proofs of Equation (4.5) for the Case $N = 3$

A.3.1. The Gaudin Method [2]

Because the integrand is a symmetric function of x_1, x_2, and x_3, let us put

$$\sigma_1 = (\tfrac{1}{2}\beta)^{1/2}(x_1 + x_2 + x_3),$$
$$\sigma_2 = (\tfrac{1}{2}\beta)(x_2x_3 + x_3x_1 + x_1x_2),$$
$$\sigma_3 = (\tfrac{1}{2}\beta)^{3/2} x_1x_2x_3,$$

so that

$$\Delta^2 \equiv [(x_2 - x_3)(x_3 - x_1)(x_1 - x_2)]^2 = p;$$

and the Jacobian is

$$\frac{\partial(\sigma_1, \sigma_2, \sigma_3)}{\partial(x_1, x_2, x_3)} = (\tfrac{1}{2}\beta)^3 \mid \Delta \mid.$$

The integral can therefore be written with the new variables as

$$\Psi_3(\beta) = 3!(\tfrac{1}{2}\beta)^{-3} \int_D e^{-(\sigma_1^2 - 2\sigma_2)} p^{(1/2)(\beta-1)}\, d\sigma_1\, d\sigma_2\, d\sigma_3 \qquad \text{(A3.1)}$$

where p is the discriminant

$$-p = (\tfrac{1}{2}\beta)^{-3}[27\sigma_3^2 - 2\sigma_3(2\sigma_1^3 - 9\sigma_1\sigma_2) + 4\sigma_2^3 - \sigma_1^2\sigma_2^2]$$

and D is the domain of integration on which p is nonnegative. Now p is a polynomial of second degree in σ_3 which must have two real roots. Let these roots be a and b:

$$p \equiv 27(\tfrac{1}{2}\beta)^{-3}(a - \sigma_3)(\sigma_3 - b), \qquad a > b;$$

so that

$$\int_b^a d\sigma_3\, p^{(1/2)(\beta-1)} = 27^{(1/2)(\beta-1)}(\tfrac{1}{2}\beta)^{-(3/2)(\beta-1)}(a - b)^\beta B\left(\frac{\beta + 1}{2}, \frac{\beta + 1}{2}\right),$$

where B is Euler's beta function. The quantity $(a - b)^2$ is the discriminant of p as a polynomial in σ_3

$$(a - b)^2 = \frac{4}{27^2}[4\sigma_1^6 - 36\sigma_1^4\sigma_2 + 108(\sigma_1^2\sigma_2^2 - \sigma_2^3)]$$
$$= \left(\frac{4}{27}\right)^2 (\sigma_1^2 - 3\sigma_2)^3.$$

Therefore we must have $\sigma_1^2 \geqslant 3\sigma_2$. Taking σ_1 and $z = \frac{1}{2}\sigma_1^2(\sigma_1^2 - 2\sigma_2)^{-1}$ as new variables of integration, we see that z varies from 0 to 1 and σ_1 from $-\infty$ to ∞. We then have

$$\begin{aligned}
&\iint (a-b)^\beta e^{-(\sigma_1^2-2\sigma_2)}\, d\sigma_1\, d\sigma_2 \\
&\quad = \left(\frac{4}{27}\right)^\beta \iint e^{-\sigma_1^2/3z} \mid \sigma_1 \mid^{3\beta} \left(\frac{1-z}{2z}\right)^{(3/2)\beta} \sigma_1^2 z^{-2\frac{1}{6}}\, d\sigma_1\, dz \\
&\quad = \frac{1}{6}\left(\frac{4}{27}\right)^\beta \left(\frac{1}{2}\right)^{3\beta/2} \int_0^1 dz \int_{-\infty}^{\infty} d\sigma_1\, e^{-\sigma_1^2/3z} \mid \sigma_1 \mid^{3\beta+2} z^{-((3/2)\beta+2)} (1-z)^{(3/2)\beta} \\
&\quad = \tfrac{1}{6} 2^{(1/2)\beta} 27^{-(1/2)(\beta-1)} \Gamma(\tfrac{3}{2}\beta + \tfrac{3}{2}) \int_0^1 dz z^{-1/2} (1-z)^{(3/2)\beta} \\
&\quad = \tfrac{1}{6} 2^{(1/2)\beta} 27^{-(1/2)(\beta-1)} \Gamma(\tfrac{3}{2}\beta + \tfrac{3}{2})\, B(\tfrac{1}{2}, \tfrac{3}{2}\beta + 1).
\end{aligned}$$

Collecting the results, we obtain

$$\Psi_3(\beta) = 2^{\beta/2}(\tfrac{1}{2}\beta)^{-(3/2)(\beta+1)} B\left(\frac{\beta+1}{2}, \frac{\beta+1}{2}\right) B(\tfrac{1}{2}, \tfrac{3}{2}\beta + 1)\, \Gamma(\tfrac{3}{2}\beta + \tfrac{3}{2}).$$

Using the duplication formula

$$\Gamma(\tfrac{1}{2}\beta)\, \Gamma(\tfrac{1}{2}\beta + \tfrac{1}{2}) = 2^{-\beta+1} \sqrt{\pi}\, \Gamma(\beta),$$

we get (4.5) for $N = 3$;

$$\Psi_3(\beta) = 6\beta^{-(3/2)(\beta+1)} (2\pi)^{3/2} \frac{\Gamma(\beta)\, \Gamma(\frac{3}{2}\beta)}{\Gamma(\frac{1}{2}\beta)\, \Gamma(\frac{1}{2}\beta)}.$$

The factor 3! in (A3.1) occurs because the correspondence between x_1, x_2, x_3, and σ_1, σ_2, σ_3 is not one to one.

A.3.2. The C. L. Mehta Method [1]

Introduce the new variables

$$y_1 = \sqrt{\frac{k}{2}}\,(x_1 - x_2),$$

$$y_2 = \sqrt{\frac{k}{6}}\,(x_1 + x_2 - 2x_3),$$

$$y_3 = \sqrt{\frac{k}{3}}\,(x_1 + x_2 + x_3),$$

so that the Jacobian is

$$\frac{\partial(y_1, y_2, y_3)}{\partial(x_1, x_2, x_3)} = k^{3/2}$$

and

$$k(x_1^2 + x_2^2 + x_2^3) = y_1^2 + y_2^2 + y_3^2.$$

The integral

$$\Psi_3(2k) = \int\int\int_{-\infty}^{\infty} \exp\left(-k\sum_1^3 x_j^2\right) \prod_{1\leqslant i<j\leqslant 3} (x_i - x_j)^{2k}\, dx_1\, dx_2\, dx_3\,, \qquad \beta = 2k,$$

therefore transforms to

$$\Psi_3(2k) = \int\int\int_{-\infty}^{\infty} \exp\left(-\sum_1^3 y_i^2\right) [\sqrt{2}\, y_1\, \tfrac{1}{2}(\sqrt{2}\, y_1 + \sqrt{6}\, y_2)\, \tfrac{1}{2}(\sqrt{2}\, y_1 - \sqrt{6}\, y_2)]^{2k}$$
$$\times\, dy_1\, dy_2\, dy_3\, k^{-(3/2)(1+2k)}\,.$$

The integration over y_3 is immediate and gives

$$\Psi_3 = \sqrt{\pi}\, 2^{-k} k^{-(3/2)(1+2k)} \int\int_{-\infty}^{\infty} e^{-y_1^2 - y_2^2} (y_1^3 - 3y_1 y_2^2)^{2k}\, dy_1\, dy_2\,.$$

Now put

$$y_1 = r^{1/2} \cos\theta, \qquad y_2 = r^{1/2} \sin\theta,$$

so that

$$y_1^3 - 3y_1 y_2^2 = r^{3/2} \cos 3\theta, \qquad \frac{\partial(y_1, y_2)}{\partial(r, \theta)} = \frac{1}{2}\,;$$

hence

$$\Psi_3 = \sqrt{\pi}\, 2^{-k-1} k^{-(3/2)(1+2k)} \int_0^{\infty} e^{-r} r^{3k}\, dr \int_0^{2\pi} (\cos 3\theta)^{2k}\, d\theta$$
$$= \sqrt{\pi}\, 2^{-k-1} k^{-(3/2)(1+2k)} (3k)!\, \frac{1}{3} \int_0^{6\pi} (\cos\phi)^{2k}\, d\phi$$
$$= \pi^{3/2} 2^{-3k} k^{-(3/2)(1+2k)} (3k)!\, (2k)!\, k!$$

which is (4.13).

A.4. The Minimum Value of W, Equation (4.6)

To get the minimum value of the potential energy W we present Stieltjes' [1] ingenious arguments.

The existence of a minimum is clear. Let the points $x_1, \ldots, x_N$ make

$$W = \frac{1}{2}\sum_{1}^{N} x_j^2 - \sum_{1 \leqslant i < j \leqslant N} \ln | x_i - x_j | \tag{A4.1}$$

a minimum; then

$$0 = -\frac{\partial W}{\partial x_j} \equiv -x_j + \sum_{i(\neq j)} \frac{1}{x_j - x_i}. \tag{A4.2}$$

Consider the polynomial

$$g(x) = (x - x_1)(x - x_2) \cdots (x - x_N), \tag{A4.3}$$

which has $x_1, x_2, \ldots, x_N$ as its zeros. Differentiation gives

$$g'(x_j) = \prod_{i(\neq j)} (x_j - x_i)$$

and

$$\frac{g''(x_j)}{g'(x_j)} = 2 \sum_{i(\neq j)} \frac{1}{x_j - x_i} \tag{A4.4}$$

so that (A4.2) can be written as

$$g''(x_j) - 2x_j g'(x_j) = 0. \tag{A4.5}$$

This means that the polynomial

$$g''(x) - 2x\, g'(x)$$

of order N has its zeros at $x_1, \ldots, x_N$ and therefore it must be proportional to $g(x)$. Comparing the coefficients of x^N, we see that

$$g''(x) - 2xg'(x) + 2Ng(x) = 0. \tag{A4.6}$$

The polynomial solution of this differential equation is uniquely determined to be the Hermite polynomial of order N:

$$H_N(x) = N! \sum_{m=0}^{[\frac{1}{2}N]} \frac{(-1)^m (2x)^{N-2m}}{m!\,(N - 2m)!}. \tag{A4.7}$$

The discriminant of $H_N(x)$ [Szegö, 1] is

$$\prod_{1 \leqslant i < j \leqslant N} (x_i - x_j)^2 = 2^{-(1/2)N(N-1)} \prod_{j=1}^{N} j^j, \tag{A4.8}$$

and from (A4.7) we get

$$\sum_{1}^{N} x_j^2 = \tfrac{1}{2}N(N-1). \tag{A4.9}$$

Thus the minimum value of W is

$$W_0 = \tfrac{1}{4}N(N-1)(1 + \ln 2) - \frac{1}{2}\sum_{1}^{N} j \ln j$$

which is (4.6).

A.5. Proof of Equation (4.15)

We have

$$\begin{aligned}
&\left(\frac{k}{\pi}\right)^{1/2} \int_{-\infty}^{\infty} \exp[-kx^2 - i(2k)^{1/2}\lambda x] x^n \, dx \\
&= \left(\frac{k}{\pi}\right)^{1/2} \int_{-\infty}^{\infty} \left(\frac{i}{\sqrt{2k}} \frac{d}{d\lambda}\right)^n e^{-\lambda^2/2} \exp[-k(x + i(2k)^{-1/2}\lambda)^2] \, dx \\
&= \left(\frac{i}{\sqrt{2k}} \frac{d}{d\lambda}\right)^n e^{-\lambda^2/2}.
\end{aligned}$$

Putting $\lambda = 0$ on both sides we get (4.15).

A.6. Proof of Equations (5.4), (5.51), and (5.52)

With $R_n(x_1, x_2, \ldots, x_n)$ and $T_n(x_1, x_2, \ldots, x_n)$ defined by (5.2) and (5.3), let us write

$$r_0 = 1$$

$$r_n = \int_{-\infty}^{\infty} \cdots \int R_n(x_1, \ldots, x_n) \prod_{1}^{n} [a(x_i)\, dx_i], \qquad n \geqslant 1,$$

$$t_n = \int_{-\infty}^{\infty} \cdots \int T_n(x_1, \ldots, x_n) \prod_{1}^{n} [a(x_i)\, dx_i], \qquad n \geqslant 1,$$

and the generating functions

$$R(z) = \sum_{n=0}^{\infty} \frac{r_n}{n!} z^n \equiv 1 + \sum_{n=1}^{\infty} \frac{r_n}{n!} z^n,$$

$$T(z) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{t_n}{n!} z^n,$$

where the function $a(x)$ is quite arbitrary.

The numbers t_n and r_n may be related by the use of (5.3). In fact,

$$t_n = \sum_G (-1)^{n-m}(m-1)! \frac{n!}{G_1! \cdots G_m!} \frac{1}{m!} r_{G_1} r_{G_2} \cdots r_{G_m}$$

$$= \sum_G (-1)^{n-m} \frac{n!}{m} \prod_{j=1}^{m} \frac{r_{G_j}}{G_j!},$$

the sum being taken over all partitions G of n:

$$n = G_1 + \cdots + G_m, \qquad G_j \geqslant 1, \qquad m \geqslant 1.$$

We can now relate the functions $T(z)$ and $R(z)$:

$$T(z) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n!} t_n z^n$$

$$= \sum_{n=1}^{\infty} \sum_{G(n)} \frac{(-1)^{m-1}}{m} \prod_{j=1}^{m} \frac{r_{G_j}}{G_j!} z^{G_j}.$$

As we are summing finally over all integers n, the restriction

$$\sum_{j=1}^{m} G_j = n$$

may be removed; we then have

$$T(z) = \sum_{m, G_1, \ldots, G_m = 1}^{\infty} \frac{(-1)^{m-1}}{m} \prod_{j=1}^{m} \frac{r_{G_j}}{G_j!} z^{G_j}$$

$$= \sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m} \left(\sum_{l=1}^{\infty} \frac{r_l}{l!} z^l \right)^m$$

$$= \ln R(z);$$

or

$$R(z) = \exp[T(z)].$$

Expanding the exponential and equating the coefficient of z^n on both sides, we get

$$\frac{r_n}{n!} = \text{coefficient of } z^n \text{ in } \sum_{m=1}^{n} \frac{1}{m!} \left(\sum_{l=1}^{n} \frac{(-1)^{l-1}}{l!} t_l z^l \right)^m$$

$$= \sum_{G} \frac{1}{m!} \prod_{j=1}^{m} \left[\frac{(-1)^{G_j - 1}}{G_j!} t_{G_j} \right]; \qquad \sum_{j=1}^{m} G_j = n$$

or

$$r_n = \sum_{G} \frac{(-1)^{n-m}}{m!} \frac{n!}{G_1!\, G_2! \cdots G_m!} t_{G_1} t_{G_2} \cdots t_{G_m}.$$

In view of the arbitrariness of $a(x)$, this relation is identical to (5.4).

Finally, expanding

$$\rho_N = \left\langle \prod_{j=1}^{N} [1 + a(x_j)] \right\rangle$$

in powers of $a(x)$ and using the definition of $R_n(x_1, x_2, \ldots, x_n)$, (5.2), we get

$$\rho_N = [R(z)]_{z=1} = R(1),$$

which is (5.51).

A.7. Proof of Equation (5.20). Expansion of a Pfaffian along Its Principal Pseudodiagonal

Consider an antisymmetric matrix $[a_{ij}]$; $i, j = 1, 2, \ldots, n$; $a_{ij} = -a_{ji}$. If its order n is an odd number, it is easy to see that its determinant is zero. On the other hand, if n is an even number, it can be shown by induction [Aitken, 1] that it is a perfect square; the square root is

$$(\det[a_{ij}])^{1/2} = \frac{1}{(\frac{1}{2}n)!} \sum \pm a_{i_1 i_2} a_{i_3 i_4} \cdots a_{i_{n-1} i_n} \tag{A7.1}$$

where the summation is extended over all permutations $i_1, i_2, \ldots, i_n$ of $1, 2, \ldots, n$, with the restrictions $i_1 < i_2$; $i_3 < i_4$;...; $i_{n-1} < i_n$, and the sign is plus or minus depending on whether the permutation

$$\begin{pmatrix} 1 & 2 \cdots n \\ i_1 & i_2 \cdots i_n \end{pmatrix}$$

is even or odd. The expression (A7.1) is known as the "pfaffian."

Now let us take up the problem of evaluating

$$I = \int \cdots \int dx_1 \cdots dx_n \det[\varphi_j(x_i), \psi_j(x_i)]_{\substack{i=1,\ldots,m \\ j=1,\ldots,2m}} ; \tag{A7.2}$$

the limits of integration are fixed. Define

$$a_{ij} = \int [\varphi_i(x)\,\psi_j(x) - \varphi_j(x)\,\psi_i(x)]\,dx, \tag{A7.3}$$

so that

$$a_{ij} = -a_{ji} .$$

Expanding the integrand in (A7.2) and integrating independently over all the variables we make the following observations:

1. The integral I is a sum of terms, each being a product of m numbers a_{ij} ;

2. The indices of the various a_{ij} occurring in any of the above terms are all different. In totality they are all the indices from 1 to $2m$.

3. We may restrict i to be less than j in each a_{ij} occurring in I; for if i is not less than j we may replace a_{ij} by $-a_{ji}$.

4. The coefficient of the term $a_{i_1 i_2} a_{i_3 i_4} \cdots a_{i_{2m-1} i_{2m}}$ in I is $+1$ or -1, depending on whether the permutation

$$\begin{pmatrix} i_1 & i_2 \cdots i_{2m} \\ 1 & 2 \cdots 2m \end{pmatrix}$$

is even or odd.

From these observations and (A7.1) we conclude that

$$I = m!\,[\det[a_{ij}]_{i,j=1,2,\ldots,2m}]^{1/2}.$$

The expansion of a determinant with large diagonal and small off-diagonal elements is well known. The coefficient of ϵ^k in the power series expansion of

$$\det[\delta_{ij} + \epsilon\alpha_{ij}]_{i,j=1,2,\ldots,N}$$

is the sum of all possible $k \times k$ principal subdeterminants of $[\alpha_{ij}]$ obtained by suppressing symmetrically $(N - k)$ rows and columns of $[\alpha_{ij}]$. This can be seen from the fact that the determinant is a sum

(with proper signs) of products of N elements, one element being taken from every row and every column. To get ϵ^k we must have $N - k$ of the elements in each product equal to unity, which can come only from the diagonal. When the rows and columns containing these elements equal to one are suppressed, a $k \times k$ principal subdeterminant of $[\alpha_{ij}]$ remains.

There is an analogous, but not so widely known, expansion of a pfaffian. To get the expansion of

$$\left\{\det \begin{bmatrix} \epsilon\lambda_{ij} & \delta_{ij} + \epsilon\nu_{ij} \\ -(\delta_{ij} + \epsilon\nu_{ji}) & \epsilon\mu_{ij} \end{bmatrix}\right\}^{1/2}$$

with

$$\lambda_{ij} = -\lambda_{ji}\,, \qquad \mu_{ij} = -\mu_{ji}$$

in powers of ϵ we proceed in a similar manner. To get ϵ^k we must have $(N - k)$ factors in (A7.1) equal to unity, whereas the remaining k factors containing ϵ can be regrouped into pfaffians. Thus

$$\begin{aligned} &\left\{\det \begin{bmatrix} \epsilon\lambda_{ij} & \delta_{ij} + \epsilon\nu_{ij} \\ -(\delta_{ij} + \epsilon\nu_{ji}) & \epsilon\mu_{ij} \end{bmatrix}\right\}^{1/2} \\ &\quad = 1 + \epsilon \sum_i \begin{vmatrix} 0 & \nu_{ii} \\ -\nu_{ii} & 0 \end{vmatrix}^{1/2} + \epsilon^2 \sum_{i<j} \begin{vmatrix} 0 & \nu_{ii} & \lambda_{ij} & \nu_{ij} \\ -\nu_{ii} & 0 & -\nu_{ji} & \mu_{ij} \\ \lambda_{ji} & \nu_{ji} & 0 & \nu_{ij} \\ -\nu_{ij} & \mu_{ji} & -\nu_{jj} & 0 \end{vmatrix}^{1/2} \\ &\qquad + \epsilon^3 \sum_{i<j<k} \left\{ \begin{vmatrix} \lambda_{i_1 i_2} & \nu_{i_1 i_2} \\ -\nu_{i_2 i_1} & \mu_{i_1 i_2} \end{vmatrix}_{i_1, i_2 = i, j, k} \right\}^{1/2} + \cdots. \end{aligned} \tag{A7.4}$$

Whenever any two indices are equal, the corresponding pfaffian is identically zero. Therefore we can sum independently over the indices and replace ϵ^k with $\epsilon^k/k!$.

A.8. The Limit of $\sum_0^{N-1} \varphi_j^2(x)$

The dominant term in $\sum_0^{N-1} \varphi_j^2(x)$ may be obtained with ease by a physical argument.

The $\varphi_j(x)$ is the normalized oscillator function, so that $\varphi_j^2(x)\,dx$ gives the probability that an oscillator in the jth state is found in the interval $(x, x + dx)$. Consider N oscillators, one each in the states $0, 1, \ldots, N - 1$, so that when N is large, $\sum_0^{N-1} \varphi_j^2(x)$ is the density of the particles at the point x. The particles are fermions and the

temperature is zero, for there is not more than one particle in each state and all states up to a certain energy (Fermi energy) are filled. The Fermi momentum corresponding to this maximum energy can be obtained from the differential equation satisfied by $\varphi_{N-1}(x)$.

$$\hbar^2 \frac{\partial^2}{\partial x^2}\varphi_{N-1}(x) + \hbar^2(2N - 1 - x^2)\,\varphi_{N-1}(x) = 0,$$

so that

$$-p_F{}^2 + (2N - 1 - x^2)\,\hbar^2 = 0.$$

Because our oscillators are one-dimensional, their density is given by

$$\sigma(x) \approx \frac{1}{(2\pi\hbar)}\int_{-p_F}^{p_F} dp = \frac{1}{2\pi\hbar}\,2p_F\,, \qquad p \lesssim p_F\,.$$

From the last two equations we get

$$\sigma(x) = \frac{1}{\pi}(2N - x^2)^{1/2}, \qquad x^2 \lesssim 2N,$$

$$= 0, \qquad x^2 \gtrsim 2N.$$

This is the dominant term. Terms of the next lower order cannot be obtained from physical arguments alone. To get further information about $\sigma(x)$ we may write from the formula of Christoffel–Darboux [Bateman, 1]

$$\sum_0^{N-1}\varphi_j{}^2(x) = N\,\varphi_N{}^2(x) - [N(N+1)]^{1/2}\,\varphi_{N-1}(x)\,\varphi_{N+1}(x),$$

and use the known [Erdyeli, 1] asymptotic behavior of the functions $\varphi_{N-1}(x)$, $\varphi_N(x)$, and $\varphi_{N+1}(x)$ for the various intervals of x.

A.9. The Limits of $\sum_0^{N-1}\varphi_j(x)\,\varphi_j(y)$, etc.

The Christoffel–Darboux formula gives [Bateman, 1]

$$\sum_0^{N-1}\varphi_j(x)\,\varphi_j(y) = (\tfrac{1}{2}N)^{1/2}\left[\frac{\varphi_N(x)\,\varphi_{N-1}(y) - \varphi_N(y)\,\varphi_{N-1}(x)}{x - y}\right]. \quad \text{(A9.1)}$$

Let $N = 2m$, $2m^{1/2}x = \pi\xi$, $2m^{1/2}y = \pi\eta$ and let us take the limit $m \to \infty$, $x \to 0$, $y \to 0$; whereas ξ and η are finite. Using the formula [Bateman, 1]

$$\lim(-1)^m m^{1/4}\,\varphi_{2m}(x) = \pi^{-1/2}\cos\pi\xi,$$

$$\lim(-1)^m m^{1/4}\,\varphi_{2m+1}(x) = \pi^{-1/2}\sin\pi\xi,$$

we get

$$\lim \sum_0^{2m-1} \varphi_j(x)\,\varphi_j(y) = \frac{2m^{1/2}}{\pi}\,\frac{\sin\pi\xi\cos\pi\eta - \cos\pi\xi\sin\pi\eta}{\pi\xi - \pi\eta}$$

$$= \frac{2\sqrt{m}}{\pi}\,\frac{\sin(\xi-\eta)\,\pi}{(\xi-\eta)\,\pi}, \tag{A9.2}$$

which is (5.61).

To obtain the limit of

$$K_m(x, y) = \sum_0^{m-1} \varphi_{2j}(x)\,\varphi_{2j}(y)$$

we observe that $K_m(x, y)$ is the even part in x of (A9.2).

$$K_m(x, y) = \frac{1}{2}\sum_0^{2m-1} [\varphi_j(x)\,\varphi_j(y) + \varphi_j(-x)\,\varphi_j(y)]$$

Therefore

$$\lim K_m(x, y) = \frac{1}{2}\,\frac{2\sqrt{m}}{\pi}\left[\frac{\sin(\xi-\eta)\,\pi}{(\xi-\eta)\,\pi} + \frac{\sin(\xi+\eta)\,\pi}{(\xi+\eta)\,\pi}\right] \tag{A9.3}$$

$$= \frac{2\sqrt{m}}{\pi} Q(\xi, \eta), \text{ say.} \tag{A9.4}$$

Also

$$\lim \sum_0^{m-1} \varphi_{2j}(x)\,\varphi'_{2j}(y) = \lim \frac{\partial}{\partial y} \sum_0^{m-1} \varphi_{2j}(x)\,\varphi_{2j}(y)$$

$$= \frac{2\sqrt{m}}{\pi}\,\frac{\partial}{\partial\eta} \lim K_m(x, y)$$

$$= \left(\frac{2\sqrt{m}}{\pi}\right)^2 \frac{\partial}{\partial\eta} Q(\xi, \eta) \tag{A9.5}$$

and

$$\lim \sum_{0}^{m-1} \varphi_{2j}(x) \int_0^y \varphi_{2j}(z)\, dz = \lim \int_0^y K_m(x, z) dz$$
$$= \int_0^\eta Q(\xi, \zeta)\, d\zeta. \tag{A9.6}$$

By writing similar equations for

$$\lim \sum_{0}^{m-1} \varphi'_{2j}(x)\, \varphi_{2j}(y) \quad \text{and} \quad \lim \sum_{0}^{m-1} \varphi_{2j}(y) \int_0^x \varphi_{2j}(z)\, dz$$

and combining, we obtain (5.62) and (5.63).

A.10. The Fourier Transforms of the Two-Point Cluster Functions

The functions $s(r)$, $(d/dr)(s(r))$, and $\int_0^r s(z)\, dz$ are given only for positive values of r by (5.64), (5.62), and (5.63); for negative values of r they are defined by the statement that they are even functions of r. Therefore

$$F[s^2(r)] \equiv \int_{-\infty}^{\infty} e^{2\pi i k r}\, s^2(r)\, dr = 2 \int_0^\infty \cos(2\pi |k| r) \frac{\sin^2 \pi r}{\pi^2 r^2}\, dr$$
$$= \int_0^\infty \cos(2\pi |k| r) \frac{1 - \cos 2\pi r}{\pi^2 r^2}\, dr$$
$$= \int_0^\infty (2\pi^2 r^2)^{-1} [2 \cos(2\pi |k| r) - \cos(2|k| + 2)\pi r$$
$$- \cos(2|k| - 2)\pi r]\, dr.$$

Now

$$\int_0^\infty r^{-2}(\cos ar - \cos br)\, dr = \int_a^b d\lambda \int_0^\infty r^{-1} \sin \lambda r\, dr$$
$$= \frac{\pi}{2} \int_a^b d\lambda \operatorname{sign} \lambda = \frac{\pi}{2} (|b| - |a|),$$

so that

$$F[s^2(r)] = (2\pi^2)^{-1} \frac{\pi}{2} 2\pi[(|k| + 1) - |k| + |(|k| - 1)| - |k|]$$
$$= \tfrac{1}{2}[(1 - |k|) + |(1 - |k|)|]$$
$$= \begin{cases} 1 - |k|, & \text{if } |k| < 1, \\ 0, & \text{if } |k| > 1. \end{cases} \tag{A10.1}$$

By partial integration we have

$$F\left[\frac{d}{dr}s(r)\right] = 2\int_0^\infty \cos(2\pi|k|r)\frac{d}{dr}\frac{\sin\pi r}{\pi r}\,dr$$

$$= 2\left\{-1 + |k|\int_0^\infty r^{-1}[\cos(2|k|-1)\pi r - \cos(2|k|+1)\pi r]\,dr\right\}$$

$$= 2\left(-1 + |k|\int_{(2|k|-1)\pi}^{(2|k|+1)\pi}\int_0^\infty dr\sin\lambda r\right).$$

Now

$$\int_0^\infty dr\sin\lambda r = \lim_{\alpha\to 0}\int_0^\infty e^{-\alpha r}\sin\lambda r\,dr = \lim_{\alpha\to 0}\frac{\lambda}{\alpha^2+\lambda^2} = \frac{1}{\lambda},$$

so that

$$F\left(\frac{d}{dr}s(r)\right) = 2\left(-1 + |k|\ln\frac{(2|k|+1)}{|(2|k|-1)|}\right). \tag{A10.2}$$

Also by partial integration

$$F\left\{\left[\int_0^r s(z)\,dz\right]\left[\frac{d}{dr}s(r)\right]\right\}$$

$$= -F[s^2(r)] + 4\pi|k|\int_0^\infty dr\sin(2\pi|k|r)\frac{\sin\pi r}{\pi r}\left[\int_0^r s(z)\,dz\right]$$

and

$$2\int_0^\infty r^{-1}\sin(2\pi|k|r)\sin\pi r\left(\int_0^r\frac{\sin\pi x}{x}\,dx\right)dr$$

$$= \int_{(2|k|-1)\pi}^{(2|k|+1)\pi} dz\int_0^\infty dr\sin zr\left(\int_0^r dx\frac{\sin\pi x}{x}\,dx\right)$$

$$= \int_{(2|k|-1)\pi}^{(2|k|+1)\pi} dz\frac{1}{2z}[\operatorname{sign}(\pi+z) + \operatorname{sign}(\pi-z)]\frac{\pi}{2}$$

$$= \begin{cases} -\dfrac{\pi}{2}\ln|(2|k|-1)|, & \text{if } |k|<1, \\ 0, & \text{if } |k|>1, \end{cases}$$

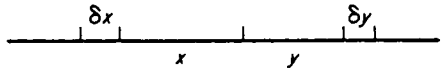
so that

$$F\left\{\left[\int_0^r s(z)\,dz\right]\left[\frac{d}{dr}s(r)\right]\right\} = \begin{cases} -1 + |k| - |k|\ln|(2|k|-1)|, & \text{if } |k|<1, \\ 0, & \text{if } |k|>1. \end{cases} \tag{A10.3}$$

Combining (A10.1), (A10.2) and (A10.3), we get all the Fourier transforms quoted in Chapter 5.

A.11. Proof of Equations (5.84) and (9.39)

Let $E(x + y)$ be the probability that an interval of length $x + y$



is empty of eigenvalues. Then $E(x + \delta x + y)$ is the probability that the interval $\delta x + x + y$ is empty, and

$$E(x + y) - E(x + \delta x + y)$$

is the probability that the interval $[x + y]$ is empty and δx is not empty. The probability that δx will contain more than one eigenvalue is of second or higher order in δx. Therefore, taking the limit $\delta x \to 0$ and keeping only the first-order terms, we get

$$-\frac{\partial E(x+y)}{\partial x}\,\delta x = \text{probability that } x + y \text{ is empty and } \delta x \text{ contains one eigenvalue}$$

By a similar argument we obtain

$$\frac{\partial^2 E(x+y)}{\partial x\,\partial y}\,\delta x\,\delta y = \text{prob}\begin{Bmatrix} x+y, & 0 \\ \delta x, & 1 \\ \delta y, & 1 \end{Bmatrix},$$

where

$$\text{prob}\begin{Bmatrix} x+y, & 0 \\ \delta x, & 1 \\ \delta y, & 1 \end{Bmatrix}$$

means the probability that $x + y$ is empty and δx, δy each contain one eigenvalue. Putting $x + y = t$ we get (5.84).

Let $E'(x + y)$ be the probability that the interval $x + y$ will contain at most one eigenvalue, so that $E' - E = P(x + y)$ is the probability that $x + y$ will contain exactly one eigenvalue. By increasing x by δx in $P(x + y)$ and subtracting it from $P(x + y)$ we find that either the eigenvalue in $x + y$ moves in δx or a new eigenvalue appears in δx. Therefore

$$-\frac{\partial P(x+y)}{\partial x}\,\delta x = -\text{prob}\begin{Bmatrix} x+y, & 0 \\ \delta x & 1 \end{Bmatrix} + \text{prob}\begin{Bmatrix} x+y, & 1 \\ \delta x, & 1 \end{Bmatrix},$$

where

$$\text{prob}\begin{Bmatrix} x+y, & r \\ \delta x, & s \end{Bmatrix}$$

means the probability that $x + y$ and δx contain r and s eigenvalues, respectively. A second differentiation gives

$$\frac{\partial^2 P(x+y)}{\partial x\,\partial y}\,\delta x\,\delta y = -\text{prob}\begin{Bmatrix} x+y, & 0 \\ \delta x, & 1 \\ \delta y, & 1 \end{Bmatrix} + \text{prob}\begin{Bmatrix} x+y, & 1 \\ \delta x, & 1 \\ \delta y, & 1 \end{Bmatrix} - \text{prob}\begin{Bmatrix} x+y, & 0 \\ \delta x, & 1 \\ \delta y, & 1 \end{Bmatrix}.$$

Putting $P = E' - E$,

$$\text{prob}\begin{Bmatrix} x+y, & 0 \\ \delta x, & 1 \\ \delta y, & 1 \end{Bmatrix} = \frac{\partial^2 E}{\partial x\,\partial y}\,\delta x\,\delta y \quad \text{and} \quad x + y = t,$$

we obtain the probability density for the next nearest neighbor spacings or for the spacings of the alternate series (9.39).

A.12. Various Probability Distribution and Probability Density Functions

The first few terms in the power-series expansions of the quantities in Table A.1 are reproduced here.

$$E(t) = 1 - t + \frac{\pi^2}{36}t^3 - \frac{\pi^4}{1200}t^5 + \frac{\pi^4}{8100}t^6 + \frac{\pi^6}{70560}t^7 - \frac{\pi^6}{264600}t^8 + \cdots$$

$$E'(t) = 1 - \frac{\pi^2}{36}t^3 + \frac{\pi^4}{1200}t^5 - \frac{\pi^6}{70560}t^7 + \cdots$$

$$E_u(t) = 1 - t + \frac{\pi^2}{36}t^4 - \frac{\pi^4}{675}t^6 + \frac{\pi^6}{17640}t^8 + \cdots$$

$$F(t) = 1 + \frac{dE}{dt} = \frac{\pi^2}{12}t^2 - \frac{\pi^4}{240}t^4 + \frac{\pi^4}{1350}t^5 + \frac{\pi^6}{10080}t^6 + \cdots$$

$$F_u(t) = 1 + \frac{dE_u}{dt} = \frac{\pi^2}{9}t^3 - \frac{2\pi^4}{225}t^5 + \frac{\pi^6}{2205}t^7 + \cdots$$

$$p(t) = \frac{d^2E}{dt^2} = \frac{\pi^2}{6}t - \frac{\pi^4}{60}t^3 + \frac{\pi^4}{270}t^4 + \frac{\pi^6}{1680}t^5 + \cdots$$

$$p_u(t) = \frac{d^2E_u}{dt^2} = 4\frac{dE}{dt}\frac{dE'}{dt} = \frac{\pi^2}{3}t^2 - \frac{2\pi^4}{45}t^4 + \frac{\pi^6}{315}t^6 + \cdots$$

TABLE A.1

$\frac{\pi}{2}t$	t	$E(t)$	$E'(t)$	$E_u(t)$	$F(t)$	$F_u(t)$	$p(t)$	$p_u(t)$
0	0	1	1	1	0	0	0	0
0.1	0.064	0.936408	0.99993	0.93634	0.00330	0.0006	0.104	0.013
0.2	0.127	0.873239	0.99944	0.87275	0.01321	0.0023	0.207	0.052
0.3	0.191	0.810904	0.99811	0.80937	0.02947	0.0074	0.303	0.114
0.4	0.255	0.749796	0.99556	0.74647	0.05168	0.0172	0.395	0.197
0.5	0.318	0.690283	0.99142	0.68436	0.07947	0.0327	0.477	0.292
0.6	0.382	0.632698	0.98537	0.62344	0.11219	0.0546	0.549	0.397
0.7	0.446	0.577337	0.97712	0.56413	0.14920	0.0832	0.6117	0.504
0.8	0.509	0.524450	0.96645	0.50685	0.18982	0.1187	0.6630	0.608
0.9	0.573	0.474248	0.95318	0.45204	0.23338	0.1604	0.7032	0.704
1.0	0.637	0.426889	0.93720	0.40008	0.27908	0.2081	0.7308	0.791
1.2	0.764	0.341117	0.89695	0.30596	0.37410	0.3161	0.7547	0.898
1.4	0.891	0.267527	0.84597	0.22632	0.46962	0.4336	0.7396	0.943
1.6	1.018	0.205888	0.78542	0.16171	0.56114	0.5505	0.6933	0.883
1.8	1.146	0.155459	0.71717	0.11149	0.64529	0.6584	0.6255	0.789
2.0	1.273	0.115153	0.64362	0.07411	0.71986	0.7515	0.5445	0.661
2.2	1.400	0.083669	0.56742	0.04748	0.78376	0.8269	0.4587	0.520
2.4	1.528	0.059626	0.49116	0.02929	0.83681	0.8845	0.3750	0.387
2.6	1.655	0.041674	0.41734	0.01739	0.87956	0.9262	0.2978	0.272
2.8	1.782	0.028563	0.34802	0.00994	0.91307	0.9548	0.2301	0.182
3.0	1.910	0.019199	0.28475	0.005467	0.93863	0.9735	0.1730	0.115
3.2	2.037	0.012654	0.22856	0.002892	0.95760	0.9851	0.1267	0.070
3.4	2.164	0.008177	0.17995	0.001471	0.97133	0.9918	0.0906	0.040
3.6	2.292	0.005182	0.13897	0.000720	0.98104	0.9959	0.0631	0.022
3.8	2.419	0.003219	0.10522	0.000329	0.98772	0.9980	0.0429	0.012
4.0	2.546	0.001961	0.07817	0.000153	0.99223	0.9990	0.0286	0.006
4.2	2.674	0.001171	0.05693	0.000067	0.99518		0.0185	
4.4	2.801	0.0006858	0.04066	0.000028	0.99708		0.0117	
4.6	2.928	0.0003937	0.02847	0.000011	0.9983		0.0062	
4.8	3.055	0.0002216	0.01955	0.000004	0.9990		0.0030	
5.0	3.183	0.0001222	0.01316	0.000001	0.9994		0.002	

A.13. Some Applications of Gram's Result

Gram's result may be stated as follows [Courant and Hilbert, 1].

Let v_i , $i = 1, 2,..., m$, be m vectors and let $v_{i\alpha}$, $\alpha = 1, 2,..., n$, be their components along some basis. Form the scalar products

$$b_{ij} = (v_i , v_j) = \sum_{\alpha=1}^{n} v_{i\alpha} v_{j\alpha}^* , \qquad i, j = 1, 2,..., m;$$

then

$$\begin{vmatrix} b_{11} & b_{12} \cdots b_{1m} \\ b_{21} & b_{22} \cdots b_{2m} \\ \cdots & \cdots \\ b_{m1} & b_{m2} \cdots b_{mm} \end{vmatrix} = \frac{1}{m!} \sum_{\alpha_1, \alpha_2, \ldots, \alpha_m} \begin{vmatrix} v_{1\alpha_1} & \cdots v_{1\alpha_m} \\ \cdots & \cdots \\ v_{m\alpha_1} & \cdots v_{m\alpha_m} \end{vmatrix}^2 \tag{A13.1}$$

where on the right-hand side we sum over all possible ways of choosing α_1 , α_2 ,..., α_m among 1, 2,..., n. The summation over α may be finite, denumerable, or continuously infinite, the summation sign being replaced by an integration over a suitable measure.

One may convince oneself of the validity of this statement as follows. By the very definition of a determinant the left-hand side of (A13.1) is

$$\frac{1}{m!} \sum_{s_i, s_j} \pm b_{i_1 j_1} b_{i_2 j_2} \cdots b_{i_m j_m} = \frac{1}{m!} \sum_{s_i, s_j, \sigma} \pm v_{i_1 \alpha_1} \cdots v_{i_m \alpha_m} v_{j_1 \alpha_1}^* \cdots v_{j_m \alpha_m}^* ,$$

where the summation s_i is taken over all permutations i_1 , i_2 ,..., i_m and s_j over all permutations j_1 , j_2 ,..., j_m of the indices 1, 2,..., m; the summation (or integration) σ is over the indices α_1 , α_2 ,..., α_m . The sign is $+$ or $-$, depending on whether the parity of the permutation

$$\begin{pmatrix} i_1 & i_2 \cdots i_m \\ j_1 & j_2 \cdots j_m \end{pmatrix} \tag{A13.2}$$

is even or odd. Interchanging σ and the s, which can certainly be done, as the $\sum_s$ contain a finite number of terms, we get

$$\frac{1}{m!} \sum_{\sigma} \left| \sum_{s_i} \pm v_{i_1 \alpha_1} \cdots v_{i_m \alpha_m} \right|^2 ,$$

which is visibly the right-hand side of (A13.1). If the reader has any doubt about the sign $+$ or $-$, which now depends on a different

permutation, we may note that the parity of the permutation (A13.2) is equal to the product of the parities of the permutations

$$\begin{pmatrix} i_1 & i_2 & \cdots & i_m \\ \alpha_1 & \alpha_2 & \cdots & \alpha_m \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} j_1 & j_2 & \cdots & j_m \\ \alpha_1 & \alpha_2 & \cdots & \alpha_m \end{pmatrix}.$$

If we make the correspondence

$$v_{i\alpha} \to \sqrt{2}\,\varphi_{2i}(x), \qquad \sum_{\alpha} \to \int_{\theta}^{\infty} dx,$$

the scalar products become

$$b_{ij} = \sum_{\alpha} v_{i\alpha}\, v_{j\alpha}^{*} \to \int_{\theta}^{\infty} 2\,\varphi_{2i}(x)\,\varphi_{2j}(x)\,dx = g_{ij}$$

and the theorem (A13.1) yields (5.119). If we make the correspondence

$$v_{i\alpha} \to \sqrt{2}\,e^{-(1/2)y^2}\frac{y^{2i+1/2}}{(y^2+\theta^2)^{1/4}}, \qquad \sum_{\alpha} \to \int_{0}^{\infty} dy,$$

the scalar products become

$$b_{ij} = \sum_{\alpha} v_{i\alpha}\, v_{j\alpha}^{*} \to \int_{0}^{\infty} dy\, 2e^{-y^2}\frac{y^{2i+2j+1}}{(y^2+\theta^2)^{1/2}} = \eta_{i+j}$$

and the theorem (A13.1) yields (5.125). If we make the correspondence

$$v_{i\alpha} \to \varphi_{\alpha-1}(x_i), \qquad \sum_{\alpha} \to \sum_{\alpha=1}^{N}$$

we obtain (6.7) with (6.8). If we make the correspondence

$$v_{i\alpha} \to \varphi_i(x), \qquad \sum_{\alpha} \to \int_{\text{out}} dx \equiv \left(\int_{-\infty}^{-\theta} + \int_{\theta}^{\infty}\right) dx,$$

the scalar products become

$$b_{ij} = \int_{\text{out}} \varphi_i(x)\,\varphi_j(x)\,dx = \delta_{ij} - \int_{-\theta}^{\theta} \varphi_i(x)\,\varphi_j(x)\,dx$$

and we have (6.15). If we make the correspondence

$$v_{j\alpha} \to (2\pi)^{-1/2}\,e^{i\alpha\theta_j}, \qquad \sum_{\alpha} \to \sum_{\alpha=0}^{N-1},$$

the scalar products become

$$b_{jk} = \sum_{\alpha=0}^{N-1} (2\pi)^{-1}\,e^{i\alpha\theta_j}e^{-i\alpha\theta_k} = K_N(\theta_j\,,\,\theta_k),$$

and we get (9.56). If we make the correspondence

$$v_{j\alpha} \to (2\pi)^{-1/2} e^{ij\theta}, \qquad \sum_{\alpha} \to \int_{\alpha}^{2\pi-\alpha} d\theta,$$

the scalar products become

$$b_{jk} = (2\pi)^{-1} \int_{\alpha}^{2\pi-\alpha} d\theta \, e^{i(j-k)\theta},$$

and we get (9.63).

A.14. Power-Series Expansion of $I_m(\theta)$[†]

Expanding the integrand in (5.126) in powers of θ and integrating, we get

$$\eta_i(\theta) = \xi_{2i} - \tfrac{1}{2}\theta^2 \xi_{2i-2} + \tfrac{3}{8}\theta^4 \xi_{2i-4} + \cdots, \tag{A14.1}$$

where

$$\xi_{2i} = 2 \int_0^{\infty} e^{-y^2} y^{2i} \, dy = \Gamma(i + \tfrac{1}{2}).$$

Taking terms only up to θ^4, we put the expansion (A14.1) in the determinant of (5.125). We see that in writing the determinant as a sum of several terms many of them vanish because two rows are proportional. Thus we may write

$$I_m(\theta) = I_m(0)[1 - \tfrac{1}{2}\theta^2 a + \tfrac{3}{8}\theta^4 (b - c) + \cdots] \tag{A14.2}$$

where

$$I_m(0) = \frac{\sqrt{\pi}}{2} (m-1)! \prod_{1}^{2m-1} [\Gamma(1 + \tfrac{1}{2}j)]^{-1} \det[\xi_{2i+2j}]_{i,j=1,\ldots,m-1},$$

$$I_m(0)\, a = \begin{vmatrix} \xi_2 & \xi_4 & \cdots & \xi_{2m-2} \\ \xi_6 & \xi_8 & \cdots & \xi_{2m+2} \\ \cdots & \cdots & \cdots & \cdots \\ \xi_{2m} & \xi_{2m+2} & \cdots & \xi_{4m-4} \end{vmatrix}, \qquad I_m(0)\, b = \begin{vmatrix} \xi_0 & \xi_2 & \cdots & \xi_{2m-4} \\ \xi_6 & \xi_8 & \cdots & \xi_{2m+2} \\ \cdots & \cdots & \cdots & \cdots \\ \xi_{2m} & \xi_{2m+2} & \cdots & \xi_{4m-4} \end{vmatrix},$$

$$I_m(0)\, c = \begin{vmatrix} \xi_2 & \xi_4 & \cdots & \xi_{2m-2} \\ \xi_4 & \xi_6 & \cdots & \xi_{2m} \\ \xi_8 & \xi_{10} & \cdots & \xi_{2m+4} \\ \cdots & \cdots & \cdots & \cdots \\ \xi_{2m} & \xi_{2m+2} & \cdots & \xi_{4m-4} \end{vmatrix}.$$

[†] M. L. Mehta [1].

The evaluation of determinants whose elements are gamma functions is almost as easy as those whose elements are the successive factorials. Taking out all the common factors, one may reduce these determinants to the triangular form by simple operations. Thus we obtain

$$I_m(0) = 1, \qquad a = \tfrac{2}{3}(m-1), \qquad b = \tfrac{2}{3}m(m-1),$$

and

$$c = \tfrac{2}{45}(m-1)(m-2).$$

Putting these values in (A14.2), we get (5.127).

A.15. Proof of the Inequalities (5.130)

Let $u_i = \theta^2/y_i^2$ so that $u_i \geqslant 0$. The second inequality in (5.130)

$$\prod_1^n (1+u_i)^{-1/2} \leqslant 1$$

is immediate, for each factor in the product lies between 0 and 1.

The first inequality can be proved by induction. Suppose that

$$1 - \frac{1}{2}\sum_{i=1}^{r} u_i \leqslant \prod_{i=1}^{r} (1+u_i)^{-1/2}$$

is true for $1 \leqslant r \leqslant n$ and let us prove then that

$$1 - \frac{1}{2}\sum_{i=1}^{n+1} u_i \leqslant \prod_{i=1}^{n+1} (1+u_i)^{-1/2}.$$

Let $(1/2)\sum_1^{n+1} u_i \leqslant 1$, for otherwise the left-hand side will be negative and therefore smaller than the right-hand side, which is positive. Thus

$$\Big(1 - \frac{1}{2}\sum_1^r u_i\Big)\Big(1 - \frac{1}{2}\sum_{r+1}^{n+1} u_i\Big) \leqslant \prod_1^r (1+u_i)^{-1/2} \prod_{r+1}^{n+1} (1+u_i)^{-1/2}, \qquad r \leqslant n,$$

for both quantities in the product on both sides are positive. Therefore we have

$$\prod_{i=1}^{n+1} (1+u_i)^{-1/2} \geqslant 1 - \frac{1}{2}\sum_{i=1}^{n+1} u_i + \Big(\frac{1}{2}\sum_1^r u_i\Big)\Big(\frac{1}{2}\sum_{r+1}^{n+1} u_i\Big) \geqslant 1 - \frac{1}{2}\sum_{i=1}^{n+1} u_i\,.$$

Also it is easy to verify that

$$1 - \tfrac{1}{2}u_1 \leqslant (1 + u_1)^{-1/2}$$

and the proof is complete.

A.16. The Confluent Alternant

Let m_1 , m_2 ,..., m_N be positive integers and let their sum be

$$m = m_1 + m_2 + \cdots + m_N \; ; \; m_j \geqslant 1, \; j = 1, 2, ..., N.$$

Let us form a determinant Δ as follows. The first row consists of the powers of x_1

$$1, x_1 , x_1^2, x_1^3, ..., x_1^{m-1}.$$

The jth row for $1 \leqslant j \leqslant m_1$ is the $(j - 1)$th derivative of the first row. The $(m_1 + 1)$th row consists of the powers of x_2 .

$$1, x_2 , x_2^2, x_2^3, ..., x_2^{m-1}.$$

The $(m_1 + j)$th row for $1 \leqslant j \leqslant m_2$ is the $(j - 1)$th derivative of the $(m_1 + 1)$th row. The next m_3 rows are formed in a similar way from the powers of x_3 and their successive derivatives, and so on.

The determinant Δ is called the "confluent alternant" and it factorizes as

$$\Delta = \prod_{i=1}^{N} \Big[\prod_{r_i=1}^{m_i} \Gamma(r_i) \Big] \prod_{1 \leqslant i < j \leqslant N} (x_j - x_i)^{m_i m_j} \tag{A16.1}$$

This last statement can be proved in three steps:

1. The determinant Δ is a homogeneous polynomial in x_1 , x_2 ,..., x_N of degree

$$\tfrac{1}{2}m(m-1) - \sum_{i=1}^{N} \frac{1}{2} m_i(m_i - 1) = \tfrac{1}{2}\Big(m^2 - \sum_{1}^{N} m_i^2\Big) = \sum_{1 \leqslant i < j \leqslant N} m_i m_j .$$

2. It contains a factor

$$(x_j - x_i)^{m_i m_j},$$

which can be seen by observing that Δ and several of its derivatives with respect to x_j contain two identical rows on setting $x_i = x_j$

Hence they vanish. The lowest derivative that does not vanish is one in which each of the rows containing x_j is differentiated m_i times. Because there are m_j rows containing x_j, the lowest nonzero derivative at $x_i = x_j$ is of the order $m_i m_j$.

3. The term arising from the $m_j \times m_j$ blocks, $j = 1, 2, \ldots, N$, along the principal diagonal can be computed to give the constant coefficient in (A16.1). Equations 6.27 and 9.42 are special cases of (A16.1) when $m_1 = m_2 = \cdots = m_N = 2$.

A.17. Proof of Equations (8.12) and (8.28)

Let dM and dM' be connected by a similarity transformation

$$dM' = A\, dM A^{-1}, \tag{A.17.1}$$

where A is nonsingular. We now show that the Jacobian

$$J = \frac{\partial(dM'_{ij})}{\partial(dM_{ij})} \tag{A17.2}$$

is unity.

Considering the various matrix elements dM'_{ij} as components of a single vector (and similarly for dM_{ij}), we can write (A17.1) as

$$dM'_{ij} = \sum_{k,l} A_{ik} A^{-1T}_{jl}\, dM_{kl}$$

or

$$dM' = (A \times A^{-1T})\, dM, \tag{A17.3}$$

where the direct or the Kronecker product $(A \times B)$ is defined by the equation

$$(A \times B)_{ij,kl} = A_{ik} \cdot B_{jl}\,.$$

The Jacobian (A17.2) is thus seen to be equal to the determinant of $(A \times A^{-1T})$.

Now it can be easily verified that if P and Q are matrices of the order $(n \times n)$, whereas R and S are of the order $(m \times m)$,

$$(P \times R) \cdot (Q \times S) = (P \cdot Q) \times (R \cdot S), \tag{A17.4}$$

where a dot means ordinary matrix multiplication. From (A17.4) we obtain

$$(R \times P) = (R \times 1_n) \cdot (1_m \times P),$$

where 1_r is the $(r \times r)$ unit matrix. Taking determinants on both sides of this equation we have

$$\det(R \times P) = (\det R)^n (\det P)^m. \tag{A17.5}$$

Thus

$$J = \frac{\partial(dM'_{ij})}{\partial(dM_{ij})} = \det(A \times A^{-1T}) = [(\det A)(\det A^{-1})]^N = 1,$$

which establishes the result we wanted.

A.18. Proof of Equation (11.29)

$$\begin{aligned}
&\prod_{1 \leqslant j < k \leqslant 2m} (\nu_j + \nu_k) \prod_{j=1}^{m} (\nu_{2j-1} + \nu_{2j})^{-1} \\
&= (\nu_1 + \nu_3)(\nu_1 + \nu_4)(\nu_1 + \nu_5) \cdots (\nu_1 + \nu_{2m})(\nu_2 + \nu_3)(\nu_2 + \nu_4) \cdots (\nu_2 + \nu_{2m}) \\
&\quad \cdots (\nu_{2m-3} + \nu_{2m-1})(\nu_{2m-3} + \nu_{2m})(\nu_{2m-2} + \nu_{2m-1})(\nu_{2m-2} + \nu_{2m}) \\
&= (\nu_2 + \nu_4)^2(-\nu_2 + \nu_4)^2(\nu_2 + \nu_6)^2(-\nu_2 + \nu_6)^2 \cdots (\nu_2 + \nu_{2m})^2(-\nu_2 + \nu_{2m})^2 \\
&\quad \times (\nu_4 + \nu_6)^2(-\nu_4 + \nu_6)^2 \cdots (\nu_4 + \nu_{2m})^2(-\nu_4 + \nu_{2m})^2 \\
&\quad \cdots (\nu_{2m-2} + \nu_{2m})^2(-\nu_{2m-2} + \nu_{2m})^2 \\
&= [(\nu_2 + \nu_4)(\nu_2 + \nu_6) \cdots (\nu_2 + \nu_{2m})(\nu_4 + \nu_6) \cdots (\nu_4 + \nu_{2m}) \\
&\quad \cdots (\nu_{2m-2} + \nu_{2m})(-\nu_2 + \nu_4)(-\nu_2 + \nu_6) \cdots (-\nu_2 + \nu_{2m})(-\nu_4 + \nu_6) \\
&\quad \cdots (-\nu_4 + \nu_{2m}) \cdots (-\nu_{2m-2} + \nu_{2m})]^2 \\
&= [2.\ 3.\ 4 \cdots m.\ 4.\ 5 \cdots (m+1) \cdots (2m-2) \\
&\quad \times 1.\ 2 \cdots (m-1)\ 1.\ 2 \cdots (m-2) \cdots 1]^2 \\
&= \left[\frac{m!}{1!}\frac{(m+1)!}{3!}\frac{(m+2)!}{5!} \cdots \frac{(2m-2)!}{(2m-3)!}(m-1)!\,(m-2)! \cdots 1!\right]^2 \\
&= \left(\prod_{j=1}^{2m-2} j!\right)^2 \left[\prod_{j=1}^{m-1} (2j-1)!\right]^{-2} \\
&= \left(\prod_{j=1}^{2m-1} j!\right)^2 \left[\prod_{j=1}^{m} (2j-1)!\right]^{-2} \\
&= \prod_{j=1}^{m-1} [(2j)!]^2 .
\end{aligned}$$

A.19. Wilson's Proof of Equation (10.11)

Let us write

$$y_j = \prod_{l(\neq j)} \left(1 - \frac{z_j}{z_l}\right), \qquad u_j = y_j^{-1}, \tag{A19.1}$$

so that when

$$P(z) \equiv P(z_1, \ldots, z_N) = \prod_j (y_j)^{a_j}, \tag{A.19.2}$$

with $a_1, a_2, \ldots, a_N$ positive integers, is expanded in positive and negative powers of the z_j the constant term is

$$K = (2\pi i)^{-N} \int dz_1 \cdots \int dz_N (z_1 \cdots z_N)^{-1} P(z), \tag{A19.3}$$

where the contours of integration are unit circles taken counterclockwise. We have to prove that K is given by (10.11):

$$K = \frac{(a_1 + \cdots + a_N)!}{a_1! \cdots a_N!}. \tag{A19.4}$$

The proof can be conveniently divided into a set of lemmas.

Lemma A1. *A polynomial that is antisymmetric in $x_1, x_2, \ldots, x_l$ is necessarily of the form*

$$G(x_1, \ldots, x_l) \prod_{1 \leqslant i < j \leqslant l} (x_i - x_j),$$

where $G(x_1, x_2, \ldots, x_l)$ is a symmetric polynomial in all the x_j; $j = 1, 2, \ldots, l$.

A polynomial that changes sign when any pair of variables x_j, x_k is interchanged vanishes when $x_j = x_k$ so that it contains $(x_j - x_k)$ as a factor; taking all possible pairs x_j, x_k it contains, in fact, $\prod_{i<j} (x_i - x_j)$ as a factor. The remaining factors form a polynomial that must now be symmetric in all the x_j.

Lemma A2. *If $G(x_1, x_2, \ldots, x_l)$ is a ratio of two polynomials in $x_1, x_2, \ldots, x_l$, such that* (a) *its denominator is the product of all the differences $\prod_{i<j} (x_i - x_j)$,* (b) *it is homogeneous of degree r in x_j; $j = 1, 2, \ldots, l$, and* (c) *it is symmetric in all x_j, then G is* (a) *identically zero, if $r < 0$,* (b) *a homogeneous polynomial of degree r, if $r \geqslant 0$.*

Because G is symmetric, and its denominator is antisymmetric, the numerator must also be antisymmetric. Therefore by Lemma A1 it contains the entire denominator as a factor. Therefore G is a homogeneous polynomial of degree r. If $r < 0$, G must vanish identically.

Lemma A3. *By regarding the u_j as functions of z_j, we have the identity*

$$\sum_{j=1}^{N} u_j = 1. \tag{A19.5}$$

Since $\sum_{j=1}^{N} u_j$, as a function of the z_j, satisfies all the conditions of Lemma A2 with $r = 0$, it is a constant. To evaluate this constant put $z_1 = 0$, so that

$$u_1 = 1, \; u_2 = u_3 = \cdots = u_N = 0.$$

Lemma A4. *The Jacobian*

$$J(z) = \frac{\partial(\ln u_2, \ldots, \ln u_N)}{\partial(\ln z_1, \ldots, \ln z_{N-1})} \tag{A19.6}$$

is given by

$$J(z) = (N-1)!\, u_1 . \tag{A19.7}$$

Proof: The Jacobian J is the determinant of the matrix

$$J_{ij} = \frac{\partial(\ln u_i)}{\partial(\ln z_j)}$$

(rows numbered $i = 2$ to N, columns $j = 1$ to $N - 1$). Without changing the value of the determinant, we may add columns $j = 2$ through $N - 1$ to column 1; because $\ln u_i$ is homogeneous of degree zero in the z_j, we now have

$$J_{i1} = -\frac{\partial(\ln u_i)}{\partial(\ln z_N)}.$$

Move this column to the right and call it J_{iN}; thus

$$J = (-1)^{N-2} \det[J_{ij}]_{i,j=2,3,\ldots,N}.$$

Now

$$J_{ij} = -z_i(z_j - z_i)^{-1}, \qquad i \neq j,$$
$$J_{ii} = \sum_{k(\neq i)} z_i(z_k - z_i)^{-1}.$$

Evidently J is a ratio of two polynomials in the z_j, the denominator being a product of factors $(z_j - z_i)$. No such factor occurs twice, for a denominator $(z_j - z_i)$ appears only in the elements J_{ii}, J_{ij}, J_{ji}, and J_{jj}, so that the term $(z_j - z_i)^2$ occurs in the denominator of J only if it occurs in the 2×2 determinant $(J_{ii}J_{jj} - J_{ij}J_{ji})$. However, the last expression does not contain any $(z_j - z_i)^{-2}$. Furthermore, J has a factor $z_2, z_3, \ldots, z_N$; J is symmetric in $z_2, z_3, \ldots, z_N$ (but not z_1) and is homogeneous of degree 0 in the z_j. By the arguments of Lemma A3, we must have

$$J = C \prod_{j=2}^{N} z_j(z_j - z_1)^{-1} = Cu_1,$$

where C is a constant. Since (A19.4) is known to be true for the special case $a_1 = a_2 = \cdots = a_N = 0$, we must have

$$C = (N - 1)!$$

Lemma A5.

$$K = (2\pi i)^{-N+1} \int du_2 \cdots \int du_N (u_2 \cdots u_N)^{-1} \frac{P(z)}{J(z)}, \tag{A19.8}$$

where the z_j are expressed in terms of the u_j; $j = 2, \ldots, N$, by (A19.1) and (A19.5). The contours of integration are the circles

$$|u_i| = R_i, \tag{A19.9}$$

taken $(i - 1)$ times counterclockwise, where the R_i are arbitrary, except that they satisfy

$$R_{i+1} \ll R_i \ll 1, \qquad 2 \leqslant i \leqslant N - 1. \tag{A19.10}$$

Lemmas A3 and A4 show that J and P are single-valued functions of the u_j, and we do not have to specify the branch of the solution of (A19.1).

The lemma is proved by introducing the new variables one at a time and using mathematical induction. For more details the reader is referred to the original paper by Wilson [1].

Theorem A1. *K is given by* (A19.4).

Proof: By the foregoing lemmas

$$K = \frac{(2\pi i)^{-N+1}}{(N-1)!} \int du_2 \cdots \int du_N \prod_{j=1}^{N} u_j^{-a_j-1}, \tag{A19.11}$$

with

$$u_1 = 1 - \sum_{j=2}^{N} u_j .$$

It is somewhat annoying to have the integrand symmetric in the variables $u_1, u_2, \ldots, u_N$ (not all independent, of course) while the integrations are not. To overcome this slight inconvenience Wilson proceeded as follows. If $K(\lambda)$ were defined by (A19.11), but with

$$u_1 = \lambda - \sum_{j=2}^{N} u_j ,$$

where $|\lambda| = 1$, then by making a change of variables $u_j' = \lambda u_j$ we obtain

$$K(\lambda) = \lambda^{-a-1} K, \qquad a = \sum_{j=1}^{N} a_j . \tag{A19.12}$$

Also

$$(2\pi i)^{-1} \int u^{-n-1} e^u \, du = \frac{1}{n!} \tag{A19.13}$$

if the contour encloses the origin once counterclockwise. From the last two equations we get

$$K = (2\pi i)^{-1} a! \int e^{\lambda} K(\lambda) \, d\lambda.$$

Now interchange the order of integration, so that keeping $u_2, u_3, \ldots, u_N$ fixed, λ is integrated out first to get

$$(2\pi i)^{-1} \int e^{\lambda} \left(\lambda - \sum_{j=2}^{N} u_j\right)^{-a_1-1} d\lambda = (a_1!)^{-1} e^{u_2+\cdots+u_N}.$$

Thus

$$K = \frac{a!}{a_1!} \frac{(2\pi i)^{-N+1}}{(N-1)!} \int du_2 \cdots \int du_N \prod_{j=2}^{N} e^{u_j} u_j^{-a_j-1}$$

$$= \frac{a!}{a_1!\, a_2! \cdots a_N!} = \frac{(a_1 + \cdots + a_N)!}{a_1! \cdots a_N!} .$$

The factor $(N-1)!$ is canceled by the requirement that u_j goes round the circle $(j-1)$ times.

A.20. Proof That the Second Term in Equation (7.28) Drops Out on Summation

By reducing to a common denominator we have

$$\sum_{\substack{i,j,l \\ (i,j,l \text{ all different})}} (x_j - x_i)^{-1}(x_j - x_l)^{-1} = \frac{N(x_i)}{D(x_i)},$$

where

$$D(x_i) = \prod_{i<j} (x_i - x_j)$$

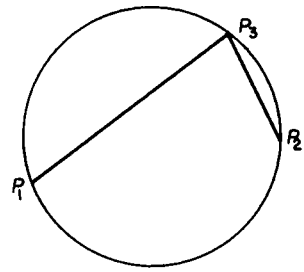
and N is a certain polynomial whose order is 2 less than that of D. As the left-hand side of this equation is symmetric in all the x_j and D is antisymmetric, we see that N must be antisymmetric.

Now the lowest order nonzero polynomial antisymmetric in the x_j is clearly $D(x_j)$. Because the order of N is less than that of D, we conclude that N must be identically zero.

A.21. Proof of the Inequality (10.5)

Consider all $\frac{1}{2}N(N-1)$ chords that join the N points $A_1, A_2, \ldots, A_N$ which lie on the unit circle. For definiteness let the angle variables of these points be in increasing order. Moreover, in the subsequent argument, whenever the index of any point exceeds N we subtract a multiple of N so that it is one of the numbers $1, 2, \ldots, N$. We want to maximize the product of the lengths of all the chords.

Let us divide the set of chords into a number of classes. In the first class we put the chords $A_1A_2, A_2A_3, \ldots, A_{N-1}A_N, A_NA_1$. In the second one we put $A_1A_3, A_3A_5, A_5A_7, \ldots$, and so on until we get back to A_1. If A_2 is left out, we construct a separate class of $A_2A_4, A_4A_6, \ldots$, until we get back to A_2. In the next class we put the chords $A_1A_4, A_4A_7, A_7A_{10}, \ldots$, until A_1 is repeated. If A_2 is left out, we construct a separate class $A_2A_5, A_5A_8, \ldots$. Similarly for A_3. And so on, until all the chords are exhausted. We maximize the product of the lengths of the chords belonging to a particular class. It is conceivable that the maximization conditions are different for different classes. If this occurs, we will really be in trouble.



Notice that if the points P_1, P_2 are fixed, but the point P_3 varies over the upper arc, the product of the chords P_1P_3, P_2P_3 is maximum when the two chords are equal. From this it follows that the product of the chords belonging to any one class is maximum when the points A_1, A_2,..., A_N lie at the vertices of a regular polygon. This condition is the same for any of the classes and the above mentioned trouble does not arise.

A.22. The Probability Density of the Spacings Resulting from a Random Superposition of *n* Unrelated Sequences of Energy Levels

Let ρ_i be the level density in the ith sequence and $p_i(\rho_i S)\,\rho_i\,dS$, the probability that a spacing in the ith sequence will have a value between S and $S + dS$. Because $p_i(x)$ is normalized and the level density is the inverse of the mean spacing, we have

$$\int_0^\infty p_i(x)\,dx = 1, \qquad \int_0^\infty x\,p_i(x)\,dx = 1. \tag{A22.1}$$

Let $F_i(x)$ and $E_i(x)$ be defined by

$$F_i(x) = \int_0^x p_i(y)\,dy = 1 - \int_x^\infty p_i(y)\,dy = 1 - \int_0^\infty p_i(x+y)\,dy \tag{A22.2}$$

and

$$E_i(x) = \int_x^\infty [1 - F_i(y)]\,dy = \int_x^\infty \left[\int_0^\infty p_i(y+z)\,dz\right] dy$$

$$= \int_0^\infty\!\!\int p_i(x+y+z)\,dy\,dz = \int_0^\infty y\,p_i(x+y)\,dy, \tag{A22.3}$$

so that $F_i(\rho_i S)$ is the probability that a spacing in the ith sequence is less than or equal to S and $E_i(\rho_i S)$ is the probability that a given interval of length S will not contain any of the levels belonging to the sequence i.

Consider the system resulting from the superposition of n sequences. The total density is

$$\rho = \sum_i \rho_i \,. \tag{A22.4}$$

Let $P(\rho S)\,\rho\, dS$ be the probability that a spacing will lie between S and $S + dS$. Analogously to (A22.2) and (A22.3), we introduce the functions $F(x)$ and $E(x)$ by

$$F(x) = \int_0^x P(y)\, dy = 1 - \int_0^\infty P(x + y)\, dy \tag{A22.5}$$

and

$$E(x) = \int_x^\infty [1 - F(y)]\, dy = \iint_0^\infty P(x + y + z)\, dy\, dz = \int_0^\infty y\, P(x + y)\, dy. \tag{A22.6}$$

From the observation that $E(\rho S)$ is the probability that a given interval of length S will not contain any of the levels and the randomness of the superposition we have

$$E(\rho S) = \prod_i E_i(\rho_i S), \tag{A22.7}$$

Introducing the fractional densities

$$f_i = \frac{\rho_i}{\rho}, \qquad \sum_i f_i = 1 \tag{A22.8}$$

and the variable $x = \rho S$, we have

$$E(x) = \prod_i E_i(f_i x). \tag{A22.9}$$

By differentiating (A22.9) twice, we obtain

$$P(x) = \frac{d^2E}{dx^2} = E(x) \left\{ \sum_i f_i^2 \frac{p_i(f_i x)}{E_i(f_i x)} + \left[\sum_i f_i \frac{1 - F_i(f_i x)}{E_i(f_i x)} \right]^2 - \sum_i \left(f_i \frac{1 - F_i(f_i x)}{E_i(f_i x)} \right)^2 \right\}, \tag{A22.10}$$

which was the purpose of this appendix.

We now consider three special cases.

1. Let the levels in each of the sequences be independent of one another so that $p_i(x) = e^{-x}$. In this case

$$1 - F_i(x) = E_i(x) = e^{-x} \tag{A22.11}$$

and (A22.10) yields $P(x) = e^{-x}$, which verifies the obvious fact that the random superposition of sequences of independent random levels produces a sequence of independent random levels.

2. Let all fractional densities be equal to $1/n$ and take the limit as n goes to ∞. Let $x = ny$ so that

$$P(ny) = [E(y)]^n \left\{ \frac{1}{n} \frac{p(y)}{E(y)} + \left(1 - \frac{1}{n}\right) \left[\frac{1 - F(y)}{E(y)} \right]^2 \right\}. \tag{A22.12}$$

From (A22.1), (A22.2), and (A22.3) we have

$$F(0) = 0, \qquad E(0) = 1, \qquad E'(0) = -1. \tag{A22.13}$$

Therefore, taking the limit as $n \to \infty$, $y \to 0$, whereas $ny = x$ is fixed, we see that the terms in the square brackets tend to 1, while from

$$E(y) \approx E(0) + yE'(0) + \cdots = 1 - y + \cdots,$$

keeping only the first term,

$$[E(y)]^n \approx (1 - y)^n = \left(1 - \frac{x}{n}\right)^n \to e^{-x}. \tag{A22.14}$$

This is a verification of the heuristic reasoning that if the number n of the sequences to be superimposed is large, a level belonging to a sequence will almost certainly be followed by a level of another sequence and these two levels will be independent, whatever $p(x)$ may be.

3. For the "Wigner surmise"

$$p(x) = \frac{\pi}{2} x \exp\left(-\frac{\pi}{4} x^2\right) \tag{A22.15}$$

we have

$$1 - F(x) = \exp\left(-\frac{\pi}{4} x^2\right), \qquad E(x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^{(1/2)x\sqrt{\pi}} e^{-y^2}\, dy. \tag{A22.16}$$

For the correct spacing distribution the functions $F(x)$ and $E(x)$ are tabulated in Appendix A.12.

Because $p(0) = 0$, we have

$$P(0) = 1 - \sum_{i=1}^{n} f_i^2 \tag{A22.17}$$

and, in particular, $P(0) \neq 0$.

A.23. Some Properties Connected with Symmetric and Antisymmetric Unitary Matrices

A few lemmas needed at various stages are given here for completeness.

1. If x is a real $n \times 1$ column matrix normalized to unity, $\sum_{i=1}^{n} x_i^2 = 1$, an $n \times n$ real orthogonal matrix can be constructed (in many ways) whose first column is x.

Take any $(n - 1)$ real $n \times 1$ column matrices which together with the given column x form a linearly independent system of n columns. This is possible due to the invariance of the number of dimensions of the space under consideration. Now apply Schmidt's orthonormalization procedure to these column matrices, starting with x.

2. If A is an $n \times n$ Hermitian (anti-Hermitian) matrix with complex elements, all its eigenvalues are real (pure imaginary).

In particular the eigenvalues of a symmetric (antisymmetric) real matrix are all real (pure imaginary).

Let λ be an eigenvalue and x, the corresponding eigenvector; that is,

$$Ax = \lambda x. \tag{A23.1}$$

Multiplying by $x^\dagger$ from the left gives

$$x^\dagger Ax = \lambda x^\dagger x. \tag{A23.2}$$

The Hermitian conjugate of the last equation reads

$$x^\dagger A^\dagger x = \lambda^* x^\dagger x. \tag{A23.3}$$

Because $x^\dagger x \neq 0$, we conclude by comparison of (A23.2) and (A23.3) that $\lambda = \lambda^*$ or $\lambda = -\lambda^*$ depending on whether $A^\dagger = A$ or $A^\dagger = -A$.

3. If A is a real symmetric matrix, all its eigenvectors may be taken to be real.

In fact, because the eigenvalue λ is real, by separating the real and imaginary parts of (A23.1) we get real eigenvectors.

4. If A is a real symmetric matrix, there exists a real orthogonal matrix B that diagonalizes it; that is,

$$B^TAB = E, \qquad B^TB = B^\dagger B = 1, \tag{A23.4}$$

where E is diagonal.

Let λ_1 be an eigenvalue and x, a normalized real eigenvector corresponding to it. Construct a real orthogonal matrix B_1 whose first column is x; $B_1 = [x, Y]$. Transforming A by B_1, we get

$$B_1{}^TAB_1 = \begin{bmatrix} x^T \\ Y^T \end{bmatrix} A[x \quad Y] = \begin{bmatrix} \lambda_1 & x^TAY \\ 0 & Y^TAY \end{bmatrix},$$

where we have used the fact that $Y^Tx = 0$. Because $B_1{}^TAB$ is symmetric, $x^TAY = 0$ and

$$B_1{}^TAB_1 = \begin{bmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{bmatrix},$$

where $A_1 = Y^TAY$ is a real symmetric matrix whose order is one less than that of A. The process can be repeated on A_1. Let B_2' be an orthogonal matrix with

$$B_2'{}^TA_1B_2' = \begin{bmatrix} \lambda_2 & 0 \\ 0 & A_2 \end{bmatrix},$$

where A_2 is a real symmetric $(n-2) \times (n-2)$ matrix. Putting

$$B_2 = B_1 \begin{bmatrix} 1 & 0 \\ 0 & B_2' \end{bmatrix},$$

we see that

$$B_2{}^TAB_2 = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & A_2 \end{bmatrix}.$$

The construction of B can thus be carried out step by step.

5. If A is a real antisymmetric matrix, there exists a real orthogonal matrix B such that it transforms A into the canonical form

$$B^TAB = E, \qquad B^TB = B^\dagger B = 1, \tag{A23.5}$$

where E is a real antisymmetric matrix whose only nonzero elements lie in the 2×2 blocks along the principal diagonal; that is,

$$E_{2j,2j-1} = -E_{2j-1,2j} = \mu_j, \qquad \mu_j \text{ real},$$

and

$$E_{jk} = 0, \qquad \text{otherwise.} \tag{A23.6}$$

Let $x + iy$ be the eigencolumn corresponding to the eigenvalue $\lambda = -i\mu$ with x, y, and μ real [see 2]:

$$A(x + iy) = -i\mu(x + iy). \tag{A23.7}$$

Separating the real and imaginary parts, we get

$$Ax = \mu y, \qquad Ay = -\mu x, \tag{A23.8}$$

so that

$$x^TAx = \mu x^Ty, \qquad x^Ty = y^Tx = \text{a number.}$$

On taking the transpose of this equation we find that the left-hand side changes sign. Therefore the right-hand side must be zero. Thus, if $\mu \neq 0$, the real and imaginary parts of the eigenvector are orthogonal.

$$x^Ty = y^Tx = 0, \quad \text{for} \quad \mu \neq 0. \tag{A23.9}$$

On the other hand, (A23.8) leads to

$$\mu y^Ty = y^TAx, \qquad \mu x^Tx = -x^TAy.$$

Taking the transpose of the first of these equations and comparing it with the second, we see that if $\mu \neq 0$ we can normalize the eigenvector in a way to satisfy

$$x^Tx = y^Ty = 1. \tag{A23.10}$$

Now let $\mu \neq 0$ and construct a real orthogonal matrix B_1 whose first two columns are x and y: $B_1 = [x \quad y \quad W]$. This is possible according to (1). Transforming A by B_1, we get

$$B_1^TAB_1 = \begin{bmatrix} x^T \\ y^T \\ W^T \end{bmatrix} A[x \quad y \quad W] = \begin{bmatrix} 0 & -\mu & x^TAW \\ \mu & 0 & y^TAW \\ 0 & 0 & W^TAW \end{bmatrix},$$

where we have used (A23.9) and (A23.10) and the orthogonality of W to both x and y. Because $B_1{}^T A B_1$ is antisymmetric, we must have $x^T A W = y^T A W = 0$, and

$$B_1{}^T A B_1 = \begin{bmatrix} 0 & -\mu & 0 \\ \mu & 0 & 0 \\ 0 & 0 & W^T A W \end{bmatrix} = \begin{bmatrix} 0 & -\mu \\ \mu & 0 \end{bmatrix} \dot{+} [W^T A W]. \quad \text{(A23.11)}$$

The process can be repeated for all nonzero eigenvalues, and we get

$$B^T A B = \begin{bmatrix} 0 & -\mu_1 \\ \mu_1 & 0 \end{bmatrix} \dot{+} \cdots \dot{+} \begin{bmatrix} 0 & -\mu_l \\ \mu_l & 0 \end{bmatrix} \dot{+} A_1 , \quad \text{(A23.12)}$$

where A_1 is a real antisymmetric matrix with all its eigenvalues equal to zero. Hence A_1 must itself be zero. Thus we have found a real orthogonal matrix B which transforms A to the canonical form.

6. Any number of commuting real symmetric matrices can be diagonalized by the same real orthogonal matrix. Any number of commuting real antisymmetric matrices can be transformed to the canonical form (A23.5), (A23.6), by the same real orthogonal matrix.

As the commuting matrices all have the same eigenvectors, the process followed in (4) and (5), if applied to all matrices, will transform them to the desired form: diagonal for real symmetric and canonical (A23.5), (A23.6) for real antisymmetric.

7. If A is a unitary symmetric matrix with complex elements then there exists a real orthogonal matrix B that diagonalizes A;

$$B^T A B = E, \qquad B^T B = B^\dagger B = 1, \quad \text{(A23.13)}$$

where E is diagonal. The diagonal elements of E are the complex number $e^{i\theta_j}$ lying on the unit circle.

Let

$$A = A_1 + iA_2 \quad \text{(A23.14)}$$

where A_1 and A_2 are real. As A is symmetric so are both A_1 and A_2. The unitarity of A gives

$$(A_1 - iA_2)(A_1 + iA_2) = 1. \quad \text{(A23.15)}$$

By separating real and imaginary parts of (A23.15) we get

$$A_1{}^2 + A_2{}^2 = 1, \qquad A_1 A_2 - A_2 A_1 = 0. \quad \text{(A23.16)}$$

Thus the real symmetric matrices A_1 and A_2 commute and therefore can be diagonalized both by the same real orthogonal matrix B;

$$B^T A_1 B = E_1 , \qquad B^T A_2 B = E_2 , \tag{A23.17}$$

where E_1 and E_2 are diagonal and real [see (2) and (6)]. If e_j and e_j' are the diagonal elements of E_1 and E_2, respectively, the first relation of (A23.16) gives

$$e_j^2 + e_j'^2 = 1,$$

so that we can write

$$e_j = \cos\theta_j , \qquad e_j' = \sin\theta_j . \tag{A23.18}$$

Equations A23.14, A23.17, and A23.18, then, give the diagonal elements of E as

$$e_j + ie_j' = e^{i\theta_j}. \tag{A23.19}$$

8. If A is a unitary antisymmetric matrix with complex elements, then there exists a real orthogonal matrix B such that

$$B^T A B = E, \qquad B^T B = B^\dagger B = 1, \tag{A23.20}$$

where E is an antisymmetric matrix with the canonical form (A23.6). The elements of E are

$$E_{2j,2j-1} = -E_{2j-1,2j} = e^{i\theta_j}, \tag{A23.21}$$

with θ_j real. All other elements of E are zero.

Let

$$A = A_1 + iA_2 \tag{A23.22}$$

where A_1 and A_2 are real. Because A is antisymmetric, so are A_1 and A_2. The unitarity of A, on separating the real and imaginary parts, gives the equations

$$A_1^2 + A_2^2 = -1, \qquad A_1 A_2 - A_2 A_1 = 0. \tag{A23.23}$$

Thus the real antisymmetric matrices A_1 and A_2 commute. A real orthogonal matrix can therefore be found such that

$$B^T A_1 B = E_1 , \qquad B^T A_2 B = E_2 ; \tag{A23.24}$$

E_1 and E_2 both having the form (A23.6). Let the elements in the subdiagonals of E_1 and E_2 be e_j and e_j', respectively:

$$(E_1)_{2j,2j-1} = e_j, \qquad (E_2)_{2j,2j-1} = e_j'. \tag{A23.25}$$

The first of the equations (A23.23) then gives

$$e_j^2 + e_j'^2 = 1,$$

so that we may choose

$$e_j = \cos\theta_j, \qquad e_j' = \sin\theta_j. \tag{A23.26}$$

Equations A23.22, A23.24, A23.25, and A23.26 then give the subdiagonal elements of E, as stated.

9. If A is a unitary symmetric matrix, then there exist unitary symmetric matrices U and V such that

$$A = UU^T \quad \text{and} \quad VAV^T = 1. \tag{A23.27}$$

Let the real orthogonal matrix B diagonalize A:

$$A = BEB^T$$

and choose

$$U = BE^{1/2}, \qquad V = U^{-1},$$

where $E^{1/2}$ is a diagonal matrix with diagonal elements $\exp(\frac{1}{2}i\theta_j)$.

The matrices U and V are not unique. We may replace U by UR, where R is any real orthogonal matrix. Moreover, the relation $VU = 1$ need not be valid.

10. Given a unitary symmetric matrix A, we can find a unitary symmetric matrix U (in many ways) such that

$$A = U^2, \qquad U^\dagger = U^* = U^{-1}.$$

Choose

$$U = BE^{1/2}B^T.$$

if the real orthogonal B diagonalizes A.

11. If A is a unitary antisymmetric matrix, then there exist unitary matrices U and V such that

$$A = UZU^T \quad \text{and} \quad VAV^T = Z, \tag{A23.28}$$

where Z is the matrix given by (2.21).

Let the real orthogonal matrix B transform A to the canonical form (A23.20), (A23.21).

$$A = BEB^T, \qquad B^TB = B^\dagger B = 1, \tag{A23.29}$$

$$\begin{aligned} E_{2j,2j-1} &= -E_{2j-1,2j} = e^{i\theta_j} \\ E_{jk} &= 0, \qquad \text{otherwise.} \end{aligned} \tag{A23.30}$$

Let X be the diagonal matrix with diagonal elements $\exp(\frac{1}{2}i\theta_j)$, each repeated twice. Then we may choose

$$U = BX, \qquad V = U^{-1}$$

which visibly have the required properties. The matrices U and V are not unique. We may replace U by UR, where R is any unitary matrix satisfying the relation

$$RZR^{-1} = Z.$$

Moreover, the relation $VU = 1$ need not be valid.

In the foregoing lemmas we can make these replacements: "real" → "quaternion real," "$x.y = \sum_j x_j^* y_j$" → "$x.y = \sum_j x_j^\dagger y_j$" "imaginary" → "quaternion imaginary," "transpose" → "dual," "symmetric" → "self-dual," "antisymmetric" → "antiself-dual," "real symmetric" → "Hermitian," "real antisymmetric" → "anti-Hermitian," "real orthogonal" → "symplectic," "diagonal" → "diagonal and scalar," and "unitary" → "unitary"; and we get new propositions about matrices with quaternion elements. (For the definitions of terms see Section 2.4.) The proofs also remain valid with the same replacements. We shall need the following in particular.

12. If A is a quaternion real self-dual (i.e., Hermitian) matrix, then there exists a symplectic matrix B such that

$$B^RAB = E, \tag{A23.31}$$

where E is diagonal, real, and scalar. The superscript R on B denotes its dual [cf. (2.36)].

13. If A is a unitary self-dual matrix, then there exist unitary matrices U and V such that

$$A = UU^R \quad \text{and} \quad VAV^R = 1. \tag{A23.32}$$

14. If A is a unitary self-dual matrix, then there exists a symplectic matrix B which diagonalizes A; that is,

$$B^R AB = E, \tag{A23.33}$$

where E is diagonal and scalar. The diagonal elements of E are the complex numbers $e^{i\theta_j}$ (θ_j real), each repeated twice.

A.24. Evaluation of the Integral (12.9) for Complex Matrices

In the exposition given here and in Appendix A.26 we follow the method of Ginibre [1]. We start with the proposition:

Any complex nonsingular $N \times N$ matrix X can be expressed in one and only one way as

$$X = UYV, \tag{A24.1}$$

where U is a unitary matrix, Y is a triangular matrix with all diagonal elements equal to unity, $y_{ij} = 0, i > j$; $y_{ii} = 1$, and V is a diagonal matrix with real positive diagonal elements.

Proof: Given $X = [x_{ij}]$, we solve the homogeneous linear equations in u_{rj} :

$$\begin{aligned} \sum_{j=1}^{N} u_{rj} x_{ji} &= 0, \qquad i < r, \\ \sum_{j=1}^{N} u_{rj} u_{ij}^* &= 0, \qquad i > r, \end{aligned} \tag{A24.2}$$

successively for $r = N, N-1, \ldots, 1$. Because the number of unknowns is always one greater than the number of equations, the u_{rj} for a fixed r are not all zero. We may normalize them to satisfy

$$\sum_{j=1}^{N} u_{rj} u_{rj}^* = 1, \tag{A24.3}$$

without disturbing the equalities (A24.2). Thus we have found a unitary matrix $U_1 = [u_{rj}]$ such that $Y_1 = U_1 X$ is triangular $(Y_1)_{ij} = 0, i > j$. Because X is nonsingular, all diagonal elements of Y_1 are different from zero. Writing the diagonal elements of Y_1 in the polar form, $(Y_1)_{jj} = v_j \exp(i\theta_j)$, we construct two diagonal

matrices, one unitary, U_2, with diagonal elements $(U_2)_{jj} = \exp(i\theta_j)$, and the other positive definite, V, with diagonal elements $V_{jj} = u_j$. Putting $Y = U_2^\dagger Y_1 V^{-1}$ and $U = U_1^\dagger U_2$, we see that $X = UYV$, where U, Y, and V have the properties required in the proposition. Next let $X = UYV = U'Y'V'$; then $U'^\dagger U = Y'V'V^{-1}Y^{-1}$ is unitary (left-hand side) as well as triangular with real positive diagonal elements (right-hand side). Thus $U'^\dagger U$ and $Y'V'V^{-1}Y^{-1}$ are unit matrices. A comparison of the diagonal elements on the two sides of $YV = Y'V'$ now gives $V = V'$. The decomposition $X = UYV$ is therefore unique.

Using the fact that $S = XEX^{-1}$, $U^\dagger U = 1$, and $EV = VE$, we can write

$$\operatorname{tr}(S^\dagger S) = \operatorname{tr}[E^\dagger X^\dagger XE(X^\dagger X)^{-1}] = \operatorname{tr}[E^\dagger Y^\dagger YE(Y^\dagger Y)^{-1}], \tag{A24.4}$$

$$dA = X^{-1}\,dX = V^{-1}Y^{-1}(U^{-1}\,dU)\,YV + V^{-1}Y^{-1}\,dYV + V^{-1}\,dV. \tag{A24.5}$$

The volume element $\prod_{i\neq j} dA_{ij}^{(0)}\,dA_{ij}^{(1)}$ needed in (12.9) is the quotient of the volume element $\prod_{i,j} dA_{ij}^{(0)}\,dA_{ij}^{(1)}$ by that of the set of all complex diagonal matrices. We put aside the quantities that do not depend on the eigenvalues, for they give only multiplicative constants. All of these constants can be adjusted in the final normalization. From (A24.5) and the structure of Y and V we see that

$$\prod_{i,j} dA_{ij}^{(0)}\,dA_{ij}^{(1)} = \prod_{i<j} (Y^{-1}\,dY)_{ij}^{(0)}\,(Y^{-1}\,dY)_{ij}^{(1)}\,a, \tag{A24.6}$$

where a depends only on U and V. We replace $\prod_{i\neq j} dA_{ij}^{(0)}\,dA_{ij}^{(1)}$ in (12.9) with $\prod_{i<j} (Y^{-1}\,dY)_{ij}^{(0)}\,(Y^{-1}\,dY)_{ij}^{(1)}$ and calculate

$$\int \exp[-\operatorname{tr}(E^\dagger HEH^{-1})] \prod_{i<j} (Y^{-1}\,dY)_{ij}^{(0)}\,(Y^{-1}\,dY)_{ij}^{(1)}, \tag{A24.7}$$

where

$$H = Y^\dagger Y. \tag{A24.8}$$

The matrix H is Hermitian. Any of its upper left diagonal block of size n is obtained from the upper left diagonal block of Y of the same size: $H_n = Y_n^\dagger Y_n$. Therefore, for every n, $\det H_n = 1$, and the diagonal elements H_{nn} are successively and uniquely determined once the off-diagonal elements are given. Thus we need $N(N-1)$ real parameters to specify H, the same number needed to specify Y.

We can further convince ourselves that the correspondence of Y and H is one to one. However, we do not need this last result, for we are omitting the constants anyway. Now we make a change of variables. First, because $\det Y = 1$,

$$\prod_{i<j} (Y^{-1}\,dY)^{(0)}_{ij}\,(Y^{-1}\,dY)^{(1)}_{ij} = \prod_{i<j} dY^{(0)}_{ij}\,dY^{(1)}_{ij}. \tag{A24.9}$$

Next, we take $H^{(0)}_{ij}$, $H^{(1)}_{ij}$ for $i<j$ as independent variables. The superscripts (0) and (1) denote, as always, the real and the imaginary parts. From

$$H_{ij} = Y_{ij} + \sum_{k<i} Y^*_{ki} Y_{kj}, \qquad i<j,$$

one can easily calculate the Jacobian of the transformation from Y to H, it being unity. The integral (12.9) is

$$J = C \int \exp[-\mathrm{tr}(E^\dagger H E H^{-1})] \prod_{i<j} dH^{(0)}_{ij}\,dH^{(1)}_{ij}, \tag{A24.10}$$

where C is a constant.

The integration over H is done in N steps. At every step we integrate over the variables of the last column and thus decrease by one the size of the matrix, whose structure remains the same. For this we need the recursion relation (A24.17) derived below.

Let $H' = Y_n{}^\dagger Y_n$, $E' = [z_i\delta_{ij}]_{i,j=1,2,\ldots,n}$, be the relevant matrices of order n and H, E be those obtained from H', E' by removing the last row and last column. Let the Greek indices run from 1 to n, and the Latin indices from 1 to $n-1$. Let $\Delta'_{\alpha\beta}$ be the cofactor of $H'_{\alpha\beta}$ in H' and Δ_{ij}, the cofactor of H_{ij} in H. Let $g_i = H'_{in}$. Because $\det H' = \det H = 1$, we have

$$\Delta'_{\alpha\beta} = H'^{-1}_{\beta\alpha}, \qquad \Delta_{ij} = H^{-1}_{ji}. \tag{A24.11}$$

Expanding $\det H'$, Δ'_{in}, Δ'_{ij} by the last row and last column, we have

$$1 = H'_{nn} - \sum_{i,j} g^*_i g_j \Delta_{ji}, \tag{A24.12}$$

$$\Delta'_{in} = -\sum_l \Delta_{il} g^*_l, \tag{A24.13}$$

$$\Delta'_{ij} = H'_{nn}\Delta_{ij} - \sum_{l,k} g^*_k g_l \Delta^{lk}_{ij}, \tag{A24.14}$$

where Δ_{ij}^{lk} is the cofactor obtained from H by removing the ith and lth rows and jth and kth columns. Sylvester's theorem [Gantmacher, 1; Mohr, 1] expresses Δ_{ij}^{lk} in terms of Δ_{rs}

$$\Delta_{ij}^{lk} = \Delta_{ij}\,\Delta_{lk} - \Delta_{ik}\,\Delta_{lj}\,. \tag{A24.15}$$

In writing (A24.15), we have replaced det H by unity on the left-hand side. Let

$$\phi_n = \mathrm{tr}(E'^{\dagger}H'E'H'^{-1}) = \sum_{\alpha,\beta} z_\alpha^* z_\beta H'_{\alpha\beta}\,\Delta'_{\alpha\beta}\,. \tag{A24.16}$$

Separating the last row and last column and making use of (A24.11) to (A24.15), we get, after some simplification,

$$\phi_n = |\,z_n\,|^2 + \phi_{n-1} + \langle g^*|\,H^{-1}(E^{\dagger} - z_n^*)\,H(E - z_n)\,H^{-1}\,|\,g\rangle, \tag{A24.17}$$

where

$$\langle g^*|\,B\,|g\rangle = \sum_{i,j} g_i^* B_{ij} g_j\,. \tag{A24.18}$$

Substituting (A24.17) for $n = N$ in (A24.10), we get

$$J = Ce^{-|z_N|^2}\int e^{-\phi_{N-1}} \prod_{1\leqslant i<j\leqslant N-1} dH_{ij}^{(0)}\,dH_{ij}^{(1)}$$
$$\times \int \exp[-\langle g^*|\,H^{-1}(E^{\dagger} - z_N^*)\,H(E - z_N)\,H^{-1}\,|\,g\rangle] \prod_{1\leqslant i\leqslant N-1} dg_i^{(0)}\,dg_i^{(1)}. \tag{A24.19}$$

The last integral is immediate and gives

$$\pi^{N-1}\{\det[H^{-1}(E^{\dagger} - z_N^*)\,H(E - z_N)\,H^{-1}]\}^{-1} = \pi^{N-1}\prod_{i=1}^{N-1} |\,z_i - z_N\,|^{-2}. \tag{A24.20}$$

The process can be repeated N times and we finally get

$$J = C\exp\left(-\sum_1^N |\,z_i\,|^2\right) \prod_{1\leqslant i<j\leqslant N} |\,z_i - z_j\,|^{-2} \tag{A24.21}$$

where C is a new constant.

A.25. A Few Remarks about the Eigenvalues of a Quaternion Real Matrix and its Diagonalization†

A quaternion-real matrix S is one whose elements are real quaternions (cf. Chapter 2). To emphasize this aspect we say that S is a q-matrix. If we replace the elements of $N \times N$ q-matrix S by their 2×2 matrix representation (2.23), we get a $2N \times 2N$ matrix S with complex elements, a c-matrix. A real quaternion is represented by a 2×2 matrix of the form

$$\begin{bmatrix} a & -b^* \\ b & a^* \end{bmatrix},$$

so that the c-matrix S has the form

$$S = \begin{bmatrix} a_{ij} & -b_{ij}^* \\ b_{ij} & a_{ij}^* \end{bmatrix}. \tag{A25.1}$$

The c-matrix S has $2N$ (complex or c) eigenvalues and at least one (complex or c) eigenvector belonging to each distinct eigenvalue. If $\begin{bmatrix} x_i \\ y_i \end{bmatrix}$ is a c-eigenvector of S belonging to the c-eigenvalue α,

$$\sum_j \begin{bmatrix} a_{ij} & -b_{ij}^* \\ b_{ij} & a_{ij}^* \end{bmatrix} \begin{bmatrix} x_j \\ y_j \end{bmatrix} = \alpha \begin{bmatrix} x_i \\ y_i \end{bmatrix}. \tag{A25.2}$$

$$\alpha = \alpha_0 + i\alpha_1 ; \qquad \alpha_0 , \alpha_1 \text{ real}, \tag{A25.3}$$

that is,

$$\sum_j (a_{ij}x_j - b_{ij}^* y_j) = \alpha x_i$$

$$\sum_j (b_{ij}x_j + a_{ij}^* y_j) = \alpha y_i. \tag{A25.2'}$$

Then, taking the complex conjugate of these equations and changing the order in which they are written, we see that

$$\sum_j \begin{bmatrix} a_{ij} & -b_{ij}^* \\ b_{ij} & a_{ij}^* \end{bmatrix} \begin{bmatrix} -y_j^* \\ x_j^* \end{bmatrix} = \alpha^* \begin{bmatrix} -y_i^* \\ x_i^* \end{bmatrix}, \tag{A25.4}$$

† Study [1].

that is $[\begin{smallmatrix} -y_i \\ x_i^* \end{smallmatrix}]$ is another c-eigenvector belonging to the c-eigenvalue α^*. We now write (A25.2) and (A25.4) together:

$$\sum_j \begin{bmatrix} a_{ij} & -b_{ij}^* \\ b_{ij} & a_{ij}^* \end{bmatrix} \begin{bmatrix} x_j & -y_j^* \\ y_j & x_j^* \end{bmatrix} = \begin{bmatrix} x_i & -y_i^* \\ y_i & x_i^* \end{bmatrix} \begin{bmatrix} \alpha & 0 \\ 0 & \alpha^* \end{bmatrix} \tag{A25.5}$$

or, in the quaternion notation,

$$Sx = x\alpha, \qquad \alpha = \alpha_0 + \alpha_1 e_1 . \tag{A25.6}$$

We see that the eigenvectors and eigenvalues of S are real quaternions. Moreover, the q-eigenvalue does not contain the e_2 and e_3 parts that give rise to off-diagonal terms in its 2×2 matrix representation. Thus the quaternion α in (A25.6) may be identified with the complex number α in (A25.3).

We say that two quaternions λ_1 and λ_2 are essentially distinct if the equation $\lambda_1 \mu = \mu \lambda_2$ implies $\mu = 0$. If $x_1 , x_2 ,..., x_r$ are q-eigenvectors belonging to the essentially distinct q-eigenvalues $\lambda_1 , \lambda_2 ,..., \lambda_r$, then $x_1 , x_2 ,..., x_r$ are right linearly independent; that is, the right linear (vector) equation

$$x_1 c_1 + x_2 c_2 + \cdots + x_r c_r = 0 \tag{A25.7}$$

implies

$$c_1 = c_2 = \cdots = c_r = 0.$$

The proof is by induction. Let the proposition be true for $(r - 1)$ vectors and let, if possible, all the c_i in (A25.7) be different from zero. Because λ are essentially distinct, not all of them are zero. Let $\lambda_1 \neq 0$. Multiplying (A25.7) from the left by S and using

$$Sx_i = x_i \lambda_i , \tag{A25.8}$$

we get

$$x_1 \lambda_1 c_1 + x_2 \lambda_2 c_2 + \cdots + x_r \lambda_r c_r = 0. \tag{A25.9}$$

Because c_1 and $\lambda_1 c_1$ are different from zero, they do have inverses. We multiply (A25.7) by c_1^{-1} and (A25.9) by $(\lambda_1 c_1)^{-1} = c_1^{-1} \lambda_1^{-1}$ from the right and subtract to get

$$\sum_{j=2}^{r} x_r (c_j c_1^{-1} - \lambda_j c_j c_1^{-1} \lambda_1^{-1}) = 0. \tag{A25.10}$$

Equation A25.10 is a right linear relation among the $(r-1)$ vectors. Therefore, by the induction hypothesis all the coefficients in (A25.10) must be zero; that is,

$$c_j c_1^{-1}\lambda_1 = \lambda_j c_j c_1^{-1}, \qquad c_j c_1^{-1} \neq 0, \qquad j = 2, 3, \ldots, r. \tag{A25.11}$$

However, this is contradictory to the hypothesis that λ_1 is essentially distinct from $\lambda_2, \lambda_3, \ldots, \lambda_r$. Because the proposition is evidently true for one single vector, the induction has a base and the demonstration is complete.

However, the right linear independence of a set of q-vectors does not necessarily lead to their left linear independence, as may be seen from the following example. The vectors

$$x_1 = \begin{bmatrix} 1 \\ e_1 \end{bmatrix}, \qquad x_2 = \begin{bmatrix} e_2 \\ -e_3 \end{bmatrix} \tag{A25.12}$$

are right linearly independent, but they are left linearly dependent. Thus we are still far from the diagonalization of S by purely quaternion means.

However, we may again use the intermediatory of the c-matrix S. If all the c-eigenvalues $z_1, z_1^*, \ldots, z_N, z_N^*$ of S are distinct, none of them being real, the c-matrix x whose columns are the eigenvectors of S belonging to these c-eigenvalues, is nonsingular and

$$x^{-1}Sx = E, \tag{A25.13}$$

where E is diagonal with diagonal elements $z_1, z_1^*, \ldots, z_N, z_N^*$. It is easy to be convinced that when x^{-1} is re-expressed as an $N \times N$ quaternion matrix all its elements will be real quaternions. Therefore, if all the N q-eigenvalues of S are essentially distinct, a quaternion real matrix x exists such that

$$S = xEx^{-1}, \tag{A25.14}$$

where E is diagonal and q-real.

A.26. Evaluation of the Integral (12.46)

As in Appendix A.24, we decompose the $2N \times 2N$ matrix X into the unique product $X = UYV$, where U is unitary, Y is triangular

with unit diagonal elements $Y_{ij} = 0, i > j$; $Y_{ii} = 1$, and V is diagonal with real positive elements. Moreover, because X has the form

$$\begin{bmatrix} a_{ij} & -b_{ij}^* \\ b_{ij} & a_{ij}^* \end{bmatrix}, \tag{A26.1}$$

the U, Y, and V all have the same form. In particular, $Y_{2i-1,2i} = -Y_{2i,2i-1}^* = 0$, and $V_{2i-1,2i-1} = V_{2i,2i}$.

Any matrix A having the form (A26.1) is equivalent to the statement that it satisfies the relation $ZA = A^*Z$, where Z is given by (2.21). From $ZX = X^*Z$ one sees that

$$U^TZU = Y^*VZV^{-1}Y^{-1} \tag{A26.2}$$

is unitary and antisymmetric (the left-hand side) and has nonzero elements only in the 2×2 blocks along the principal diagonal (comparison of elements on the two sides). One also has V_{ii} real and positive. Thus $U^TZU = Z$ and $V_{2i-1,2i-1} = V_{2i,2i}$. Substituting these in (A26.2), we finally get $Y^{*-1}ZY = VZV^{-1} = Z$. Thus U, Y, and V all have the form of (A26.1).

If we let $H = Y^\dagger Y$, then, because $H^\dagger = H$ and H has the form (A26.1), $H_{2i-1,2i-1} = H_{2i,2i} = h_i$ and $H_{2i-1,2i} = H_{2i,2i-1} = 0$. Moreover, h_i is completely determined by the condition $\det H_{2i-1} = \det Y_{2i-1}^\dagger Y_{2i-1} = 1$. Thus we may consider H_{ij}, $i < j$ and $(i, j) \neq (2k - 1, 2k)$ as independent complex variables. As in Appendix A.24, we change from the volume element

$$\prod_{i<j} \prod_{\lambda=0}^{1} dA_{2i-1,2j}^{(\lambda)}\, dA_{2i,2j}^{(\lambda)} \qquad \text{to} \qquad \prod dY_{ij}^{(0)}\, dY_{ij}^{(1)}$$

and finally to $\prod dH_{ij}^{(0)}\, dH_{ij}^{(1)}$, where the product $\prod$ over the elements of dY or dH are taken over all $i < j$ except the pairs $(i, j) = (2k - 1, 2k)$.

To calculate the integral

$$\int \exp[-(1/2)\,\mathrm{tr}(E^\dagger HEH^{-1})] \prod dH_{ij}^{(0)}\, dH_{ij}^{(1)}, \tag{A26.3}$$

a recurrence relation similar to (A24.17) is needed. Let H', E' denote matrices of order $2n$; H'', E'', their upper left diagonal blocks of order $(2n - 1)$, and H, E, their upper left diagonal blocks of order $2n - 2$.

The cofactors of H', H'', and H are denoted, respectively, by Δ', Δ'', and Δ. Let

$$g = (g_i), \qquad g' = (g'_i), \tag{A26.4}$$

where

$$g_i = H'_{i,2n-1}, \qquad g'_i = H'_{i,2n}, \qquad i = 1, 2, \ldots, 2n-2,$$
$$g_{2n-1} = g_{2n} = g'_{2n-1} = g'_{2n} = 0. \tag{A26.5}$$

Then

$$g'_{2i} = g^*_{2i-1}, \qquad g'_{2i-1} = -g^*_{2i},$$

or

$$g' = Zg^*, \tag{A26.6}$$

where Z is given by (2.21). Using the facts

$$H'_{2n-1,2n-1} = H'_{2n,2n} = h_n, \qquad H'_{2n-1,2n} = H'_{2n,2n-1} = 0, \tag{A26.7}$$

and

$$\det H' = \det H'' = \det H = 1, \tag{A26.8}$$

we get by expanding according to the last row and last column

$$1 = h_n - \sum_{i,j} g^*_i g_j \Delta_{ji}, \tag{A26.9}$$

$$\Delta''_{2n-1,l} = \sum_k g_k \Delta_{kl}, \tag{A26.10}$$

$$\Delta''_{ij} = h_n \Delta_{ij} - \sum_{k,l} g^*_k g_l \Delta^{lk}_{ij}$$
$$= \Delta_{ij} + \sum_{k,l} g^*_k g_l \Delta_{ik} \Delta_{lj}, \tag{A26.11}$$

where in the last step of (A26.11) we have used (A26.9) and (A24.15). The matrices H', H'', and H are positive definite and so are their inverses. In particular,

$$\sum_{i,j} f^*_i \Delta_{ij} f_j > 0 \tag{A26.12}$$

if not all f_i are zero. By equating to zero the expansion of a determinant whose first $2n-1$ rows are identical with those of H' and whose last row is identical with the last but one of H', we get

$$0 = \sum_i g_i^* \Delta'_{2n,i} + h_n \Delta'_{2n,2n-1}$$

$$= \sum_{i,j} g_i^* g_j' \Delta''_{ji} + h_n \sum_j g_j' \Delta''_{j,2n-1}, \tag{A26.13}$$

which on making use of (A26.9), (A26.10), and (A26.11) gives

$$0 = \sum_{j,p} g_j' g_p^* \Delta_{jp} \Big(1 + \sum_{i,q} g_i^* \Delta_{qi} g_q\Big). \tag{A26.14}$$

In view of (A26.12), this is equivalent to

$$0 = \sum_{j,p} g_j' \Delta_{jp} g_p^* = \Big(\sum_{j,p} g_j'^* \Delta_{pj} g_p\Big)^* \tag{A26.15}$$

or

$$\sum_j \Delta''_{j,2n-1} g_j' = (H''^{-1} g')_{2n-1} = 0. \tag{A26.15'}$$

Next we put

$$\phi_n = \tfrac{1}{2}\mathrm{tr}(E'^{\dagger} H' E' H'^{-1}) \tag{A26.16}$$

and apply (A24.17) twice to get

$$\phi_n = |z_n|^2 + \phi_{n-1} + \tfrac{1}{2}\langle g^* | U | g\rangle + \tfrac{1}{2}\langle g'^* | V | g'\rangle. \tag{A26.17}$$

The notation is that of (A24.18)

$$\langle f^* | B | f\rangle = \sum_{i,j} f_i^* B_{ij} f_j, \tag{A24.18}$$

where

$$U = H^{-1}(E^{\dagger} - z_n^*) H (E - z_n) H^{-1}, \tag{A26.18}$$

$$V = H''^{-1}(E''^{\dagger} - z_n) H'' (E'' - z_n^*) H''^{-1}. \tag{A26.19}$$

From (A26.15) we see that V is essentially equal to U.

$$\langle g'^* | V | g'\rangle = \langle g'^* | U | g'\rangle. \tag{A26.20}$$

Last, from (A26.6), $U^{\dagger} = U$, and $UZ = ZU^*$, Z given by (2.21), we have

$$\langle g'^* | U | g' \rangle = \langle g^* | U | g \rangle. \tag{A26.21}$$

Collecting (A26.17), (A26.20) and (A26.21), the recurrence relation becomes

$$\phi_n = | z_n |^2 + \phi_{n-1} + \langle g^* | U | g \rangle \tag{A26.22}$$

where U is given by (A26.18).

The rest of the integration is identical to that in Appendix A.24.

A.27. The Proof of Equation (12.70)

Let us put

$$\psi(x, y) = (2\pi)^{-1/2} \sum_{k=0}^{\infty} I_{k+1/2}(x)[y^{k+1/2} - y^{-(k+1/2)}]. \tag{A27.1}$$

Differentiating (A27.1) with respect to x and using the relation

$$I'_{\nu}(x) = \tfrac{1}{2}[I_{\nu+1}(x) + I_{\nu-1}(x)], \tag{A27.2}$$

we have

$$\begin{aligned}
\frac{\partial \psi}{\partial x} &= (2\pi)^{-1/2} \frac{1}{2} \sum_{k=0}^{\infty} [I_{k+3/2}(x) + I_{k-1/2}(x)][y^{k+1/2} - y^{-(k+1/2)}] \\
&= (2\pi)^{-1/2} \frac{1}{2} \sum_{k=0}^{\infty} I_{k+1/2}(x)\,[y^{k+1/2} - y^{-(k+1/2)}](y + y^{-1}) \\
&\quad + (2\pi)^{-1/2} \tfrac{1}{2}(y^{1/2} - y^{-1/2})(I_{1/2}(x) + I_{-1/2}(x)) \\
&= \tfrac{1}{2}(y + y^{-1})\,\psi(x, y) + (2\pi)^{-1/2}\,\tfrac{1}{2}(y^{1/2} - y^{-1/2}) \left(\frac{2}{\pi x}\right)^{1/2} e^x,
\end{aligned} \tag{A27.3}$$

where we have used the fact that

$$I_{1/2}(x) + I_{-1/2}(x) = \left(\frac{2}{\pi x}\right)^{1/2} (\sinh x + \cosh x). \tag{A27.4}$$

The differential equation (A27.3) can be immediately solved to give us

$$\psi(x, y) = \frac{1}{\pi} \exp[\tfrac{1}{2}(y + y^{-1})x](y^{1/2} - y^{-1/2}) \int_0^x \exp[x' - \tfrac{1}{2}(y + y^{-1})\, x'] \frac{dx'}{\sqrt{x'}}. \tag{A27.5}$$

Changing the integration variable from x' to $t = 1 - x'/x$, we get

$$\psi(x, y) = \frac{1}{\pi} e^x (y^{1/2} - y^{-1/2}) \sqrt{x} \int_0^1 \exp[\tfrac{1}{2}(y^{1/2} - y^{-1/2})^2\, xt] \frac{dt}{\sqrt{1 - t}} \tag{A27.6}$$

Finally putting $x = zz^*$ and $y = z^*/z$ we have

$$\psi(x, y) = \phi(z, z^*), \tag{A27.7}$$

and (A27.6) gives (12.70).

A.28. The Case of Random Real Matrices

For $N \times N$ real matrices the linear measure is

$$\mu(dS) = \prod_{i,j=1}^{N} dS_{ij}\,.$$

The procedure leading to (12.4) is the same as in the complex case, and instead of (12.8) we have

$$\mu(dS) = \prod_{i \neq j} |\, z_i - z_j \,|\, dA_{ij}\,, \qquad dA = X^{-1}\, dX,$$

where X diagonalizes the matrix S and $z_1\,, z_2\,, \ldots, z_N$ are the distinct eigenvalues of S. We then have to evaluate

$$\int e^{-\mathrm{tr}(S^\dagger S)} \prod_{i \neq j} dA_{ij}\,.$$

In case all the eigenvalues of a real matrix S are real, the corresponding eigenvectors can all be taken to be real and the real X that satisfies $S = XEX^{-1}$ can be written uniquely in the form $X = UYV$ (as in the Appendix A.24), where U, Y, and V are real, U is unitary, hence orthogonal, Y is triangular with unit diagonal elements, and V is diagonal positive. As in Appendix A.24, we can derive the recurrence relation

$$\phi_n = z_n^{\,2} + \phi_{n-1} + \langle g|\, H^{-1}(E - z_n)\, H(E - z_n)\, H^{-1}\, |g\rangle,$$

where $\phi_n = \text{tr}(E'H'E'H'^{-1})$, E' is an $n \times n$ real diagonal matrix with diagonal elements $z_1, z_2, \ldots, z_n$, $H' = Y_n^{\dagger} Y_n$, and Y_n is an $n \times n$ real triangular matrix with unit diagonal elements. The integral

$$\int \exp[-\text{tr}(EHEH^{-1})] \prod_{i<j} dH_{ij}$$

is immediate and comes out to be proportional to

$$\prod_{i<j} [(z_i - z_j)^2]^{-1/2}.$$

The joint probability density of the eigenvalues is therefore

$$P(z_1, \ldots, z_N) = C \exp\left(-\sum_1^N z_i^2\right) \prod_{i<j} | z_i - z_j | \qquad \text{(A28.1)}$$

in the case when all the eigenvalues are real. Equation A28.1 has the same form as (3.18).

In case some of the eigenvalues are complex, this procedure can still be carried out for the real eigenvalues. One is then left with the evaluation of an integral

$$\int e^{-\text{tr}(S^{\dagger}S)} \, d\mu(H)$$

where S is a real matrix, none of whose eigenvalues is real. For details the reader may refer to the original paper of Ginibre [1].

A.29. The Density of Eigenvalues of a Random Matrix Whose Elements All Have the Same Mean Square Value

Consider a matrix H with elements H_{ij} all having an average value zero and a mean square value V^2. Let the order N be large enough so that the density of its eigenvalues may be taken to be a continuous function. Let this function be $\sigma(\epsilon, V^2)$, so that the number of eigenvalues lying between ϵ and $\epsilon + d\epsilon$ is given by $\sigma(\epsilon, V^2)\, d\epsilon$. If we change the matrix elements by small quantities δH_{ij} such that the δH_{ij} themselves all have the average value zero and a mean square value v^2, the change in a particular eigenvalue at ϵ_i can be calculated by the second order perturbation theory

$$Z(\epsilon_i, V^2) = \delta H_{ii} + \sum_{j \neq i} \frac{| \delta H_{ij} |^2}{\epsilon_i - \epsilon_j} + \cdots. \qquad \text{(A29.1)}$$

The δH_{ii} do not produce, on the average, any change in ϵ_i. The eigenvalues ϵ_j which lie nearest to ϵ_i give the largest contribution to (A.29.1) with an absolute value $v^2\,\bar{s}$ where $\bar{s}$ is the mean spacing at ϵ_i. But as there are eigenvalues on both sides of ϵ_i the two contributions arising from the two nearest eigenvalues nearly cancel out, leaving quantities of a higher order in v^2. The sum in (A.29.1) can therefore be approximated by

$$Z(\epsilon, V^2) \approx v^2 \int \frac{\sigma(\epsilon', V^2)}{\epsilon - \epsilon'} d\epsilon' \tag{A29.2}$$

where the integral in (A.29.2) is a principal value integral and

$$V^2 = \langle | H_{ij} |^2 \rangle, \qquad v^2 = \langle | \delta H_{ij} |^2 \rangle \tag{A29.3}$$

the ensemble averages being indicated by $\langle\ \rangle$. Let us calculate the change in the number of eigenvalues lying in an interval $(\epsilon, \epsilon + \delta\epsilon)$. This can be done in two ways; one gives, as is obvious from the way of writing,

$$\sigma(\epsilon, V^2)\, Z(\epsilon, V^2) - \sigma(\epsilon + \delta\epsilon, V^2)\, Z(\epsilon + \delta\epsilon, V^2) \approx -\frac{\partial(\sigma Z)}{\partial \epsilon}\, \delta\epsilon$$

while the other gives in a similar way

$$v^2 \frac{\partial \sigma}{\partial V^2}. \tag{A29.4}$$

If all the matrix elements H_{ij} are multiplied by a constant c, the values ϵ_i are also multiplied by c, while V^2 is multiplied by c^2. Hence,

$$\sigma(c\epsilon, c^2 V^2)\, c\, d\epsilon = \sigma(\epsilon, V^2)\, d\epsilon. \tag{A29.5}$$

Setting $cV = 1$ the last equation gives

$$\sigma(\epsilon, V^2) = \frac{1}{V}\, \sigma\left(\frac{\epsilon}{V}, 1\right)$$

which could have been inferred by dimensional arguments. Putting

$$Z(\epsilon, V^2) = \frac{v^2}{V} Z_1\left(\frac{\epsilon}{V}\right), \qquad \sigma(\epsilon, V^2) = \frac{1}{V}\, \sigma_1\left(\frac{\epsilon}{V}\right) \tag{A29.6}$$

in (A.29.2) and (A.29.4), we obtain

$$\frac{\partial(Z_1\sigma_1)}{\partial x} = \frac{1}{2}\frac{\partial(x\sigma_1)}{\partial x}, \qquad x = \frac{\epsilon}{V} \tag{A29.7}$$

$$Z_1(x) = P\int \frac{\sigma_1(x')}{x - x'}\,dx'. \tag{A29.8}$$

When $x = 0$, by symmetry requirement $Z_1 = 0$; therefore (A29.7) gives, on integration,

$$Z_1(x) = \tfrac{1}{2}x. \tag{A29.9}$$

Finally we have the boundary condition

$$\int \sigma(\epsilon, V^2)\,d\epsilon = \int \sigma_1(x)\,dx = N. \tag{A29.10}$$

Equations (A.29.8), (A.29.9) and (A.29.10) together are equivalent to the integral equation (4.25) together with (4.21). The solution, as in Chapter 4, is the semicircle law (4.30):

$$\sigma(\epsilon, V^2) = \begin{cases} \dfrac{1}{2\pi V^2}(2NV^2 - \epsilon^2)^{1/2}, & \epsilon^2 < 2NV^2, \\ 0, & \epsilon^2 > 2NV^2. \end{cases}$$

A.30. Values of the Functions $\mathscr{B}(x_1, x_2)$ (Table A.30.1), and $\mathscr{P}(x_1, x_2)$ (Table A.30.2)

The functions $\mathscr{B}(x_1, x_2)$ and $\mathscr{P}(x_1, x_2)$ are symmetric in the variables x_1 and x_2. Their values can thus be read from the following tables for all positive values of x_1 and x_2 such that $\pi/2(x_1 + x_2) \leqslant 5$.

TABLE A.30.1

	$\pi/2\, x_1$	0.	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\pi/2\, x_2$	x_1 / x_2	0.	0.064	0.127	0.191	0.255	0.318	0.382	0.446	0.509	0.573
0.	0.	1.	0.996673	0.986770	0.970517	0.948273					
0.1	0.064	0.996673	0.993347	0.983448	0.967204	0.944975					
0.2	0.127	0.986770	0.983448	0.973565	0.957354	0.935182					
0.3	0.191	0.970517	0.967204	0.957354	0.941211	0.919149					
0.4	0.255	0.948273	0.944975	0.935182	0.919149	0.897263					
0.5	0.318	0.920509	0 917236	0 907528	0.891656	0.870019	0.843124	0.811560	0.775987	0.737102	0.695624
0.6	0.382	0.887796	0.884559	0.874967	0.859312	0.838005	0.811560	0.780578	0.745717	0.707674	0.667162
0.7	0.446	0.850776	0.847584	0.838145	0.822766	0.801871	0.775987	0.745717	0.711721	0.674690	0.635330
0.8	0.509	0.810140	0.807006	0.797756	0.782713	0.762317	0.737102	0.707674	0.674690	0.638836	0.600805
0.9	0.573	0.766610	0.763547	0.754522	0.739878	0.720065	0.695624	0.667162	0.635330	0.600805	0.564265
1.0	0.637	0.720917	0.717936	0.709174	0.694987	0.675839	0.652271	0.624889	0.594338	0.561279	0.526374
1.2	0.764	0.625886	0.623106	0.614971	0.601863	0.584259	0.562699	0.537776	0.510106	0.480318	0.449025
1.4	0.891	0.530374	0.527836	0.520442	0.508593	0.492763	0.473479	0.451307	0.426825	0.400610	0.373222
1.6	1.019	0.438855	0.436588	0.430019	0.419547	0.405634	0.388783	0.369515	0.348361	0.325840	0.302445
1.8	1.146	0.354707	0.352729	0.347024	0.337982	0.326038	0.311654	0.295305	0.277461	0.258575	0.239074
2.0	1.273	0.280137	0.278449	0.273610	0.265984	0.255970	0.243983	0.230439	0.215745	0.200289	0.184426
2.2	1.401	0.216245	0.214840	0.210830	0.204550	0.196351	0.186597	0.175643	0.163833	0.151486	0.138895
2.4	1.528	0.163196	0.162053	0.158809	0.153758	0.147204	0.139454	0.130805	0.121539	0.111912	0.102157
2.6	1.655	0.120437	0.119529	0.116966	0.112999	0.107883	0.101871	0.095204	0.088105	0.080778	0.073401
2.8	1.783	0.086932	0.086228	0.084251	0.081208	0.077309	0.072755	0.067736	0.062426	0.056981	0.051535
3.0	1.910	0.061384	0.060851	0.059361	0.057082	0.054179	0.050809	0.047120	0.043242	0.039291	0.035365
3.2	2.037	0.042408	0.042014	0.040917	0.039250	0.037139	0.034704	0.032055	0.029288	0.026488	0.023724
3.4	2.165	0.028670	0.028385	0.027596	0.026405	0.024905	0.023186	0.021328	0.019400	0.017462	0.015561
3.6	2.292	0.018969	0.018768	0.018214	0.017381	0.016341	0.015155	0.013882	0.012570	0.011259	0.009982
3.8	2.419	0.012285	0.012145	0.011765	0.011197	0.010491	0.009693	0.008840	0.007967	0.007101	0.006263
4.0	2.546	0.007787	0.007694	0.007439	0.007060	0.006592	0.006066	0.005508	0.004940	0.004381	0.003843
4.2	2.674	0.004833	0.004771	0.004604	0.004357	0.004054	0.003715	0.003358	0.002998	0.002645	
4.4	2.801	0.002937	0.002897	0.002789	0.002632	0.002440	0.002227	0.002004			
4.6	2.928	0.001747	0.001722	0.001655	0.001557	0.001438					
4.8	3.056	0.0010177	0.0010023	0.0009611							
5.0	3.183	0.0005805									

TABLE A.30.1 *(continued)*

	$\pi/2\, x_1$	1.0	1.2	1.4	1.6	1.8	2.0	2.2	2.4
$\pi/2\, x_2$	x_2 \ x_1	0.637	0.764	0.891	1.019	1.146	1.273	1.401	1.528
1.0	0.637	0.490265	0.416818	0.345188	0.278636	0.219347			
1.2	0.764	0.416818	0.351814	0.289074	0.231395	0.180558			
1.4	0.891	0.345188	0.289074	0.235515	0.186831	0.144411			
1.6	1.019	0.278636	0.231395	0.186831	0.146802	0.112341			
1.8	1.146	0.219347	0.180558	0.144411	0.112341	0.085076			
2.0	1.273	0.168479	0.137413	0.108825	0.083782	0.062763	0.045785	0.032545	0.022555
2.2	1.401	0.126316	0.102044	0.079993	0.060927	0.045133	0.032545	0.022860	0.015650
2.4	1.528	0.092475	0.073971	0.057380	0.043223	0.031652	0.022555	0.015650	0.010581
2.6	1.655	0.066127	0.052361	0.040180	0.029926	0.021658	0.015247	0.010448	0.006975
2.8	1.783	0.046200	0.036203	0.027475	0.020228	0.014465	0.010057	0.006805	
3.0	1.910	0.031544	0.024456	0.018352	0.013353	0.009432	0.006476		
3.2	2.037	0.021052	0.016145	0.011977	0.008610	0.006007			
3.4	2.165	0.013736	0.010419	0.007639	0.005425				
3.6	2.292	0.008764	0.006573	0.004762					
3.8	2.419	0.005469	0.004055						
4.0	2.546	0.003338							

	$\pi/2\ x_1$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\pi/2\ x_2$	x_1 / x_2	0.064	0.127	0.191	0.255	0.318	0.382	0.446	0.509	0.573
0.1	0.064	0.00	0.004	0.006	0.10					
0.2	0.127	0.004	0.00849	0.01595	0.02474					
0.3	0.191	0.006	0.01595	0.02793	0.04207					
0.4	0.255	0.010	0.02474	0.04207	0.06249					
0.5	0.318	0.016	0.03484	0.05842	0.08446	0.11249	0.14205	0.17185	0.20074	0.22829
0.6	0.382	0.020	0.04640	0.07560	0.10768	0.14205	0.17725	0.21195	0.24554	0.27713
0.7	0.446	0.027	0.05811	0.09341	0.13184	0.17185	0.21195	0.25135	0.28898	0.32374
0.8	0.509	0.033	0.07016	0.11175	0.15582	0.20074	0.24554	0.28898	0.32978	0.36680
0.9	0.573	0.039	0.08241	0.12966	0.17867	0.22829	0.27713	0.32374	0.36680	0.40534
1.0	0.637	0.044	0.09421	0.14642	0.20014	0.25380	0.30580	0.35468	0.39926	0.43856
1.2	0.764	0.055	0.11484	0.17578	0.23672	0.29583	0.35163	0.40277	0.44805	0.48647
1.4	0.891	0.064	0.13065	0.19708	0.26182	0.32317	0.37963	0.42983	0.47296	0.50824
1.6	1.019	0.070	0.14003	0.20867	0.27400	0.33434	0.38850	0.43541	0.47426	0.50457
1.8	1.146	0.072	0.14273	0.21029	0.27317	0.32998	0.37961	0.42124	0.45444	0.47899
2.0	1.273	0.071	0.13906	0.20287	0.26094	0.31213	0.35569	0.39106	0.41805	0.43672
2.2	1.401	0.067	0.13015	0.18804	0.23962	0.28404	0.32077	0.34957	0.37045	0.38362
2.4	1.528	0.061	0.11730	0.16800	0.21219	0.24936	0.27919	0.30164	0.31693	0.32547
2.6	1.655	0.054	0.10210	0.14500	0.18161	0.21162	0.23495	0.25177	0.26237	0.26721
2.8	1.783	0.046	0.08600	0.12113	0.15046	0.17392	0.19156	0.20360	0.21045	0.21261
3.0	1.910	0.038	0.07017	0.09808	0.12089	0.13863	0.15146	0.15972	0.16379	0.16417
3.2	2.037	0.030	0.05557	0.07708	0.09426	0.10727	0.11629	0.12169	0.12382	0.12314
3.4	2.165	0.023	0.04274	0.05884	0.07142	0.08067	0.08680	0.09013	0.09100	0.08979
3.6	2.292	0.018	0.03195	0.04369	0.05263	0.05901	0.06301	0.06493	0.06505	0.06371
3.8	2.419	0.013	0.02325	0.03156	0.03775	0.04200	0.04451	0.04554	0.04532	0.04392
4.0	2.546	0.009	0.01647	0.02219	0.02635	0.02911	0.03068	0.03098	0.03054	0.03051
4.2	2.674	0.006	0.01135	0.01520	0.01799	0.01951	0.02032	0.02182	0.01811	
4.4	2.801	0.004	0.00771	0.00998	0.01160	0.01406	0.01099			
4.6	2.928	0.002	0.00499	0.00738	0.00622					
4.8	3.056	0.00	0.00271							
5.0	3.183									

TABLE A.30.2 *(continued)*

	$\pi/2\, x_1$	1.0	1.2	1.4	1.6	1.8	2.0	2.2	2.4
$\pi/2\, x_2$	x_1 / x_2	0.637	0.764	0.891	1.019	1.146	1.273	1.401	1.528
1.0	0.637	0.47173	0.51757	0.53524	0.52626	0.49501			
1.2	0.764	0.51757	0.55629	0.56405	0.54434	0.50292			
1.4	0.891	0.53524	0.56405	0.56133	0.53206	0.48297			
1.6	1.019	0.52626	0.54434	0.53206	0.49548	0.44217			
1.8	1.146	0.49501	0.50292	0.48297	0.44217	0.38796			
2.0	1.273	0.44735	0.44659	0.42166	0.37955	0.32751	0.27196	0.21795	0.16889
2.2	1.401	0.38956	0.38240	0.35501	0.31429	0.26678	0.21795	0.17182	0.13098
2.4	1.528	0.32777	0.31640	0.28893	0.25164	0.21015	0.16889	0.13098	0.09848
2.6	1.655	0.26689	0.25345	0.22771	0.19514	0.16032	0.12676	0.09698	0.07311
2.8	1.783	0.21065	0.19685	0.17404	0.14674	0.11861	0.09252	0.07086	
3.0	1.910	0.16137	0.14842	0.12912	0.10713	0.08542	0.06647		
3.2	2.037	0.12009	0.10871	0.09309	0.07615	0.06023			
3.4	2.165	0.08688	0.07745	0.06534	0.05258				
3.6	2.292	0.06120	0.05366	0.04401					
3.8	2.419	0.04184	0.03506						
4.0	2.546	0.02627							

A.31. Proof of Equations (6.25″) and (16.69)

Consider the integral equation

$$\lambda g(x) = \int_{-t}^{t} K(x, y)\, g(y)\, dy \tag{A31.1}$$

whose solutions can always be chosen to be either even, $g(-x) = g(x)$, or odd, $g(-x) = -g(x)$. The set of even solutions, $g_{2n}(x)$, labeled by even subscripts, consists of all the solutions of an integral equation obtained from (A31.1) when $K(x, y)$ there is replaced by

$$K_{\text{even}}(x, y) = \tfrac{1}{2}\{K(x, y) + K(-x, y)\}. \tag{A31.2}$$

Similarly the set of odd solutions $g_{2n+1}(x)$ of (A31.1), labeled by odd subscripts, consists of all the solutions of an integral equation with the kernel

$$K_{\text{odd}}(x, y) = \tfrac{1}{2}\{K(x, y) - K(-x, y)\}. \tag{A31.3}$$

Let us denote the Fredholm determinants of $K(x, y)$, $K_{\text{even}}(x, y)$ and $K_{\text{odd}}(x, y)$, by $\Psi(t)$, $\Psi_{\text{even}}(t)$ and $\Psi_{\text{odd}}(t)$, respectively. We have

$$\Psi(t) = \det[1 - K] = \prod_{n=0}^{\infty} (1 - \lambda_n)$$

$$= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int \cdots \int dx_1 \cdots dx_n \det[K(x_i, x_j)]_{i,j=1,\ldots,n} \tag{A31.4}$$

and similar expressions for $\Psi_{\text{even}}(t)$ and $\Psi_{\text{odd}}(t)$. In (A31.4) and in what follows the integrals will be understood to be taken from $-t$ to t. We want to show that

$$(\alpha) \qquad \Psi_{\text{even}} \frac{d^2\Psi_{\text{odd}}}{dt^2} + \Psi_{\text{odd}} \frac{d^2\Psi_{\text{even}}}{dt^2} = 2 \frac{d\Psi_{\text{even}}}{dt} \frac{d\Psi_{\text{odd}}}{dt}$$

and

$$(\beta) \qquad \frac{\Psi_{\text{odd}}}{\Psi_{\text{even}}} = 1 + \sum_{n=0}^{\infty} \frac{\lambda_{2n}}{1 - \lambda_{2n}} g_{2n}(t) \int g_{2n}(x)\, dx$$

where the g_{2n} are normalized even solutions of (A31.1):

$$\int g_{2n}^2(x)\, dx = 1; \tag{A31.5}$$

provided the kernel $K(x, y)$ is an even function of the difference of its arguments:

$$K(x, y) = K(x - y) = K(y - x). \tag{A31.6}$$

Proof: (i) Let $Q(x, y)$ be any of the three kernels K, K_{even} or K_{odd} and $\Phi(t)$ be the corresponding Fredholm determinant

$$\Phi(t) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int \cdots \int dx_1 \cdots dx_n \det[Q(x_i, x_j)]_{i,j=1,\ldots,n}. \tag{A31.7}$$

Differentiating the above equation with respect to t we have

$$\Phi'(t) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} n \int \cdots \int dx_1 \cdots dx_{n-1} \{G_{n-1}(t) + G_{n-1}(-t)\}$$

$$= -2 \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int \cdots \int dx_1 \cdots dx_n \, G_n(t) \tag{A31.8}$$

where

$$G_n(t) = \det \begin{bmatrix} Q(t, t) & Q(t, x_j) \\ Q(x_i, t) & Q(x_i, x_j) \end{bmatrix}_{i,j=1,\ldots,n} \tag{A31.9}$$

and in arriving at $G_n(-t) = G_n(t)$ we have used the relations (A31.6). A comparison of (A31.8), (A31.9) with the expansions of the "resolvant" and the "minors" in the Fredholm theory of integral equations will show that

$$\frac{\Phi'(t)}{\Phi(t)} = -2 \left(\frac{Q}{1 - Q}\right)(t, t) \tag{A31.10}$$

with

$$\left(\frac{Q}{1 - Q}\right)(x, y) = \frac{\text{minor } (x, y) \text{ in } (1 - Q)}{\det(1 - Q)}$$

$$= \sum_{n=0}^{\infty} \int \cdots \int dx_1 \cdots dx_n \, Q(x, x_1) Q(x_1, x_2) \cdots Q(x_n, y). \tag{A31.11}$$

Applying this result to the even case we get an expression for $\Psi'_{\text{even}}/\Psi_{\text{even}}$ as

$$-2\sum_{n=0}^{\infty}\int\cdots\int dx_1\cdots dx_n\,K_{\text{even}}(t, x_1)\cdots K_{\text{even}}(x_n, t)$$

$$= -2\sum_{n=0}^{\infty}\int\cdots\int dx_1\cdots dx_n\,K(t, x_1)\cdots K(x_{n-1}, x_n)\,K_{\text{even}}(x_n, t) \tag{A31.12}$$

In a similar way

$$\frac{\Psi'_{\text{odd}}}{\Psi_{\text{odd}}} = -2\sum_{n=0}^{\infty}\int\cdots\int dx_1\cdots dx_n\,K(t, x_1)\cdots K(x_{n-1}, x_n)\,K_{\text{odd}}(x_n, t). \tag{A31.13}$$

Adding and subtracting the last two equations we get

$$a(t) + b(t) = -2\sum_{n=0}^{\infty}\int\cdots\int dx_1\cdots dx_n\,K(t, x_1)\cdots K(x_n, t)$$

$$= -2\left(\frac{K}{1-K}\right)(t, t), \tag{A31.14}$$

$$a(t) - b(t) = -2\sum_{n=0}^{\infty}\int\cdots\int dx_1\cdots dx_n\,K(t, x_1)\cdots K(x_n, -t)$$

$$= -2\left(\frac{K}{1-K}\right)(t, -t), \tag{A31.15}$$

where for convenience we have put

$$a(t) = \frac{\Psi'_{\text{even}}(t)}{\Psi_{\text{even}}(t)}, \qquad b(t) = \frac{\Psi'_{\text{odd}}(t)}{\Psi_{\text{odd}}(t)}. \tag{A31.16}$$

Differentiating (A31.14) once more we get

$$a' + b' = A_1 + A_2 + B_1 + B_2 \tag{A31.17}$$

where

$$A_1 = -2\sum_{n=1}^{\infty}\sum_{j=1}^{n}\int\cdots\int dx_1\cdots dx_{j-1}\,dx_{j+1}\cdots dx_n$$

$$\cdot\,K(t, x_1)\cdots K(x_{j-1}, t)\,K(t, x_{j+1})\cdots K(x_n, t), \tag{A31.18}$$

$$A_2 = -2 \sum_{n=1}^{\infty} \sum_{j=1}^{n} \int \cdots \int dx_1 \cdots dx_{j-1}\, dx_{j+1} \cdots dx_n$$
$$\cdot K(t, x_1) \cdots K(x_{j-1}, -t)\, K(-t, x_{j+1}) \cdots K(x_n, t), \tag{A31.19}$$

$$B_1 = -2 \sum_{n=0}^{\infty} \int \cdots \int dx_1 \cdots dx_n \frac{\partial K(t, x_1)}{\partial t} K(x_1, x_2) \cdots K(x_n, t), \tag{A31.20}$$

and

$$B_2 = -2 \sum_{n=0}^{\infty} \int \cdots \int dx_1 \cdots dx_n\, K(t, x_1) \cdots K(x_{n-1}, x_n) \frac{\partial K(x_n, t)}{\partial t}. \tag{A31.21}$$

Since $K(x, y)$ depends only on $x - y$, we can in B_1 replace $\partial K(t, x_1)/\partial t$ by $-\partial K(t, x_1)/\partial x_1$ and integrate partially with respect to x_1. Next in the integral

$$\int \cdots \int dx_1 \cdots dx_n\, K(t, x_1) \frac{\partial K(x_1, x_2)}{\partial x_1} K(x_2, x_3) \cdots K(x_n, t)$$

we replace $\partial K(x_1, x_2)/\partial x_1$ by $-\partial K(x_1, x_2)/\partial x_2$ and integrate partially with respect to x_2, and so on till the partial derivation is pushed to the extreme right. These step-by-step partial integrations give finally

$$B_1 = -A_1 + A_2 - B_2. \tag{A31.22}$$

Equations (A31.17), (A31.18), (A31.19), (A31.22), and (A31.15) give therefore

$$a' + b' = 2A_2 = -(a - b)^2 \tag{A31.23}$$

which in view of (A31.16) is the relation (α).

(ii) Provided the summation and integration can be interchanged, relation (β) can be written as

$$\frac{\Psi_{\text{odd}}}{\Psi_{\text{even}}} = 1 + \int \left(\frac{K}{1 - K}\right)(t, x)\, dx. \tag{A31.24}$$

One has only to expand the resolvant in terms of the normalized eigenfunctions

$$\left(\frac{K}{1 - K}\right)(t, x) = \sum_{n=0}^{\infty} \frac{\lambda_n}{1 - \lambda_n} g_n(t)\, g_n(x) \tag{A31.25}$$

and note that the odd functions contribute nothing on integration. We will prove relation (β) in the form (A31.24).

Let us calculate the logarithmic derivatives with respect to t of the two sides of (A31.24).

$$\left(\frac{\Psi_{\text{even}}}{\Psi_{\text{odd}}}\right)\frac{d}{dt}\left(\frac{\Psi_{\text{odd}}}{\Psi_{\text{even}}}\right) = b - a = 2\left(\frac{K}{1-K}\right)(t, -t) \qquad \text{(A31.26)}$$

from equations (A31.16) and (A31.15). Also the derivative of the right-hand side is

$$\begin{aligned}
&\frac{d}{dt}\left\{1 + \sum_{n=1}^{\infty}\int\cdots\int dx_1\cdots dx_n\, K(t, x_1)\cdots K(x_{n-1}, x_n)\right\} \\
&\quad = \sum_{n=1}^{\infty}\sum_{j=1}^{n}\int\cdots\int dx_1\cdots dx_{j-1}\, dx_{j+1}\cdots dx_n \\
&\qquad \cdot\{K(t, x_1)\cdots K(x_{j-1}, t)\, K(t, x_{j+1})\cdots K(x_{n-1}, x_n) \\
&\qquad\qquad + K(t, x_1)\cdots K(x_{j-1}, -t)\, K(-t, x_{j+1})\cdots K(x_{n-1}, x_n)\} \\
&\qquad + \sum_{n=1}^{\infty}\int\cdots\int dx_1\cdots dx_n\,\frac{\partial K(t, x_1)}{\partial t}\, K(x_1, x_2)\cdots K(x_{n-1}, x_n) \qquad \text{(A31.27)}
\end{aligned}$$

In the last line of the above equation one can again shift the partial derivation to the extreme right by successively replacing $\partial K(x, y)/\partial x$ by $-\partial K(x, y)/\partial y$ and integrating by parts. The expression in the last line of (A31.27) is therefore

$$\begin{aligned}
&\sum_{n=1}^{\infty}\sum_{j=1}^{n}\int\cdots\int dx_1\cdots dx_{j-1}\, dx_{j+1}\cdots dx_n \\
&\qquad \cdot\{-K(t, x_1)\cdots K(x_{j-1}, t)\, K(t, x_{j+1})\cdots K(x_{n-1}, x_n) \\
&\qquad\quad + K(t, x_1)\cdots K(x_{j-1}, -t)\, K(-t, x_{j+1})\cdots K(x_{n-1}, x_n)\}. \qquad \text{(A31.28)}
\end{aligned}$$

Also

$$\begin{aligned}
&\sum_{n=1}^{\infty}\sum_{j=1}^{n}\int\cdots\int dx_1\cdots dx_{j-1}\, dx_{j+1}\cdots dx_n \\
&\qquad \cdot K(t, x_1)\cdots K(x_{j-1}, -t)\, K(-t, x_{j+1})\cdots K(x_{n-1}, x_n) \\
&\quad = \left(\frac{K}{1-K}\right)(t, -t)\left\{1 + \int dx\left(\frac{K}{1-K}\right)(-t, x)\right\} \qquad \text{(A31.29)}
\end{aligned}$$

and

$$\int dx \left(\frac{K}{1-K}\right)(-t, x) = \int dx \left(\frac{K}{1-K}\right)(t, -x)$$

$$= \int dx \left(\frac{K}{1-K}\right)(t, x). \qquad \text{(A31.30)}$$

Putting together equations (A31.26)–(A31.30) we see that the logarithmic derivatives of the two sides of (A31.24) are equal. In addition (A31.24) is obviously valid for $t = 0$. Thus (A31.24) is valid for all t.

(iii) The formal manipulations encountered in this appendix are valid of course only when the various expansions are uniformly convergent so that term by term differentiation is allowed. This will be so if we take

$$K(x, y) = \frac{\sin(x-y)\pi}{(x-y)\pi}$$

so that relations (α) and (β) are the same as the equations (6.25″) and (16.69).

Lately the relations (α) and (β) were circulated as conjectures. This appendix is based on a letter of Gaudin which he wrote in response to those conjectures.

References

Aitken, A. C.
1. *Determinants and Matrices*, pp. 49–50. Wiley (Interscience), New York, 1954.

Baranger, M.
1. *Phys. Rev.* **120**, 957 (1950).

Bateman, H.
1. *Higher Transcendental Functions* (A. Erdelyi, ed.), Vol. 2, Chapters 10.10 and 10.13. McGraw-Hill, New York, 1953.
2. *Ibid.*, Vol. 2, formula 7.2(56).
3. *Ibid.*, Vol. 1, formulas 1.17(2), 1.18(1), 1.2(2), 1.9(10), and 1.7(11).
4. *Ibid.*, Vol. 1, formula 1.1(4).

Bethe, H. A.
1. *Rev. Mod. Phys.* **9**, 69 (1937).

Bohr, A.
1. *Proc. Intern. Conf. Peaceful Uses At. Energy, Geneva*, 1955, Vol. 2, p. 151. Columbia Univ. Press (I.D.S.), New York, 1956.

Bronk, B. V.
1. Unpublished thesis, Princeton University (1964).
2. *J. Math. Phys.* **5**, 215 (1964).
3. *J. Math. Phys.* **6**, 228 (1965).

Bruijn, N. G.
1. *J. Indian Math. Soc.* **19**, 133 (1955).

Burke, J. E.
1. *J. Math. and Phys.* **45**, 425 (1966).

Cameron, A. G. W.
1. *Can. J. Phys.* **36**, 1040 (1956).

Chandrasekhar, S.
1. *Rev. Mod. Phys.* **15**, 1 (1943). Reprinted in *Noise and Stochastic Processes*. Dover, New York, 1954.

Chevalley, C.
1. *Theory of Lie Groups*, pp. 16–24. Princeton Univ. Press, Princeton, New Jersey, 1946.

Courant, R. and D. Hilbert,
1. *Methoden der Mathematischen Physik*, Vol. 1, Chapter 1, Sect. 5. Springer, Berlin, 1924.

Dieudonné, J.
1. *La Géométrie des Groupes Classiques*, Ergeb. d. Math., Vol. 5. Springer, Berlin, 1955.

Dyson, F. J.
1. *J. Math. Phys.* **3**, 140 (1962).
2. *Ibid.*, p. 157.
3. *Ibid.*, p. 166.
4. *Ibid.*, **3**, 1191 (1962).
5. *Ibid.*, p. 1199.
6. *Phys. Rev.* **92**, 1331 (1953).

Dyson F. J. and M. L. Mehta
1. *J. Math. Phys.* **4**, 701 (1963).

Englman, R.
1. *Nuovo Cimento* **10**, 615 (1958).

Erdyeli, A.
1. *J. Indian Math. Soc.*, Golden Jubilee Commemoration Vol. **24**, 235 (1960).

Euler, L.
1. *Novi Commentarii Acad. Sci. Petropolit.* **3**, 125 (1753). Reprinted in *Leonhardi Euleri Opera Omnia*, Ser. I, Opera Mathematica, Vol. II, Commentationes Arithemeticae, Volumen primum, p. 254, Berlin, 1915.

Favro, L. D., P. B. Kahn, and M. L. Mehta
1. Brookhaven Natl. Lab. Rept. No. BNL-757 (T-280) (1962).

Fox, D., and P. B. Kahn
1. *Phys. Rev.* **134**, B1151 (1964).

Fröhlich, H.
1. *Physica* **4**, 406 (1937).

Garg, J. B., J. Rainwater, J. S. Petersen, and W. W. Havens, Jr.
1. *Phys. Rev.* **134**, B985 (1964) and the references cited therein. A detailed description of the recent experimental work on Th^{232} and U^{238} energy levels is discussed.

Gantmacher, F. R.
1. *Matrix Theory* (transl. from Russian), Vol. I, pp. 31–33. Chelsea, New York, 1959.
2. *Ibid.*, Vol. II, p. 18.

Gaudin, M.
1. *Nucl. Phys.* **25**, 447 (1961).
2. Private communication (1962).

Ginibre, J.
1. *J. Math. Phys.* **6**, 440 (1965).

Gorkov, L. P., and G. M. Eliashberg
1. *Zh. Eksptl. i Teoret. Fiz.* **48**, 1407 (1965); see *Soviet Phys. JETP* **21**, 940 (1965).

Goursat, E.
1. *Cours d'analyse mathématique*, Vol. III, pp. 389, 454. Gauthier-Villars, Paris, 1956.

Gunsen, J.
1. *J. Math. Phys.* **3**, 752 (1962).

Gurevich, I. I., and M. I. Pevsner
1. *Zh. Eksptl. i Teoret. Fiz.* **31**, 162 (1956); *Nucl. Phys.* **2**, 575 (1957).

Hardy, G. H.
1. *Ramanujan*, Chapter VIII. Cambridge Univ. Press, London and New York, 1940.

Hardy, G. H., and S. Ramanujan
1. *Proc. London Math. Soc.* **16**, 130 (1917).

Hsu, P. L.
1. *Ann. Eug.* **9**, 250 (1939).

Kahn, P. B.
1. *Nucl. Phys.* **41**, 159 (1963).

Kisslinger, L. S. and R. A. Sorenson
1. *Kgl. Danske Videnskab. Selskab. Mat.-fys. Medd.* **32**, No. 9 (1960).

Kramer, H. A.
1. *Proc. Acad. Sci. Amsterdam* **33**, 959 (1930).

Kubo, R.
1. *J. Phys. Soc. Japan* **17**, 975 (1962).

Landau, L. D., and Ya. Smorodinski
1. *Lectures on the Theory of the Atomic Nucleus* (English transl.), p. 55. Consultants Bureau, New York, 1958.

Lane, A. M.
1. Oak Ridge Natl. Lab. Rept. ORNL-2309, p. 113 (1957).

Lane, A. M., and R. G. Thomas
1. *Rev. Mod. Phys.* **30**, 257 (1958).

Lang, J., and K. Lecouteur
1. *Proc. Phys. Soc. London* **A67**, 586 (1954).

Leff, H. S.
1. Unpublished thesis, State University of Iowa, SUI-63-23 (1963).

Maradudin, A. A., P. Mazur, E. W. Montroll, and G. H. Weiss
1. *Rev. Mod. Phys.* **30**, 175 (1958), Sec. IV.

Mayer, M. G., and J. H. D. Jensen
1. *Elementary Theory of Nuclear Shell Structure*. Wiley, New York, 1955.

Mehta, C. L.
1. Private communication (1962).

Mehta, M. L.
1. *Nucl. Phys.* **18**, 395 (1960).

Mehta, M. L., and F. J. Dyson
1. *J. Math. Phys.* **4**, 713 (1963).

Mehta, M. L., and M. Gaudin
1. *Nucl. Phys.* **18**, 420 (1960).

Mehta, M. L., and P. K. Srivastava
1. *J. Math. Phys.* **7**, 341 (1966).

Mohr, E.
1. *Math. Nachr.* **10**, 257 (1953).

Moore, C. E.
1. *Natl. Bur. Std. (U.S.)*, Circ. 467, Vol. I (1949); Vol. II (1952); Vol. III (1958).

Mushkhelishvili, N. I.
1. *Singular Integral Equations*, Noordhoff, Groningen, The Netherlands, 1953.

Porter, C. E., and N. Rosenzweig
1. *Ann. Acad. Sci. Fennicae*: Ser. A VI, No. 44 (1960).

Robin, L.
1. *Fonctions sphériques de legendre et fonctions sphéroidales*. Vol. 3, p. 250, formula 255. Gauthier-Villars, Paris, 1959.

Rosen, J. L., J. S. Desjardins, J. Rainwater, and W. W. Havens, Jr.
1. *Phys. Rev.* **118**, 687 (1960).

Rosenzweig, N.
1. In *Statistical Physics*, Vol. 3. Benjamin, New York, 1963.

Rosenzweig, N., and C. E. Porter
1. *Phys. Rev.* **120**, 1698 (1960).

Scott, J. M. C.
1. *Phil. Mag.* **45**, 1322 (1954).

Shohat, J. A., and J. D. Tamarkin
1. *The Problem of Moments*, p. 8. Am. Math. Soc., Providence, Rhode Island, 1943.

Smythe, W. R.
1. *Static and Dynamic Electricity*, p. 104; Problems 29 and 30, McGraw-Hill, New York, 1950.

Stieltjes, T. J.
1. *Sur quelques théorèmes d'Algèbre, Œuvres complètes*, Vol. I, p. 440. Noordhoff, Groningen, The Netherlands, 1914.

Stratton, J. A., P. M. Morse, L. J. Chu, J. D. C. Little, and F. J. Corbato
1. *Spheroidal Wave Functions*, M.I.T. Press, Cambridge, Massachusetts, 1956.

Study, E.
1. *Acta Math.* **42**, 1 (1918).

Szegö, G.
1. *Orthogonal Polynomials*, Sect. 6.7, p. 139 and Sect. 6.71, p. 142. Am. Math. Soc. Publ., New York, 1959.

Tischmarsh, E. C.
1. *Theory of Functions*, p. 186. Oxford Univ. Press, London and New York, 1939.

Thomas, R. G., and C. E. Porter
1. *Phys. Rev.* **104**, 483 (1956).

Uhlenbeck, G. E., and L. S. Ornstein
1. *Phys. Rev.* **36**, 823 (1930). Reprinted in *Noise and Stochastic Processes*, Dover, New York, 1954.

Ullah, N.
1. *Nucl. Phys.* **58**, 65 (1964).

Ullah, N., and C. E. Porter
1. *Phys. Letters* **6**, 301 (1963).

Wang, M. C., and G. E. Uhlenbeck
1. *Rev. Mod. Phys.* **17**, 323 (1945). Reprinted in *Noise and Stochastic Processes.* Dover, New York, 1954.

Weyl, H.
1. *Classical Groups.* Princeton Univ. Press, Princeton, New Jersey, 1946.

Wigner, E. P.
1. *Group Theory*, Chapter 26. Academic, New York, 1959.
2. *Ibid.*, Chapter 3, pp. 26–28.
3. *Gatlinberg Conf. on Neutron Physics*, Oak Ridge Natl. Lab. Rept. No. ORNL-2309, p. 59.
4. *Proc. 4th Can. Math. Congr., Toronto*, 1959, p. 174.
5. *Ann. Math.* **62**, 548 (1955); **65**, 203 (1957).
6. Distribution laws for the roots of a random Hermitian matrix. Unpublished (1962).
7. Unpublished (perhaps 1959–1960).

Wilson, K. G.
1. *J. Math. Phys.* **3**, 1040 (1962).

Wishart, J.
1. *Biometrika* **20**, 32 (1928).

Index