J. Phys. A: Math. Gen. 18 (1985) 2087-2094. Printed in Great Britain

Two-dimensional neutral Coulomb gas

Prabodh Shukla

Physics Department, North Eastern Hill University, Shillong 793003, India

Received 31 July 1984

Abstract. The problem of the two-dimensional neutral Coulomb gas is analysed in two separate approximations: the one-atom approximation at low temperatures and the Thomas-Fermi approximation at high temperatures. A possible connection between the Kosterlitz-Thouless transition and the thermal ionisation of a single atom is discussed. An expression for the position dependence of the dielectric constant of the gas is obtained in the Thomas-Fermi approximation.

1. Introduction

The two-dimensional neutral Coulomb gas has been studied by several authors [1-4]. The main point of interest is the unusual nature of the phase transition in this system. Usually, the correlations between fluctuations at two different points in a system decay rapidly exponentially with the spatial distance between the points, except at the critical point of the system where the correlations decay slowly algebraically. In the two-dimensional Coulomb gas the correlations decay slowly algebraically over an entire range of temperatures from $T=0$ to $T=T_c$. In this sense, the whole range of temperatures, from $T=0$ to $T=T_c$, may be thought of as a line of critical points of the two-dimensional Coulomb gas. The critical exponents, however, depend on the temperature and are therefore not universal. At the end of the line of critical points, a phase transition takes place into a new phase where the correlations decay rapidly exponentially with distance. Similar behaviour also arises in many other cases, such as the two-dimensional XY model of ferromagnetism, the roughening model and the one-dimensional Fermi gas. All these models can be related to the two-dimensional Coulomb gas [5].

The existing work on the two-dimensional Coulomb gas and other related systems is largely in the classical approximation. We shall show in this paper that some quantum effects can be incorporated in the theory without increasing the mathematical complexity. The quantum effects do not alter the essential physics of the problem but give us a different perspective of the problem. In the limit of low density and low temperature, the neutral Coulomb gas can be approximated by a gas of hydrogen atoms (i.e. bound pairs of equal and opposite charges) with negligible interaction between the atoms. In this limit the properties of the gas can be related simply to the properties of an individual atom. We shall see that there is a basic difference between the three-dimensional hydrogen atom and the two-dimensional hydrogen atom, which lies at the heart of the distinct thermal behaviour of the two- and the three-dimensional Coulomb gases. A three-dimensional hydrogen atom, in thermal equilibrium in an

0305-4470/85/112087+08$02.25

infinite volume, is completely ionised (i.e. ionised with probability 1) at any finite temperature [6]. A two-dimensional hydrogen atom, as we shall see below, in similar circumstances is completely ionised only above a critical temperature T_c. There are several similarities between the Kosterlitz-Thouless transition [2] in the two-dimensional Coulomb gas and the ionisation of the two-dimensional hydrogen atom. Though there are no true scattering states for a pair of opposite charges in two dimensions, above the thermal ionisation temperature the charges may be considered effectively free because the probability for the pair to exist in any definite state is zero. It is also clear that above the thermal ionisation temperature the many-body effects will become so important that the properties of even a dilute gas may not be simply related to just a pair of charges. In this temperature regime we shall show that the Thomas-Fermi approximation [7] yields a simple and elegant expression for the dielectric scaling function of the Coulomb gas.

The relationship between the dilute gas at low temperatures and the hydrogen atom makes it intuitively simple to understand the unusual properties of the two-dimensional Coulomb gas. In two dimensions the electrostatic potential energy of a pair of charges varies logarithmically with the spatial separation between the charges. The thermodynamic probability, i.e. the Boltzmann factor for finding a pair of charges a distance r apart, varies slowly algebraically with r, unlike in the three-dimensional case where it vanishes rapidly exponentially. The algebraic or power law decay of the thermodynamic probability is borne out by calculations which take into account not only the potential energy of the charges but also their kinetic energy. Thus algebraic decays are the rule rather than the exception in two-dimensional Coulomb systems. What is surprising is that there exists a critical temperature above which the behaviour of the system changes over to the exponential form. The exponential form has to be understood in terms of the thermal ionisation of the atom and the many-body effects which come into play above the ionisation temperature.

This paper is organised as follows. Section 2 contains a brief review of the classic Kosterlitz-Thouless theory. In § 3, we calculate the partition function of a two-dimensional hydrogen atom in the WKB approximation, and show that it diverges above a critical temperature. This is interpreted as the thermal ionisation of the atom. The moments of the charge density distribution in the atom as well as their thermal averages are also calculated in the WKB and s-wave approximation. In this approximation, the thermal average of each moment diverges at a distinct temperature. The physical significance of this result is discussed. In § 4, we consider a neutral Coulomb gas above the thermal ionisation temperature and obtain the dielectric scaling function in the Thomas-Fermi approximation. The question of the screening length of the two-dimensional gas and its independence from the density of the gas is also discussed. Section 5 contains a summary and some concluding remarks.

2. Kosterlitz–Thouless theory

Kosterlitz and Thouless [2] were among the first to elucidate the thermal behaviour of the two-dimensional neutral Coulomb gas. Subsequently, several authors [8-10] have studied the problem, and as a result of their work the physical content of the Kosterlitz-Thouless theory has become clearer, and the mathematical steps have also become simplified. The main features of the Kosterlitz-Thouless theory can now be summarised as follows.

Consider a neutral Coulomb gas of rather low density. At low temperatures the gas consists of bound dipolar pairs of equal and opposite charges e and $-e$. The potential energy of a dipole pair of separation r is

$$V(r) = e^2 \int_{r_0}^{r} \frac{dr'}{r'\varepsilon(r')} \tag{2.1}$$

where r_0 is an arbitrary constant and $\varepsilon(r)$ is the dielectric constant of the gas. The dielectric constant depends upon the separation r because of the screening effects of other dipoles in the system. The dipoles in the system occur in various states of thermal excitation and, therefore, with various separations r. Let $n(r)$ be the density of dipoles of separation r:

$$n(r) = n_0^2 \exp(-\beta V(r)) \tag{2.2}$$

where n_0^2 is a normalisation constant and $\beta = 1/k_B T$. Let $p(r)$ be the polarisability of the dipole:

$$p(r) = \tfrac{1}{2}\beta e^2 r^2. \tag{2.3}$$

Now the susceptibility of a circular area of radius r of the neutral Coulomb gas can be calculated by summing over the polarisabilities of all pairs which lie in this area:

$$\chi(r) = \int_{r_0}^{r} n(r')p(r')2\pi r'\,dr'. \tag{2.4}$$

According to the dielectric theory, the susceptibility is related to the dielectric constant by the equation

$$\varepsilon(r) = 1 + 4\pi\chi(r). \tag{2.5}$$

Equation (2.5) gives an integral equation for $\varepsilon(r)$ which contains the essential physics of the problem. The equation for $\varepsilon(r)$ can be put in a simple form if we define

$$y(r) = n_0 r^2 \exp(-\tfrac{1}{2}\beta V(r)). \tag{2.6}$$

After a little algebra, equation (2.5) can be rewritten in the form of two coupled differential equations:

$$r\frac{dy}{dr} = 2\left(1 - \frac{\beta e^2}{4\varepsilon(r)}\right)y \tag{2.7}$$

$$r\frac{d\varepsilon}{dr} = 4\pi^2 e^2 \beta y^2. \tag{2.8}$$

Equations (2.7) and (2.8) are subject to the initial conditions:

$$\varepsilon(r_0) = 1 \qquad y(r_0) = n_0 r_0^2. \tag{2.9}$$

As we have already pointed out, r_0 is an arbitrary constant which enters the expression for the two-dimensional Coulomb potential. In writing the initial conditions (2.9), we have taken r_0 to represent the smallest distance of physical interest in the problem so that at this distance the screening effect of other charges in the system is zero. With these initial conditions

$$r\frac{dy}{dr}\bigg|_{r=r_0} = 2\left(1 - \frac{e^2}{4k_B T}\right)y \tag{2.10}$$

$$r\frac{d\varepsilon}{dr}\bigg|_{r=r_0} = 4\pi^2 e^2 \beta y^2. \tag{2.11}$$

The above equations reveal that the solutions of (2.7) and (2.8) will fall into two categories. If $k_B T > \frac{1}{4}e^2$, then $y(r)$ will increase with increasing r, and so will $\varepsilon(r)$. Eventually, as r goes to infinity, $\varepsilon(r)$ will also go to infinity. If, on the other hand, $k_B T < \frac{1}{4}e^2$, then y will decrease with increasing r, and $\varepsilon(r)$ will also decrease. In this case the rate at which $\varepsilon(r)$ decreases as compared with $y(r)$ is important, and numerical solutions show that as r goes to infinity, $\varepsilon(r)$ approaches a constant value between zero and $\frac{1}{4}e^2$ depending upon $y(r_0)$. A finite value of $\varepsilon(r)$ as r goes to infinity means that the gas is an insulator, while an infinite value of $\varepsilon(r)$ as r goes to infinity means that the gas is a conductor. The transition from an insulator phase below $k_B T_c = \frac{1}{4}e^2$ to a conductor above this temperature is the Kosterlitz-Thouless transition. It should be noted that, at the starting point of the above analysis, the Coulomb gas was visualised as a dilute assembly of dipolar pairs of charges of varying sizes. A self-consistent equation for the dielectric constant of the system was obtained by considering the polarisation effects of larger dipoles on the smaller dipoles. There are several approximations inherent in such an approach. The physical picture of the system on which the mathematical approximations are based deteriorates as the temperature increases and the dipoles begins to dissociate. It is quite difficult to solve the problem without making any approximations at all but in the following we study the low-temperature and high-temperature phases of the Coulomb gas separately in a somewhat different set of approximations. As explained in § 1, the properties of a sufficiently dilute gas at low temperatures may be related to the properties of an individual two-dimensional atom. In the next section we therefore study the quantum mechanics and the thermal behaviour of a two-dimensional atom in some detail.

3. Two-dimensional atom and thermal ionisation

The bound states of a pair of charges e and $-e$ in two dimensions are described by the Schrödinger equation:

$$-\frac{\hbar^2}{2\mu}\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}\right)\psi(r,\theta)+e^2\ln\left(\frac{r}{r_0}\right)\psi(r,\theta)=E\psi(r,\theta) \tag{3.1}$$

where μ is the reduced mass of the two charges, r and θ are the polar coordinates of the vector joining the two charges. Equation (3.1) can be easily separated into r and θ parts. The solutions are of the form:

$$\psi(r,\theta) = R(r)\exp(\pm i m\theta) \qquad m = 0, 1, 2, \ldots \tag{3.2}$$

where $R(r)$ satisfies the equation:

$$\frac{d^2R}{dr^2}+\frac{1}{r}\frac{dR}{dr}+\left[\frac{2\mu E}{\hbar^2}-\frac{2\mu e^2}{\hbar^2}\ln\left(\frac{r}{r_0}\right)-\frac{m^2}{r^2}\right]R=0. \tag{3.3}$$

Equation (3.3) can also be rewritten into a slightly more convenient form by transforming the independent variable from r to $t=\ln(r/r_0)$. We get

$$\frac{d^2R}{dt^2}+\left(\frac{2\mu}{\hbar^2}(E-e^2t)r_0^2e^{2t}-m^2\right)R=0. \tag{3.4}$$

Equation (3.4) cannot be solved exactly. However, it is in a form to which the WKB approximately (see, for example, [11]) can be applied readily. The energy eigenvalues are given by the formula:

$$\int_a^b \left(\frac{2\mu}{\hbar^2}(E - e^2 t) r_0^2 e^{2t} - m^2\right) \mathrm{d}t = (n + \tfrac{1}{2})\pi \tag{3.5}$$

where $n = 0, 1, 2, \ldots$, and a and b are the classical turning points given by the zeros of the integrand. The case of circularly symmetric s waves, i.e. $m = 0$, is particularly simple. In this case the classical turning points are $t = -\infty$ and $t = E/e^2$. The energy eigenvalues of the s waves are given by the WKB result:

$$E_{n,m=0} = e^2 \ln[(2n+1)\rho_0] \tag{3.6}$$

where $\rho_0 = R_0/r_0$, $R_0 = (\pi\hbar^2/2\mu e^2)^{1/2}$.

Equation (3.6) can be derived easily from (3.5). The case of non-zero m is somewhat difficult, even in the WKB approximation. It may be handled following the method of Quigg and Rosner [12]. If we substitute $R(r) = u(r)/r^{1/2}$ in equation (3.3), the resulting equation for $u(r)$ becomes identical with the one analysed by Quigg and Rosner. They express the logarithmic potential as a limiting case of a power law potential:

$$\ln\left(\frac{r}{r_0}\right) = \lim_{\nu \to 0} \frac{1}{\nu}\left[\left(\frac{r}{r_0}\right)^{\nu} - 1\right]. \tag{3.7}$$

The above identity holds for all finite values of r, but its use in the region of vanishing r constitutes some error. Quigg and Rosner have solved the equation for $u(r)$ for all values of ν, positive and negative, in the WKB approximation. They find that the limit $\nu = 0$ can be taken without any problem. Furthermore, ν approaching zero from the positive side and ν approaching zero from the negative side give the same result:

$$E_{n,m} = e^2 \ln[(2n + m + 1)\rho_0] \qquad (n = 0, 1, 2, \ldots; m = 0, 1, 2, \ldots). \tag{3.8}$$

Equation (3.8) may be rewritten as

$$E_p = e^2 \ln[p\rho_0] \qquad p = 1, 2, 3, \ldots. \tag{3.9}$$

The partition function Z of a two-dimensional hydrogen atom in thermal equilibrium at temperature T can now be easily evaluated:

$$Z = \sum_p D_p \exp(-E_p/k_B T) \tag{3.10}$$

where D_p denotes the degeneracy of the level E_p; $D_p = \frac{1}{2}(p+1)$ if p is odd and $D_p = \frac{1}{2}(p+2)$ if p is even. After some algebra, we obtain

$$Z = \tfrac{1}{2}\rho_0^{-K}[\xi(K-1) + (1 - 2^{-K})\xi(K)] \tag{3.11}$$

where $K = e^2/k_B T$, and ξ denotes the Riemann zeta function [13]. The partition function diverges for $K < 2$ or $k_B T > \frac{1}{2}e^2$. The probability for an atom in thermal equilibrium at a temperature T to exist in the state E_p is $\exp(-E_p/k_B T)/Z$. Thus the divergence of Z means that the atom has zero probability to be in any definite state, or in other words the atom is ionised. This result is strictly correct for a single atom in an infinite volume. We may take it as a useful approximation for the thermal behaviour of atoms in a dilute gas. The reason why this result is not strictly true for the gas is the following. In evaluating the partition function of a gas, we can use the eigenstates of a free atom only if their extension is less than the mean distance between the atoms. Eigenstates of greater extension can not be used because these overlap

other atoms and are modified by these atoms. Therefore, for the correct evaluation of the partition function of even a dilute gas, the interactions among atoms have to be taken into account. However, for understanding the thermal behaviour of the two-dimensional gas, the approximation in which interactions among atoms are altogether omitted appears useful in view of its simplicity.

The Kosterlitz-Thouless transition is marked by the divergence of the susceptibility of the gas at the transition temperature. In order to see if this effect can be seen in a one-atom approximation, it is necessary to calculate the thermal expectations of the moments of the charge density distribution of the atom. The electric field couples to the first moment of the charge density distribution, and divergent electric susceptibility implies that thermal expectation of the first moment should diverge at a critical temperature. For simplicity, we calculate the moments and their thermal averages only for $m = 0$ states in the WKB approximation. The effects we are looking for are already present in these states, and the inclusion of $m \neq 0$ states is not expected to make any qualitative change in the following discussion.

The WKB expressions for the expectation value of the kth moment in the nth energy state, and for the thermal average of the kth moment, are, respectively,

$$\langle r^k \rangle_n = \int_0^{r_c} r^k [E_n - V(r)]\,\mathrm{d}r \Big(\int_0^{r_c} [E_n - V(r)]\,\mathrm{d}r \Big)^{-1} \tag{3.12}$$

and

$$\langle r^k \rangle_T = \sum_n \langle r^k \rangle_n \exp(-E_n / k_B T). \tag{3.13}$$

We find for s waves

$$\langle r^k \rangle_n = (k+1)^{-1/2} [R_0(2n+1)]^k \tag{3.14}$$

and

$$\langle r^k \rangle_T = R_0^k (k+1)^{-1/2} \rho_0^{-K} (1 - 2^{k-K}) \xi(K-k). \tag{3.15}$$

Equation (3.15) shows that $\langle r^k \rangle_T$ diverges unless $k_B T < e^2/(k+1)$. Thus not only the thermal average of the first moment diverges but all other moments also diverge at distinct temperatures. The moments of the charge density distribution couple to various gradients of the external electric field applied to the system. It follows from (3.15) that the two-dimensional atom undergoes a series of phase transitions at critical temperatures given by $k_B T_c^k = e^2/(k+1)$, $k = 1, 2, \ldots$, which correspond to divergent susceptibilities with respect to the gradients of the externally applied inhomogeneous electric field. This result will remain qualitatively true even when $m \neq 0$ states are taken into account. However, the inclusion of $m \neq 0$ states will systematically suppress all the transition temperatures.

4. Thomas–Fermi approximation

Above the thermal ionisation temperature, the atoms dissociate and the charges become itinerant. In this case, the one-atom approximation is no longer useful. Instead the problem becomes that of an electron gas of a given density superimposed on a uniform background of positive charge of the same density. It is more appropriate now to label the electrons by wavevectors, rather than stationary states of a single atom. Let $n(r, \theta)$ and $\phi(r, \theta)$ denote, respectively, the number density of electrons and the potential field at the point (r, θ) referred to some conveniently chosen origin in the

electron gas. The local potential energy of the electron is $-e\phi(r, \theta)$, and its kinetic energy is $\hbar^2 k_F^2/2\mu$ where $\hbar k_F$ is the local Fermi momentum of the electron. The Fermi momentum of the electron is related to the number density through the relation

$$k_F^2 = 2\pi n(r, \theta). \tag{4.1}$$

Thus the local kinetic energy is $\pi n(r, \theta)\hbar^2/\mu$. In equilibrium, the sum of the local kinetic energy and the local potential energy should be a constant independent of the spatial position of the electron. We may choose this constant to be zero since the energy is arbitrary up to an additive constant. Thus

$$\pi n(r, \theta)\hbar^2/\mu = e\phi(r, \theta). \tag{4.2}$$

Equation (4.2) is the Thomas-Fermi approximation. The potential field $\phi(r, \theta)$ is determined by the two-dimensional Poisson equation:

$$\nabla^2\phi(r, \theta) = 2\pi e n(r, \theta)$$

or

$$\nabla^2\phi(r, \theta) = (1/a^2)\phi(r, \theta) \tag{4.3}$$

where

$$a = (\hbar^2/2\mu e^2)^{1/2}. \tag{4.4}$$

It is seen from (4.3) that a plays the role of the screening length in the two-dimensional electron gas. It is a special feature of the two-dimensional Coulomb gas that the screening length is a universal constant independent of the density of the gas. The many-body screening effects are best embodied in the position dependence of the dielectric constant of the gas. If we limit ourselves to the circularly symmetric case, $\varepsilon(r)$ is defined by

$$\phi(r) = -e\int_{r_0}^{r} \frac{dr'}{r'\varepsilon(r')}. \tag{4.5}$$

There is a boundary condition on $\varepsilon(r)$, namely $\varepsilon(r_0) = 1$, where r_0 is the smallest distance of physical significance in the problem. In other words, we require $\phi(r)$ to be the unscreened two-dimensional Coulomb potential at very short distances:

$$\phi(r) = -e\ln(r/r_0) \qquad r \approx r_0. \tag{4.6}$$

Equation (4.6) determines the relevant solution of (4.3) to be $K_0(r/a)$, where K_0 is the modified Bessel function of order zero [13]. Consequently the dielectric constant is given by

$$\varepsilon(r) = -r\frac{d}{dr}K_0\left(\frac{r}{a}\right) = \frac{r}{a}K_1\left(\frac{r}{a}\right). \tag{4.7}$$

It can be verified that $\varepsilon(r)$ possesses the desired asymptotic forms:

$$\lim_{r\to 0} \varepsilon(r) = 1$$

$$\lim_{r\to\infty} \varepsilon(r) = (\pi a/2r)^{1/2}\exp(-r/a). \tag{4.8}$$

The explicit form for $\varepsilon(r)$ obtained in equation (4.7) is a direct result of the Thomas-Fermi approximation. It is not possible to obtain this from the classical Kosterlitz-Thouless equations (2.7) and (2.8). However, the physics embodied in the Kosterlitz-Thouless equations remains unchanged by the inclusion of quantum effects.

5. Summary and concluding remarks

The properties of the two-dimensional neutral Coulomb gas have been analysed in two separate approximations: the one-atom approximation and the Thomas-Fermi approximation. The one-atom approximation is applicable at low temperatures to a gas of low density in which the interactions among atoms can be neglected. This would happen if the interatomic separation determined by the density of the gas is larger than the size of the atom in its lowest state. The one-atom approximation shows that the gas is an insulator at low temperatures and makes a phase transition to a conducting state above a critical temperature. In our approximation, the phase transition is due to the thermal ionisation of the atom but it is so similar to the Kosterlitz-Thouless transition that the two transitions are very probably the same. Above the transition temperature, charges become itinerant and interactions among charges become most important. These interactions give rise to screening effects which have been analysed in the Thomas-Fermi approximation. The screening length turns out to be a universal constant independent of the density of the gas. The screening length is of the same order of magnitude as the size of a single atom in its ground state. The screening length is a fraction (1.25) larger than the size of the atom in its lowest state. This suggests that gases of even moderately high densities may remain insulators at low temperatures. There is a possibility that just as a gas of low (fixed) density makes an insulator to conductor transition as its temperature is raised, a gas at a fixed temperature (say, $T = 0$) may make an insulator to conductor transition as its density is increased. However, we are not in a position to answer this question here because the approximations made in our analysis are not appropriate for gases of very high density.

Acknowledgments

It is a pleasure to thank Professor R J Elliott, Dr R B Stinchcombe, Dr S L A de Queiroz and Dr P M Duxbury for helpful conversations.

References

[1] Hauge E H and Hemmer P C 1971 *Phys. Norvegica* **5** 209
[2] Kosterlitz J M and Thouless D J 1973 *J. Phys. C: Solid State Phys.* **6** 1181
[3] Chui S T and Weeks J D 1976 *Phys. Rev.* B **14** 4978
[4] Zittartz J 1978 *Z. Phys.* B **31** 63
[5] Solyom J 1979 *Adv. Phys.* **28** 201
[6] Peierls R 1979 *Surprises in Theoretical Physics* (Princeton: Princeton University Press) p 53
[7] Thomas L H 1926 *Proc. Camb. Phil. Soc.* **23** 542
Fermi E 1928 *Z. Phys.* **48** 73
[8] Kosterlitz J M 1974 *J. Phys. C: Solid State Phys.* **7** 1046
[9] Young A P 1978 *J. Phys. C: Solid State Phys.* **11** L453
[10] Jose J V, Kadanoff L P, Kirkpatrick S and Nelson D R 1977 *Phys. Rev.* B **16** 1217
[11] Merzbacher E 1970 *Quantum Mechanics* 2nd edn (New York: Wiley)
[12] Quigg C and Rosner J L 1979 *Phys. Rep.* **56C** 167
[13] Abramowitz M and Stegun I A (ed) 1972 *Handbook of Mathematical Functions* (New York: Dover)