Occupation probability under the canonical ensemble

1995 J. Phys.: Condens. Matter 7 6707

(http://iopscience.iop.org/0953-8984/7/33/009)

J. Phys.: Condens. Matter 7 (1995) 6707–6716. Printed in the UK

Occupation probability under the canonical ensemble

Hui-Ping Chen† and Ho-Fai Cheung
Department of Physics and Materials Science, City University of Hong Kong, Hong Kong

Received 20 February 1995, in final form 12 May 1995

Abstract. We have studied the predictions of the canonical ensemble concerning the occupation probability for a model with an infinite uniform ladder of degenerate levels. Results show that in the high-degeneracy or high-temperature limit, the canonical ensemble occupation probability would be the same as that under the grand canonical ensemble (given by the Fermi–Dirac distribution). The difference between them is of the order of $O(1/g)$ where g is the degeneracy of the levels. In the low-temperature limit, this difference depends on the corresponding Fermi energy. If the number of electrons is such that the Fermi energy is equal to the energy of a highly degenerate level, then this difference is small. The maximum difference occurs when the Fermi energy is in between two levels. These results are applicable to 1D perfect mesoscopic rings.

1. Introduction

The microcanonical ensemble is applicable for systems with total energy and number of particles fixed, the canonical ensemble is applicable for systems with temperature and number of particles fixed, and the grand canonical ensemble is applicable for systems with temperature and chemical potential fixed. When applied to a macroscopic system the three ensembles yield the same occupation probability, i.e. the Fermi–Dirac distribution. Therefore one would expect the predicted physical properties of a macroscopic system to be the same, regardless of the ensemble used. This result does not necessarily apply to smaller systems at the mesoscopic scale. Denton *et al* have studied the behaviour of small metal clusters using the canonical ensemble (Denton *et al* 1973). They found that there are quantitative differences between the results of the canonical and the grand canonical ensemble. Recent studies of magnetoresistance of mesoscopic systems (see, e.g., Sharvin and Sharvin 1981, Altshuler *et al* 1982, Webb *et al* 1985) have revealed that the properties of a system before and after ensemble averaging can be qualitatively different. Studies of persistent current of mesoscopic rings (see, e.g., Levy *et al* 1990, Bouchiat and Montambaux 1989, Chandrasekhar *et al* 1991, Mailly *et al* 1993) have revealed that there are qualitative differences when averaging using different ensembles. The canonical ensemble must be used to predict the average persistent current in isolated mesoscopic rings.

In deriving the occupation probability for a macroscopic system under the canonical ensemble, the result will be given by the Fermi–Dirac distribution only if the single-particle levels of the system occur together so that there are many levels in a small energy range whenever they occur. This condition can be satisfied in the following two cases: (i) the levels are discrete but highly degenerate, or (ii) the levels form continuum bands with

† Permanent address: Centre for Fundamental Physics, University of Science and Technology of China, Hefei, Anhui 230026, People's Republic of China.

0953-8984/95/336707+10$19.50

or without energy gaps in between. For a macroscopic system it is generally true that the levels form bands with the energy spacings between the adjacent levels being of the order of the band width divided by the number of unit cells. Hence the energy spacings would be too small to produce any observable deviation from the occupation probability given by the Fermi–Dirac distribution since thermal fluctuation at reasonable temperatures would be sufficient to cause excitation to many levels above the Fermi surface, or any inelastic scattering would be sufficient to broaden the levels beyond the energy spacings. In both cases the spectrum behaves essentially like a continuous one. However the situation may be different for a mesoscopic system. If the system is sufficiently small, the energy spacings could be larger than both the energy of thermal fluctuation and the amount of level broadening caused by inelastic scattering. These two conditions may be satisfied in isolated mesoscopic systems. Then one may expect that the occupation probability would be significantly different from the Fermi–Dirac distribution.

Suppose an isolated mesoscopic ring is built on an insulating substrate, would the number of electrons in the ring be fixed? On first thought it would seem that a mesoscopic ring could have a different number of electrons from time to time, even though it is isolated by an insulator. The reason is that since the energy level spacings are small, it would not cost too much energy to have extra electrons on the ring. However Schmid pointed out that the Coulomb charging energy of putting in one extra electron could make the process energetically unfavourable at low temperatures (Schmid 1991). For example, the charging energy of a conductor is roughly equal to the square of the charge divided by the capacitance. In this way, the charging energy of putting in one more electron on a conductor of size 1 μm is estimated to be 10 K. So at lower temperatures the number of electrons should be considered fixed and thus the canonical ensemble should be used. An extreme example is provided by the electrons in a microscopic system such as an atom. Normally, ionized atoms are extremely rare in comparison with the number of neutral atoms except at very high temperatures. Hence in describing the atom's equilibrium properties or electronic transition processes at nearly equilibrium, the canonical ensemble should be used. The Fermi–Dirac distribution should not be used in this case since it would not be a good approximation for the electron occupation probability.

In the following we discuss the occupation probability of a simple model under both the canonical ensemble and the grand canonical ensemble. The model is physically realistic and yet simple enough so that analytic results can be obtained. In this way some general properties of the occupation probability from the two ensembles can be studied. In particular we would like to know how good it would be if we approximate the canonical ensemble occupation probability by that from the grand canonical ensemble.

2. The model

The model we studied is a model of non-interacting electrons with single-particle energy levels that are uniformly spaced and of degeneracy g. This model has been analysed before (Denton *et al* 1973). The model can be applied to one-dimensional (1D) perfect rings or to fictitious crystals having very narrow energy bands separated by uniform band gaps. For 1D perfect rings the energy levels may be approximated by a uniform ladder near the Fermi level. In the mesoscopic regime, the persistent current in such 1D perfect rings at non-zero temperatures has been calculated by using the grand canonical ensemble (Cheung *et al* 1988). Even for this simple model, the persistent current at non-zero temperatures under the canonical ensemble has not been calculated because the occupation probability is not known exactly. In this paper we present in more details than in the previous studies the occupation

probability of the levels of the infinite uniform energy ladder model in the canonical ensemble for the cases of degeneracy $g = 1$ and $g = 2$. This two cases describe the physical situations where the spin degeneracy is lifted or not lifted respectively. We also present new results in the high-degeneracy limit at both high and low temperatures. Comparison with the results of the grand canonical ensemble (i.e. the Fermi–Dirac distribution) are also considered.

Consider electrons in a semi-infinite uniformly spaced energy ladder with g-fold degenerate levels and energy spacing Δ. Let the number of electrons in the system be fixed at $Ng - r$ where N and r are integers. Without loss of generality, r is restricted to lie between zero and $g - 1$. The canonical ensemble is used to calculate the occupation probability of the levels. First note that the ground state of the system (not the ground state of a single electron) consists of the first $(N - 1)$ levels filled up, $g - r$ electrons in the Nth level, and the $(N + 1)$th levels onward empty. We divide the levels into two groups: (i) the levels starting from the Nth down to the first are labelled with $0, 1, 2, \ldots, N - 1$ respectively and they are called the lower levels; (ii) the levels from the $(N + 1)$th onward are labelled with $1, 2, 3, \ldots, \infty$ and they are called the upper levels. We will specify a state by the number of holes in the lower levels and the number of electrons in the upper levels. In this way the ground state of the system consists of r holes at the uppermost lower level, no holes at the other lower levels, and no electrons at any upper levels. In the following we measure the energy of a level from that of the uppermost lower level. In this way the energy of a hole on the lower levels labelled with j is $j\Delta$, and that the energy of an electron on the upper levels labelled with k is $k\Delta$. We also measure the energy of the whole system from the system's ground state energy. In this way the ground state energy is zero, and the energy of any excited state is equal to the sum of energy of electrons in the upper levels plus the sum of energy of holes in the lower levels. After this transformation, the upper limit of the label of the lower levels can be extended from $(N - 1)$ to ∞.

Following the method used by Denton *et al* (1973) and originally discussed by Kubo (1962), the partition function can be written as

$$Z(\beta) = \frac{1}{2\pi i}\oint \frac{dz}{z} z^r \prod_{j=0}^{\infty}\left[1 + \frac{1}{z}e^{-j\beta\Delta}\right]^g \prod_{k=1}^{\infty}[1 + ze^{-k\beta\Delta}]^g \tag{1}$$

where β is the inverse of the temperature with the Boltzmann factor, and z is an integration variable. The integration is over a contour on the complex z plane enclosing the point $z = 0$. The various terms can be understood as follows. Firstly the contribution to the final Boltzmann factor of an upper level labelled with k is either unity or $\exp(-k\beta\Delta)$ depending on whether there is no or one electron in that level. In the same way the contribution to the final Boltzmann factor of a lower level labelled with j is either unity or $\exp(-j\beta\Delta)$ depending on whether there is no or one hole in that level. The various ways of putting electrons and holes in the levels lead to different Boltzmann factors which are completely summed up in the term $\prod[1 + \exp(-j\beta\Delta)]^g \prod[1 + \exp(-k\beta\Delta)]^g$. Each $[1 + \exp(-j\beta\Delta)]$ and $[1 + \exp(-k\beta\Delta)]$ is raised to the gth power because all the levels are g-fold degenerate. The number of electron in the levels is counted via the power of z in $[1 + (z)\exp(-k\beta\Delta)]$ and $[1 + (z^{-1})\exp(-j\beta\Delta)]$. As the ground state of the system has r holes, the factor z^r and the integration in z is used to pick up this term and all the other terms that have the same total number of electrons.

Before considering the occupation probability, let us define the kth level as the level labelled by k and state that k takes a positive value for the upper levels and a negative value for the lower levels. When calculating the occupation probability of the kth level, we sum up the Boltzmann factor of all the states in which the kth level is definitely occupied. The

procedure is very similar to that for calculating the partition function, except now that the Boltzmann factor is for the kth level is $(z)\exp(-k\beta\Delta)$ instead of $[1+(z)\exp(-k\beta\Delta)]$. The occupation probability of the kth level is equal to the above sum divided by the partition function, and can be written as

$$P(k)=\frac{1}{2\pi \mathrm{i}Z}\oint\frac{\mathrm{d}z}{z}z^r\left[\prod_{j=0}^{\infty}\left(1+\frac{1}{z}\mathrm{e}^{-j\beta\Delta}\right)^g\right]\left[\prod_{m=1}^{\infty}(1+z\mathrm{e}^{-m\beta\Delta})^g\right]\frac{z\mathrm{e}^{-k\beta\Delta}}{1+z\mathrm{e}^{-k\beta\Delta}}. \tag{2}$$

After substituting z by $\exp(\mathrm{i}\phi)$ and $\exp(-\beta\Delta/2)$ by q, the partition function can be written as

$$Z=\frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\phi\, 2^g\mathrm{e}^{\mathrm{i}(r-g/2)\phi}\cos^g\left(\frac{\phi}{2}\right)\prod_{n=1}^{\infty}\left[1+2q^{2n}\cos 2\left(\frac{\phi}{2}\right)+q^{4n}\right]^g. \tag{3}$$

The infinite product in (3) can be simplified using the following identity (Whittaker and Watson 1935):

$$\prod_{n=1}^{\infty}(1+2q^{2n}\cos 2z+q^{4n})=\frac{1}{2\cos z}q^{-1/4}\prod_{m=1}^{\infty}(1-q^{2m})^{-1}\sum_{n=-\infty}^{\infty}q^{(n+1/2)^2}\mathrm{e}^{\mathrm{i}(2n+1)z}. \tag{4}$$

Then the integration over ϕ can be done readily and the partition function is obtained as

$$Z=Z_{\mathrm{B}}^g\mathrm{e}^{\beta g\Delta/8}\sum_{n_1=-\infty}^{\infty}\cdots\sum_{n_g=-\infty}^{\infty}\delta_{r+\sum_{i=1}^g n_i,0}\exp\left[-\frac{\beta\Delta}{2}\sum_{i=1}^{g}\left(n_i+\frac{1}{2}\right)^2\right] \tag{5}$$

where

$$Z_{\mathrm{B}}=\prod_{n=1}^{\infty}(1-\mathrm{e}^{-n\beta\Delta})^{-1}. \tag{6}$$

The integral for the occupation probability (2) can be simplified in the same way. We obtained

$$P(k)=\frac{1}{Z}Z_{\mathrm{B}}^g\mathrm{e}^{\beta g\Delta/8}\sum_{n_1=-\infty}^{\infty}\cdots\sum_{n_g=-\infty}^{\infty}\sum_{m=1}^{\infty}\delta_{r+m+\sum_{i=1}^g n_i,0}(-1)^{m+1}$$
$$\times\exp\left\{-\frac{\beta\Delta}{2}\left[2mk+\sum_{i=1}^{g}\left(n_i+\frac{1}{2}\right)^2\right]\right\}. \tag{7}$$

This is the basic equation for the occupation probability. We will present results for the $g=1$ and $g=2$ cases in the next section, and then high- and low-temperature results in later sections. Before we do that we first discuss symmetries in the system.

There is particle–hole symmetry and 'translational' symmetry of the energy ladder. We find that these symmetries are reflected in the occupation probability. Under a mapping that turns the energy ladder upside down and then transforms particles into holes and holes into particles, the occupation probability of the levels after the mapping should be equal to the probability of putting a hole into the corresponding levels before the mapping. This is one example of the particle–hole symmetry and it manifests itself by the relationship

$$P(k,r)=1-P(1-k,-r). \tag{8}$$

Thus once the occupation probability for positive k is known, the occupation probability for negative k can be deduced from (8). Another symmetry is the 'translational' symmetry of the energy ladder: adding g holes to the system is equivalent to shifting the ladder downward by one step. This symmetry manifests itself by the relationship

$$P(k,r)=P(k-1,r+g). \tag{9}$$

3. Results

We present results for the $g = 1$ and 2 cases. The occupation probability obtained from the canonical ensemble ($P(k)$) and that from the grand canonical ensemble ($P_{FD}(k)$) will be compared. The objective is to find out how well the canonical ensemble occupation probability is approximated by the Fermi–Dirac distribution. To evaluate $P_{FD}(k)$ we choose the chemical potential (μ) in such a way that the average number of electrons in the grand canonical ensemble is equal to the number of electrons fixed in the canonical ensemble. This can be achieved by requiring that on average there are r more holes in the lower levels than electrons in the upper levels, i.e.

$$\sum_{n=0}^{\infty} \frac{g}{e^{\beta(n\Delta+\mu)}+1} - \sum_{n=1}^{\infty} \frac{g}{e^{\beta(n\Delta-\mu)}+1} = r. \tag{10}$$

From this equation the corresponding chemical potential (μ) can be deduced. Note that normally the chemical potential would depend on temperature. However for the special cases where r is equal to zero or $g/2$ (the uppermost lower level is half filled or completely filled in the ground state of the system), the chemical potential turns out to be independent of temperature.

Considering the $g = 1$ case (non-degenerate levels), the occupation probability obtained from the canonical ensemble was found by setting g to unity in (7). The result is

$$P(k) = \left(\sum_{m=1}^{\infty} (-1)^{m+1} \exp\left\{ -\frac{\beta\Delta}{2}\left[2mk + \left(m + r - \frac{1}{2} \right)^2 \right] \right\} \right) \Big/ \exp\left\{ -\frac{\beta\Delta}{2}\left(r - \frac{1}{2} \right)^2 \right\}. \tag{11}$$

$P(k)$ is evaluated numerically from (11). When evaluating $P_{FD}(k)$, the chemical potential is obtained from (10) as $(-r + \frac{1}{2})\Delta$. Putting these together,

$$P_{FD}(k) = 1/(\exp[(k + r - \tfrac{1}{2})\beta\Delta] + 1). \tag{12}$$

Changing r is equivalent to shifting the energy ladder, as described by (9). Without loss of generality we can set r to zero. Results for $P(k)$ and $P_{FD}(k)$ given by (11) and (12) are shown as a function of temperature in figure 1. The graphs are plotted as a function of $\exp[-\beta(E_k - \mu)]$, which is equal to $\exp[-(k - \frac{1}{2})\beta\Delta]$ in this case. The advantage is that the graphs for all $P_{FD}(k)$ fall into a single line. From figure 1, we notice that $P(k)$ approaches $P_{FD}(k)$ as k increases.

The results for the doubly degenerate case follow. Setting g to two in equation (7), the occupation probability in the canonical ensemble is obtained as

$$P(k) = \left(\sum_{n=-\infty}^{\infty} \sum_{m=1}^{\infty} (-1)^{m+1} \exp\left\{ -\frac{\beta\Delta}{2}\left[2mk + \left(n + \frac{1}{2} \right)^2 + \left(n + r + m - \frac{1}{2} \right)^2 \right] \right\} \right) \times \left(\sum_{n=-\infty}^{\infty} \exp\left\{ -\frac{\beta\Delta}{2}\left[\left(n + \frac{1}{2} \right)^2 + \left(n + r - \frac{1}{2} \right)^2 \right] \right\} \right)^{-1}. \tag{13}$$

When there is no extra hole in the system ($r = 0$), the chemical potential is obtained from equation (10) as $\Delta/2$. The corresponding occupation probability $P_{FD}(k)$ is the same as that given by (12) with r set to zero. Results for $P(k)$ and $P_{FD}(k)$ are shown in figure 2. We notice that as k increases, $P(k)$ approaches $P_{FD}(k)$. This feature is similar to that in the $g = 1$ case. We also notice another feature, that, for the same k, $P(k)$ is closer to $P_{FD}(k)$ in the $g = 2$ case than in the $g = 1$ case. The difference between $P(k)$ and $P_{FD}(k)$ seems to scale as $1/g$.

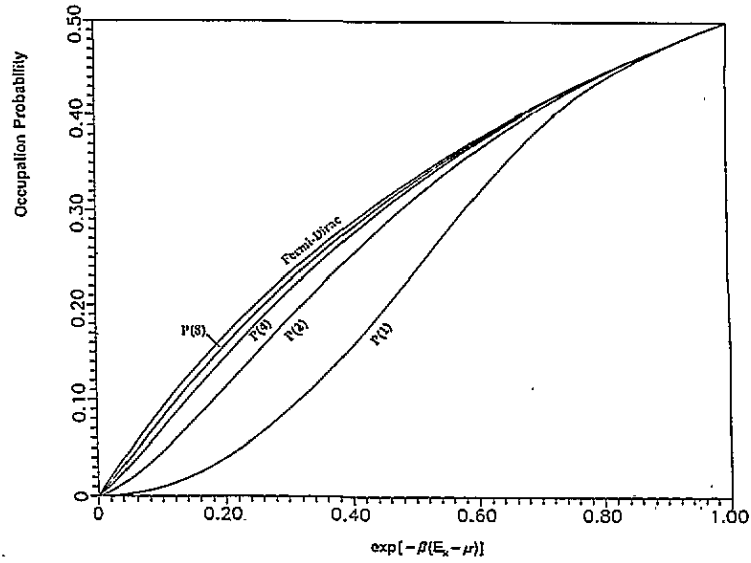

Figure 1. The occupation probability of the $k = 1, 2, 4$ and 8 levels obtained from the canonical ensemble and that from the Fermi–Dirac distribution as a function of $\exp[-\beta(E_k - \mu)]$ for the case $g = 1$ and $r = 0$.

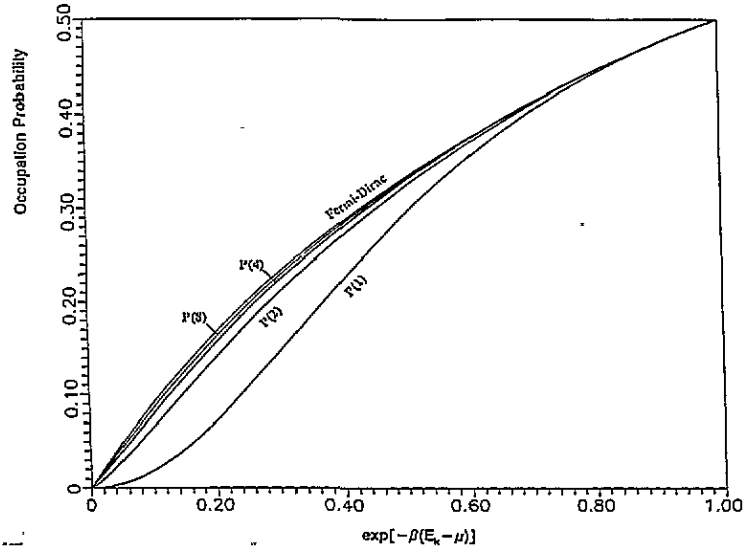

Figure 2. The occupation probability of the $k = 1, 2, 4$ and 8 levels obtained from the canonical ensemble and that from the Fermi–Dirac distribution as a function of $\exp[-\beta(E_k - \mu)]$ for the case $g = 2$ and $r = 0$.

When there is one extra hole in the system ($r = 1$), the chemical potential is obtained from equation (10) as zero. The occupation probability is obtained as

$$P_{\mathrm{FD}}(k) = 1/[\exp(k\beta\Delta) + 1]. \tag{14}$$

Results for $P(k)$ and $P_{\mathrm{FD}}(k)$ are shown in figure 3. The above-mentioned features are also present in this case. One additional feature is that $P(k)$ is closer to $P_{\mathrm{FD}}(k)$ in the $r = 1$ case

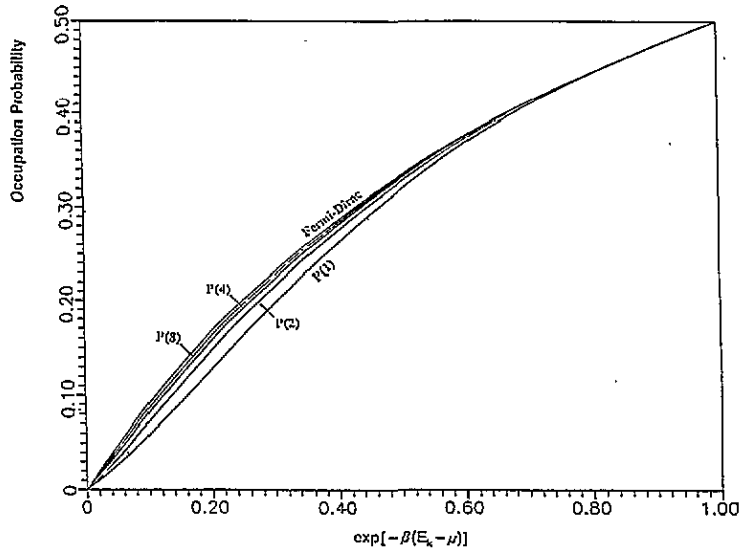

Figure 3. The occupation probability of the $k = 1$, 2, 4 and 8 levels obtained from the canonical ensemble and that from the Fermi–Dirac distribution as a function of $\exp[-\beta(E_k - \mu)]$ for the case $g = 2$ and $r = 1$.

than in the $r = 0$ case. $P(k)$ and $P_{\mathrm{FD}}(k)$ are closer to each other when the Fermi energy (chemical potential at $T = 0$) is equal to the energy of a level. This happens for the $g = 2$, $r = 1$ case. This point will be discussed in relation to the low-temperature behaviour in section 5.

From the above result, we expected that the difference between $P(k)$ and $P_{\mathrm{FD}}(k)$ would scale as $1/g$ in general. To test this prediction we have worked out the solution of a two-level model with degeneracy g. We found that the leading correction to the Fermi–Dirac distribution does scale as $1/g$ as g approaches infinity. In the following two sections we will present similar analysis for the infinite energy ladder model introduced in section 2.

4. The high-temperature behaviour

In this section we look at the high-temperature behaviour of $P(k)$ and $P_{\mathrm{FD}}(k)$ for general degeneracy. The major objectives are to study $P(k)$ and $P_{\mathrm{FD}}(k)$ at high temperatures and also in the high-degeneracy limit. Starting from equations (5) and (7), $P(k)$ can be written as

$$P(k) = \Bigg(\sum_{n_1=-\infty}^{\infty} \cdots \sum_{n_{g-1}=-\infty}^{\infty} \sum_{m=1}^{\infty} (-1)^{m+1} \exp\Bigg\{ -\frac{\beta\Delta}{2}\Bigg[2mk + \Bigg(\sum_{i=1}^{g-1} \Big(n_i + \frac{1}{2} \Big)^2 \Bigg) + \Bigg(m + \Bigg(\sum_{i=1}^{g-1} n_g \Bigg) + r - \frac{1}{2} \Bigg)^2 \Bigg] \Bigg\} \Bigg) \times \Bigg(\sum_{n_1=-\infty}^{\infty} \cdots \sum_{n_{g-1}=-\infty}^{\infty} \exp\Bigg\{ -\frac{\beta\Delta}{2}\Bigg[\sum_{i=1}^{g-1} \Big(n_i + \frac{1}{2} \Big)^2 + \Bigg(\Bigg(\sum_{i=1}^{g-1} n_i \Bigg) + r - \frac{1}{2} \Bigg)^2 \Bigg] \Bigg\} \Bigg)^{-1}. \tag{15}$$

At high temperatures when $\beta\Delta \ll 1$, the sums over the n_i can be replaced by integrals (corresponds to using the leading term in the Poisson summation formula). This will introduce an error which is of the order of $\exp[-\pi^2/\beta\Delta]$ multiplied by the leading term. After integrating out all the n_i, $P(k)$ is obtained as

$$P(k) = \sum_{m=1}^{\infty}(-1)^{m+1}e^{-\beta\Delta m(k-1/2+r/g)}e^{-(\beta\Delta/2g)m^2}. \tag{16}$$

Let us compare $P(k)$ in (16) with $P_{FD}(k)$. $P_{FD}(k)$ follows the Fermi–Dirac distribution with the chemical potential deduced from (10). The chemical potential at high temperatures, when $\beta\Delta \ll 1$, is obtained as

$$\mu = \left(\frac{1}{2} - \frac{r}{g} - \frac{4\pi}{\beta\Delta}e^{-2\pi^2/\beta\Delta}\sin\frac{2\pi r}{g} + \cdots\right)\Delta. \tag{17}$$

Neglecting the exponentially small terms in equation (17), $P_{FD}(k)$ can be written as

$$P_{FD}(k) = \frac{1}{e^{\beta\Delta(k-1/2+r/g)}+1} = \sum_{m=1}^{\infty}(-1)^{m+1}e^{-\beta\Delta(k-1/2+r/g)}. \tag{18}$$

Results from figures 1–3 for the $g = 1$ and $g = 2$ cases have already shown that $P_{FD}(k)$ is always larger than $P(k)$. Let us concentrate on the difference between $P(k)$ and $P_{FD}(k)$ at high temperatures. Define $\Delta P(k)$ as $P_{FD}(k) - P(k)$ and from equations (16) and (18), we obtain

$$\Delta P(k) = \sum_{m=1}^{\infty}(-1)^{m+1}e^{-\beta\Delta m(k-1/2+r/g)}(1 - e^{-(\beta\Delta/2g)m^2}). \tag{19}$$

Expanding the last term in (19) as a Taylor series, and evaluating the sum over m by the Poisson summation formula, we obtain

$$\Delta P(k) = \sum_{p=-\infty}^{\infty}\sum_{n=1}^{\infty}\int_0^{\infty} e^{-[\beta\Delta(k-1/2+r/g)-i(2p+1)\pi]x}\frac{1}{n!}\left(-\frac{\beta\Delta}{2g}x^2\right)^n dx. \tag{20}$$

At high temperatures only the $n = 1$ term is important. The other terms are at least $\beta\Delta/g$ times smaller. After integrating over x, $\Delta P(k)$ is obtained as

$$\Delta P(k) = \sum_{p=-\infty}^{\infty} -\frac{\beta\Delta}{g}[\beta\Delta(k-\tfrac{1}{2}+r/g) - i(2p+1)\pi]^{-3}. \tag{21}$$

For low-lying levels such that $\beta\Delta(k-\frac{1}{2}+r/g) \ll 1$, $P_{FD}(k)$ is roughly equal to $\frac{1}{2}$ and we find

$$\Delta P(k) = [(\beta\Delta)^2/16g](k-\tfrac{1}{2}+r/g)[1+O(\beta\Delta/g)]. \tag{22}$$

Hence the ratio of $\Delta P(k)$ to $P_{FD}(k)$ is roughly equal to $[(\beta\Delta)^2/8g](k-\frac{1}{2}+r/g)$, which is much smaller than unity in this limit. For higher-lying levels such that $\beta\Delta(k-\frac{1}{2}+r/g) \gg 1$, $P_{FD}(k)$ is roughly equal to $\exp[-(\beta\Delta/2)(k-\frac{1}{2}+r/g)]$ and from (21) we find that

$$\Delta P(k) = (\beta\Delta/2g)e^{-\beta\Delta(k-1/2+r/g)}[1+O(\beta\Delta/g)]. \tag{23}$$

Hence the ratio of $\Delta P(k)$ to $P_{FD}(k)$ is roughly $\beta\Delta/2g$, which is again much smaller than unity.

The above analysis reveals that $\Delta P(k)$ vanishes in the high-temperature limit. For lower-lying levels such that $P(k)$ is close to $\frac{1}{2}$, the percentage difference between $P(k)$ and $P_{FD}(k)$ is small and is of the order of $(\beta\Delta)^2k/g$. The percentage difference is actually larger for higher k values if temperature is fixed. However this does not contradict the data

in figures 1–3, where $P(k)$ is plotted as a function of $\exp[-\beta(E_k-\mu)]$. Since $E_k = k\Delta$ and $\mu < \Delta$, the percentage difference between $P(k)$ and $P_{FD}(k)$ is roughly $(\beta(k\Delta-\mu))^2/kg$. This result indicates that the percentage difference between $P(k)$ and $P_{FD}(k)$ scales roughly as $1/k$ when $\exp[-\beta(E_k-\mu)]$ is fixed, which confirms the data shown in figures 1–3.

For high-lying levels such that $P(k)$ is close to zero, the percentage difference between $P(k)$ and $P_{FD}(k)$ is small and of the order of $\beta\Delta/g$. Since $\beta\Delta/g$ is roughly equal to $(\beta(k\Delta-\mu))/kg$, the percentage difference between $P(k)$ and $P_{FD}(k)$ scales roughly as $1/k$ when $\exp[-\beta(E_k-\mu)]$ is fixed, which again confirms the data shown in figures 1–3. From this and the above analysis, the percentage difference between $P(k)$ and $P_{FD}(k)$ scales as $1/g$ for any level. We expect that this behaviour holds in general, both for the low-temperature case and also for cases where the levels are not uniformly spaced.

5. The low-temperature behaviour

The analysis in section 4 is for the high-temperature limit when $\beta\Delta \ll 1$. However it seems the difference between the $P(k)$ and $P_{FD}(k)$ still scales as $1/g$ at low temperatures in the high-degeneracy limit. We study this in this section.

At low temperatures, by considering the relevant lowest excited states, the partition function and the occupation probability can be expanded in powers of $\exp(-\beta\Delta)$ as in the following:

$$Z = [g!/(g-r)!r!][1+O(e^{-\beta\Delta})]$$
$$P(k) = \frac{1}{Z}\frac{g!}{(g-r-1)!(r+1)!}e^{-k\beta\Delta}[1+O(e^{-\beta\Delta})] = \frac{g-r}{r+1}e^{-\beta\Delta}[1+O(e^{-\beta\Delta})]. \tag{24}$$

When comparing $P(k)$ in (24) to that in the grand canonical ensemble, one needs the chemical potential, which is deduced from equation (10). At low temperatures when $\beta\Delta \gg 1$, the chemical potential μ for the $r \neq 0$ case is

$$\mu = (1/\beta)[\ln((g-r)/r) + [g^3(2r-g)/r^2(g-r)^2]e^{-\beta\Delta} + \cdots] \tag{25}$$

while for the $r = 0$ case, μ is obtained as $\Delta/2$. Hence the occupation probability for the $r \neq 0$ case in the grand canonical ensemble at low temperatures can be written as

$$P_{FD}(k) = [(g-r)/r]e^{-k\beta\Delta}[1+O(e^{-\beta\Delta})]. \tag{26}$$

Comparing $P_{FD}(k)$ in (26) to $P(k)$ in (24), one notices that they are equal in the large g and r limit. This trend agrees with the data shown in figures 2 and 3. The large g and r limit corresponds to the case where the Fermi energy (chemical potential at $T = 0$) is equal to the energy of a highly degenerate level. The maximum difference between $P(k)$ and $P_{FD}(k)$ occurs when the Fermi energy is in between two levels (i.e. the $r = 0$ case), where the occupation probability is obtained as

$$P_{FD}(k) = e^{-(k-1/2)\beta\Delta}[1+O(e^{-\beta\Delta})]$$
$$P(k) = ge^{-k\beta\Delta}[1+O(e^{-\beta\Delta})]. \tag{27}$$

In the extremely low-temperature limit when $\exp(\beta\Delta/2) \gg g$, $P_{FD}(k)$ is many times larger than $P(k)$ although they are both very close to zero. However the percentage difference between $P(k)$ and $P_{FD}(k)$ will decrease rapidly in the large-g limit when temperature increases. At intermediate temperatures when $\exp(\beta\Delta/2) \ll g$, the percentage difference between $P(k)$ and $P_{FD}(k)$ is small and scales as $1/g$. At still higher temperatures the results in section 4 will hold.

6. Discussion

Our analysis reveals that within the model that we have discussed, the approximation of the occupation probability in the canonical ensemble by that from a suitably chosen Fermi–Dirac distribution is appropriate in the high-degeneracy limit or the high-temperature limit. This approximation does not hold in the limit when both g and T are low. When this happens, there can be significant differences between the two. These differences will be small if the number of electrons is such that the corresponding Fermi energy is equal to the energy of a highly degenerate level (for example when both g and r are large). The maximum difference occurs when the corresponding Fermi energy is in between two levels (i.e. the $r = 0$ case). When applying to a system with non-uniform energy spacing (typical energy spacing Δ), our conclusion should still be valid.

As mentioned before in section 2 the energy levels in 1D perfect rings can be approximated by an infinite uniform energy ladder. Using the results on the occupation probability obtained in this paper, it becomes possible to deduce the persistent current on 1D perfect rings averaged under the canonical ensemble.

Acknowledgments

We wish to thank C Rudowicz and P Chu for careful comments on the manuscript. This work was supported by Hong Kong UPGC Research Grants Council under grant No 904033.

References

Altshuler B L, Aronov A G, Spivak B Z, Sharvin D Yu and Sharvin Yu V 1982 *JETP Lett.* **35** 588
Bouchiat H and Montambaux G 1989 *J. Physique* **50** 2695
Chandrasekhar V, Webb R A, Brady M J, Ketchen M B, Gallagher W J and Kleinsasser A 1991 *Phys. Rev. Lett.* **67** 3578
Cheung H F, Gefen Y, Riedel E K and Shih W H 1988 *Phys. Rev.* B **37** 6050
Denton R, Mühlschlegel B and Scalapino D J 1973 *Phys. Rev.* B **7** 3589
Kubo R 1962 *J. Phys. Soc. Japan* **17** 975
Levy L P, Dolan G, Dunsmuir J and Bouchiat H 1990 *Phys. Rev. Lett.* **64** 2074
Mailly D, Chapelier C and Benoit A 1993 *Phys. Rev. Lett.* **70** 2020
Schmid A 1991 *Phys. Rev. Lett.* **66** 80
Sharvin D Yu and Sharvin Yu V 1981 *JETP Lett.* **34** 272
Webb R A, Washburn S, Umbach C and Laibowitz R A 1985 *Phys. Rev. Lett.* **54** 2696
Whittaker E T and Watson G N 1935 *A Course of Modern Analysis* (London: Cambridge University Press)