

Predicting the ionization threshold for carriers in excited semiconductors

David Snoke*

Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15260, United States

Received 14 September 2007; received in revised form 12 January 2008; accepted 14 January 2008 by R. Merlin
Available online 18 January 2008

Abstract

A simple set of formulas are presented which allow prediction of the fraction of ionized carriers in an electron–hole–exciton gas in a photoexcited semiconductor. These results are related to recent experiments with excitons in single and double quantum wells.

© 2008 Elsevier Ltd. All rights reserved.

PACS: 71.35.Ee; 71.35.Lk

Keywords: A. Semiconductors; D. Mott transition; D. Exciton; D. Plasma

1. Two “Mott” transitions

Many researchers in semiconductor physics talk of “the” Mott transition density in a system of excitons and electron–hole plasma, but do not have a clear handle on exactly how to predict that density as a function of temperature and material parameters in a given system. While numerical studies have been performed for the fraction of free carriers as a function of carrier density and temperature [1,2], these do not give a readily-accessible intuition for the transition. In this paper I present a simple approach which does not involve heavy numerical methods, but is still fairly realistic. The theory is based on two well-known approximations, which are the mass-action equation for equilibrium in which different species can form bound states, and the static (Debye) screening approximation. In addition, simple approximations are used for numerical calculations of the excitonic Rydberg as a function of screening length.

The first issue to deal with is what Mott transition we are concerned with. There are actually two very different conductor–insulator transitions which go under the name of the “Mott” transition. The first, originally envisioned by Mott, occurs when the wave function overlap of bound states increases to the point that banding occurs. This Mott transition is relevant in the case of doping of semiconductors — when the

doping density increases beyond a certain point, a mini-band generated from the dopant states will arise, in which case the dopant carriers are no longer localized but instead can move in Bloch-like delocalized states. The condition for this is $r_s \sim a$, where a is the characteristic length over which the bound state wave function extends, and r_s is the average distance between bound states. If we cube both sides of this, we obtain the condition, for a three-dimensional system,

$$n \sim a^3, \quad (1)$$

where n is the density of dopant electrons.

This Mott transition condition is equivalent to the condition that the Fermi level of the electrons, if they were not in bound states, is much larger than the binding energy in the bound states. We can see this by assuming hydrogenic bound states. In this case a is given by the Bohr radius,

$$a = \frac{4\pi\epsilon\hbar^2}{e^2m}, \quad (2)$$

where ϵ is the zero-frequency dielectric constant of the medium and m is the effective mass of the bound carrier (MKS is used everywhere in this paper). The hydrogenic Rydberg for the bound state is

$$\text{Ry} = \frac{e^2}{8\pi\epsilon a}. \quad (3)$$

By comparison, the Fermi energy for a three-dimensional electron gas with two spin states, determined by the condition

* Tel.: +1 412 624 9007; fax: +1 412 624 9163.
E-mail address: snoke@pitt.edu.

$N = \int_0^{E_F} D(E)dE$, where $D(E)$ is the density of states, is

$$E_F = \left[n \frac{(3\pi^2)^{2/3} \hbar^3}{\sqrt{2} m^{3/2}} \right]^{2/3}, \quad (4)$$

where n is the electron density. If the Fermi energy is much larger than the Rydberg energy, e.g. 10 times greater, we have

$$\left(n \frac{(3\pi^2)^{2/3} \hbar^2}{\sqrt{2} m} \right) \frac{e^2}{8\pi\epsilon a} = 10$$

$$2a \left(n \frac{(3\pi^2)^{2/3} 4\pi\epsilon \hbar^2}{\sqrt{2} e^2 m} \right) = 10$$

or

$$na^3 = \frac{10^{3/2}}{6\pi^2} \sim 1.$$

Essentially, if the Fermi level is large compared to the binding energy in the bound states, the binding energy becomes just a perturbation on the Fermi gas. Since the Fermi gas is a conductor, the Coulomb binding energy has little effect on that. To put it another way, Fermi phase space filling effects, which lead to a Fermi level, become important when there is substantial wave function overlap of the electrons, which is the same point at which banding occurs.

By contrast, there is another, *thermodynamic* insulator–conductor transition which is sometimes also called the Mott transition, although Rice has named it the “ionization catastrophe.” [3] This is usually the relevant transition for photoexcited carriers when the excited carrier density is low enough that there is no Fermi level, which is to say, the chemical potential is below the ground state energy. In this case, depending on the temperature, there can still be a conductor–insulator transition from the carriers mostly in excitons to the carriers mostly in plasma. The transition in this case comes about due to screening out of the Coulomb interaction which leads to formation of excitons, not due to wave function overlap. The general condition is $a \sim l$, where l is the screening length, which enters into the screened Coulomb (Yukawa) potential, $U(r) = q^2 e^{-r/l}/r$. The screening length depends on the density; in the low-density limit in three dimensions it is given by the Debye screening formula (e.g. [4])

$$\frac{1}{l^2} = \frac{e^2 n}{\epsilon k_B T}. \quad (5)$$

We can envision this transition in two different directions. If we start with a fully ionized, high-density neutral plasma, and reduce the density while keeping the temperature constant (as, for example, typically happens due to carrier recombination) then an instability will occur when the total pair density drops to a point such that the screening length is long compared to the excitonic Bohr radius, and exciton formation is no longer screened out. Then essentially all the carriers can form into excitons. The condition for this is found by reversing the above formula, assuming a is a constant:

$$n = \frac{\epsilon k_B T}{a^2 e^2}. \quad (6)$$

On the other hand, we can imagine starting with a low-density gas of only excitons. Since excitons are neutral, we do not expect them to contribute to screening, at least not until their density is high enough that there is substantial wave function overlap, which is the condition (1) above. However, long before they reach this density, free electrons and holes can be generated by thermal ionization. These free carriers can lead to screening, which can lead to a reduced exciton binding energy, which in turn leads to further ionization. At some critical density this can become a runaway process known as the ionization catastrophe, or a thermodynamic Mott transition. To determine the condition for this, we must start with the mass-action equation, or Saha equation, for the number of free carriers in equilibrium,

$$\frac{n_e^2}{n_{ex}} = \frac{n_Q^{(e)} n_Q^{(h)}}{n_Q^{(ex)}} e^{-Ry/k_B T} \quad (7)$$

where n_e is the density of free electrons (equal to the number of free holes), n_{ex} is the density of excitons, and the other factors are the effective density of state factors, defined by

$$n_Q^{(i)} = \int_0^\infty D_i(E) e^{-E/k_B T} dE, \quad (8)$$

where $D_i(E)$ is the density of states for species i , which depends in general on its band mass m_i and spin degeneracy. The mass-action equation follows straightforwardly from the equilibrium condition that the chemical potentials for the different species in a chemical reaction must balance, that is, $\mu_e + \mu_h = \mu_{ex}$, and from the Maxwellian (low-density) formula for the total density,

$$n_i = e^{-(E_i - \mu)/k_B T} n_Q^{(i)}, \quad (9)$$

where E_i is the ground state energy of species i , and from the definition of the binding energy $Ry = (E_c - E_v) - E_{ex}$. Number conservation requires that $n_{ex} + n_e = n$, where n is the total pair density at any point in time.

To find the critical density, we set the *free* carrier density n_e to the value (6) above, instead of the total density. Above this density, screening will make exciton formation impossible. Assuming Ry is constant and $n_e \ll n_{ex}$, so that $n \approx n_{ex}$, this means the critical density for the transition is approximately

$$n = \left(\frac{\epsilon k_B T}{a^2 e^2} \right)^2 \frac{e^{Ry/k_B T}}{n_Q}, \quad (10)$$

where $n_Q = n_Q^{(e)} n_Q^{(h)} / n_Q^{(ex)}$. This critical density, for starting at low density and going up, clearly gives a very different threshold than (6) above, for starting at high density and going down. For example, at low temperature, very few excitons will ionize and therefore the critical density becomes exponentially large. Of course, at some point the Fermi-level Mott transition density (1) will be reached. In general, though, the density predicted by (10) is well below the density predicted by (1).

The difference between (6) and (10) depending on whether the system is starting with excitons or neutral plasma leads to the possibility of *hysteresis*, that is, the possibility that the

ionized fraction of excitons can depend not only on the total excitation density and temperature of the carrier gas, but also on its history. This possibility was first presented in Ref. [5] and has been recently revived by Portnoi [6] for two-dimensional systems; experimental evidence has been reported in Ref. [7].

As seen in this short survey, talking of the Mott transition density is nontrivial. The conductor–insulator transition in a carrier gas depends on density, temperature, the depth of the bound states, and the history of the system.

2. Self-consistent calculations for a three-dimensional system

In this section we ignore the possibility of hysteresis and ask for the prediction of the mass-action equation for the fraction of ionized carriers as a function of density and temperature. If we are to be self-consistent, we must account for the fact that the binding energy used in the mass-action Eq. (7) can depend on density through the density-dependent screening length. A very useful approximation, based on numerical calculations [8], is

$$\text{Ry}(l) = \begin{cases} \text{Ry}(0) \left(1 - \frac{2}{1 + l/a}\right), & l/a > 1 \\ 0, & l/a \leq 1, \end{cases} \quad (11)$$

where l is the screening length. Using this formula, and Eq. (5) for the screening length, we can rewrite (7) using number conservation as

$$n_e^2 = (n - n_e)n_Q e^{-\text{Ry}(n_e, T)/k_B T}, \quad (12)$$

which can be solved for n_e .

It turns out that we must adjust this equation slightly. The mass-action Eq. (7) breaks down in the limit $\text{Ry} \rightarrow 0$, which is exactly the limit we are interested in for the ionization catastrophe. This is easily seen by solving for n_{ex} when $\text{Ry} = 0$, which gives $n_{ex} = n_e^2/n_Q$, which is not zero. We expect, however, that the number of excitons vanishes when the exciton binding energy is zero. The reason for the breakdown is that exciton states with extremely shallow binding are still counted as exciton states, when in reality such states would be unstable, with very large orbits, and electrons and holes in such orbits would be indistinguishable in practice from free electrons and holes. As a first-order correction, one can subtract from n_{ex} the fraction of excitons which are in states with energy greater than Ry , namely a fraction equal to $e^{-\text{Ry}/k_B T}$, in which case the mass-action equation becomes

$$\frac{n_e n_h}{n_{ex}} = \frac{n_Q^{(e)} n_Q^{(h)} e^{-\text{Ry}/k_B T}}{n_Q^{(ex)} (1 - e^{-\text{Ry}/k_B T})} = \frac{n_Q}{e^{\text{Ry}/k_B T} - 1}, \quad (13)$$

which has the expected behavior that $n_{ex} \rightarrow 0$ when $\text{Ry} \rightarrow 0$. Much theoretical effort has been given, especially in the astrophysics community, to justifying this approach more rigorously, for example using Planck–Larkin partition functions, which give higher-order corrections [9,10].

Using this adjusted mass-action equation, we then have

$$n_e^2 + n_e \left(\frac{n_Q}{e^{\text{Ry}(n_e, T)/k_B T} - 1} \right) - \frac{nn_Q}{e^{\text{Ry}(n_e, T)/k_B T} - 1} = 0. \quad (14)$$

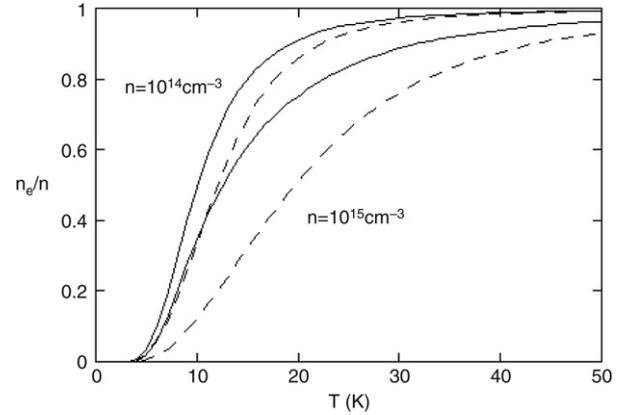


Fig. 1. Solid lines: fraction of free carriers in bulk GaAs predicted by Eq. (14) in the text, as a function of temperature for the two total pair densities labeled. Parameters used are $m_e = 0.07m_0$, $m_h = 0.08m_0$, $\epsilon/\epsilon_0 = 12$, and $\text{Ry}(0) = 4$ meV, which implies $a = 150$ Å. Dashed lines: the prediction for constant excitonic Rydberg $\text{Ry} = 4$ meV, for the same two densities.

If $\text{Ry}(n_e, T)$ is taken as a constant, then this equation can be solved simply as a quadratic equation. This leads to a value for n_e which can then be used to recompute the value of $\text{Ry}(n_e, T)$. This leads to an iterative process to converge on a self-consistent value of n_e . This iteration converges very rapidly, typically in less than 10 iterations. These formulas can be simply programmed in a conditional loop to generate the curves shown in this paper.

Fig. 1 shows the fraction of free carriers in bulk GaAs, with the expected behavior that the fraction increases toward 100% as temperature is increased and the excitons ionize. For simplicity we assume that all holes have the light-hole mass; the presence of heavy holes will give a correction but not a substantial change to the forms of the curves. The dashed curves show the prediction of the mass-action equation for the same parameters, but constant excitonic Rydberg. As expected, the self-consistent Rydberg solution gives a small correction at low density, and a larger correction at higher density. The self-consistent prediction becomes the same as the constant-Rydberg prediction at low temperature; for the parameters used in Fig. 1 the two predictions are identical below 4 K.

We can also plot the ionized fraction as a function of density at fixed temperature. Fig. 2 shows the fraction of ionized carriers as a function of density for various temperatures predicted for bulk GaAs. At low temperature, the insulator–conductor transition is quite sharp, while at higher temperature it is less so, but at all temperatures there is a “catastrophe” at which n_e/n becomes strictly equal to unity. Below $T = 4$ K, the critical density predicted by these equations becomes greater than that of the phase space filling Mott transition density, $n \simeq 1/a^3 = 3 \times 10^{17} \text{ cm}^{-3}$. Surprisingly, the mass-action predicts that at extremely low densities, the fraction of free carriers approach 100% at all temperatures. This is because the rate of formation of excitons is proportional to n_e^2 , and at low density the electrons and holes simply cannot find each other to form excitons.

Fig. 3 shows the predicted phase boundary. At low temperatures, this is a sharp conductor–insulator phase

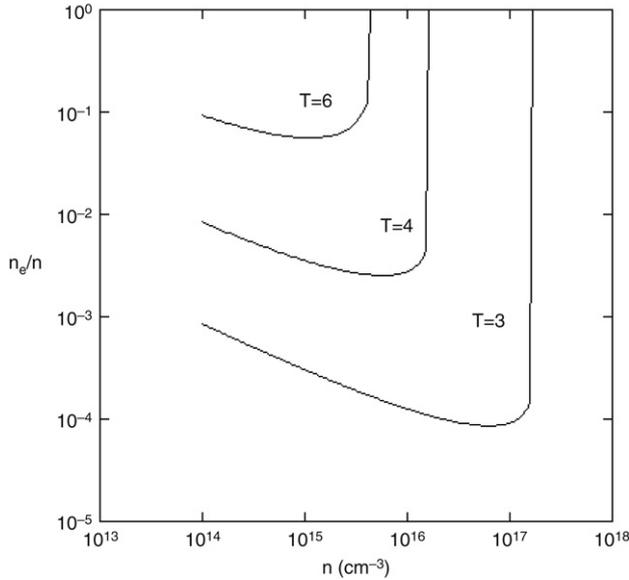


Fig. 2. Fraction of free carriers in bulk GaAs predicted by Eq. (14) in the text, as a function of density for the four temperatures shown. Parameters for bulk GaAs are the same as for Fig. 1.

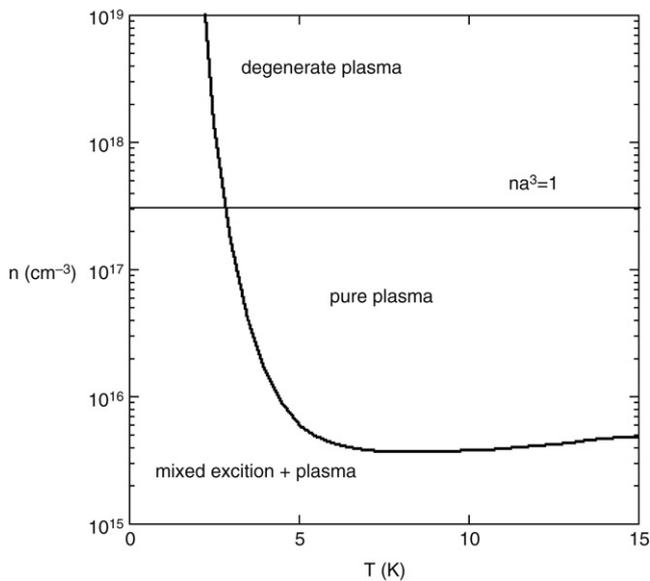


Fig. 3. Predicted phase diagram of carriers in bulk GaAs. Heavy line: the critical threshold for 100% ionization implied by Eq. (14) in the text. Above the boundary, the system is a pure plasma. Light line: the phase space filling condition $n = 1/a^3 \sim 3 \times 10^{17} \text{ cm}^{-3}$. Parameters for bulk GaAs are the same as for Figs. 1 and 2.

boundary. As seen in Fig. 1, above around 6 K a substantial fraction of free carriers coexist with excitons even below the ionization catastrophe boundary. The effect of the transition will be much more dramatic on the luminescence spectrum, however, since excitons dominate the luminescence due to their stronger oscillator strength. At the ionization catastrophe, the exciton line will disappear.

3. Calculations for a two-dimensional system

We can use the same approach for a two-dimensional system, e.g. carriers in quantum wells. In this case the screening

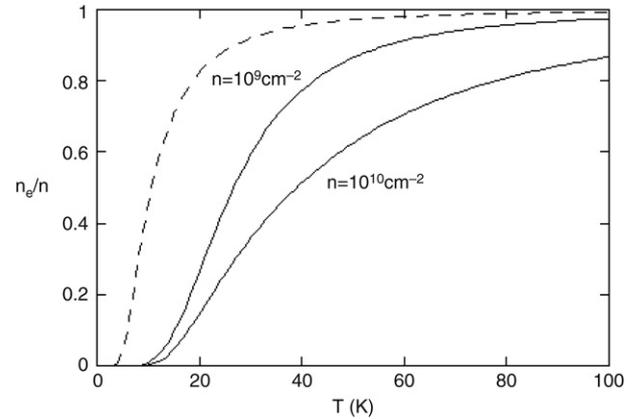


Fig. 4. Fraction of free carriers in single GaAs quantum wells predicted by Eq. (14) in the text, as a function of temperature for the densities shown. Solid lines: parameters for GaAs are the same as for Fig. 1, but $Ry = 10 \text{ meV}$, $a = 120 \text{ \AA}$, typical values for an exciton in a narrow single quantum well. Dashed line: the same parameters, but $Ry = 4 \text{ meV}$, for the lower density.

length is given by [11]

$$\frac{1}{l} = \frac{e^2 n}{2\epsilon k_B T}. \quad (15)$$

(Note that the Coulomb interaction is still assumed to be three-dimensional, but the carriers are constrained to move only in two dimensions.) The two-dimensional exciton Bohr radius is related to the two-dimensional excitonic Rydberg by [12]

$$Ry = \frac{e^2}{4\pi\epsilon a}. \quad (16)$$

In addition, we must use two-dimensional densities of states. A numerical solution for the variation of the Rydberg as screening increases in two dimensions has been performed by Portnoi and Galbraith [1], which can be approximated as

$$Ry(l) = \begin{cases} Ry(0) \left(1 - \frac{2}{1 + \sqrt{2l/a}} \right), & l/a > \frac{1}{2} \\ 0, & l/a \leq \frac{1}{2}, \end{cases} \quad (17)$$

where a is the two-dimensional excitonic Bohr radius.

Fig. 4 shows the fraction of free carriers as a function of temperature for the parameters of narrow GaAs quantum wells, with $Ry = 10 \text{ meV}$, for two densities. In this case only one in-plane mass is relevant at low temperature. This figure also shows a curve for the same parameters but $Ry = 4 \text{ meV}$, with the same exciton Bohr radius, which are realistic parameters for indirect excitons in double quantum wells, which have been used in recent experiments [13,14]. It is perhaps surprising that excitons exist all the way up to 100 K or above at these densities. This is a simple consequence of the prediction of the mass-action equation (13), which in the limit of low density and high temperature is

$$\frac{n_{ex}}{n} = \frac{n}{n_Q} \frac{\Delta}{k_B T}. \quad (18)$$

Although higher temperature favors ionization, the free carrier fraction is never 100% unless there is an ionization catastrophe.

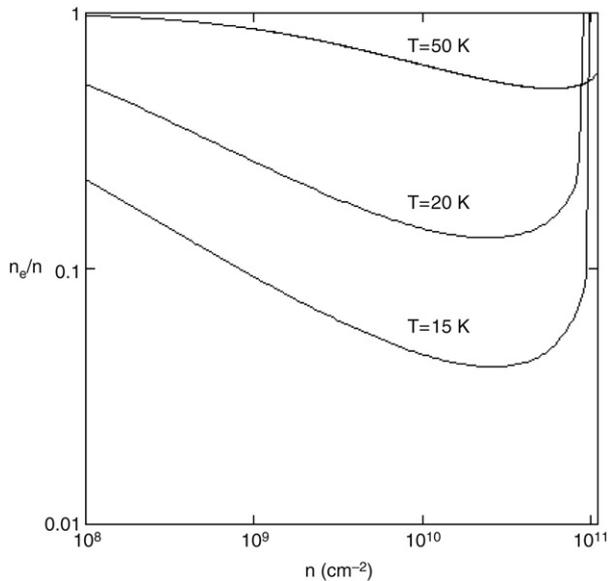


Fig. 5. Fraction of free carriers in single GaAs quantum wells predicted by Eq. (14) in the text, as a function of density for the temperatures shown. Parameters for GaAs are the same as for Fig. 4.

The fraction of the gas which is excitons is very low at these high temperatures, (e.g. about 1% at $T = 90$ K for 4 meV binding energy), but since they have a strong oscillator strength they can still have a strong emission in the luminescence in many semiconductors.

Fig. 5 shows the fraction of free carriers as a function of total pair density for several temperatures, for the case of single quantum wells. A sharp transition occurs at all temperatures, above which n_e/n is exactly equal to unity. The behavior is similar to that of bulk, but with the temperature scale shifted to higher temperature and density. Fig. 6 shows the predicted phase diagram, for the two different values of exciton binding energy discussed above.

4. Conclusions

The self-consistent mass-action equation presented here allows a relatively simple way of calculating the fraction of excitons in equilibrium in a population of excited carriers, and also the critical density for the thermodynamic Mott transition, or ionization catastrophe, in which all excitons are screened out.

Although the phase boundaries for the thermodynamic Mott transition are not simple, their form allows us to speak of two critical thresholds. At moderate densities (within an order of magnitude or so of the phase space filling density $n = 1/a^d$, where d is the dimension), one can speak of a critical temperature for the Mott transition, since the phase boundary corresponds to a roughly constant temperature in this range. On the other hand, at high temperature, one can speak of a critical density for the Mott transition, since the phase boundary corresponds to a nearly constant density in this region. This critical density is typically an order of magnitude or more below the phase space filling density $n = 1/a^d$ in both two and three dimensions.

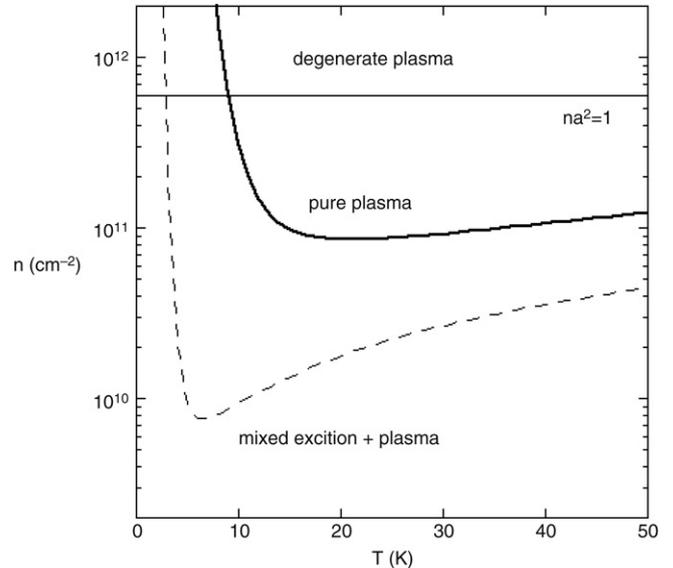


Fig. 6. Predicted phase diagram of carriers in GaAs quantum wells, for two different exciton binding energies. (Heavy solid line: $Ry = 10$ meV; dashed line: $Ry = 4$ meV.) Heavy line: the critical threshold for 100% ionization according to the model discussed in the text. Above the boundary, the system is a pure plasma. Light line: the phase space filling condition $n = 1/a^2$.

At moderate densities, the critical temperature for the Mott transition is much less than Ry/k_B . On the other hand, at low densities, excitons persist as a small fraction of the gas to temperatures well above Ry/k_B .

Acknowledgment

This work has been supported by the Department of Energy under grant DE-GF02-99ER45780.

References

- [1] M.E. Portnoi, I. Galbraith, Phys. Rev. B 60 (1999) 5570.
- [2] S.W. Koch, W. Hoyer, M. Kira, V.S. Filinov, Phys. Status Solidi C 238 (2003) 404.
- [3] T.M. Rice, in: H. Ehrenreich, F. Seitz, D. Turnbull (Eds.), Solid State Physics: Advances in Research and Applications, vol. 32, Academic, 1977.
- [4] B.K. Ridley, Quantum Processes in Semiconductors, 2nd edition, Oxford, 1988.
- [5] D.W. Snoko, J.D. Crawford, Phys. Rev. E 52 (1995) 5796.
- [6] V.V. Nikolaev, M.E. Portnoi, A.V. Kavokin, Proc. 27th ICPS, AIP Conference Proceedings 772 (2005) 1156.
- [7] P. Nemeč, P. Maly, M. Niki, K. Nitsch, Appl. Phys. Lett. 76 (2000) 2850.
- [8] C.R. Smith, Phys. Rev. 134 (1964) A1235.
- [9] Ebeling, in: W. Ebeling, W.-D. Kraft, D. Kremp (Eds.), Theory of Bound States and Ionization Equilibria in Plasmas and Solids, Akademie-Verlag, Berlin, 1976.
- [10] W.-D. Kraft, D. Kremp, W. Ebeling, G. Röpke, Quantum Statistics of Charged Particle Systems, Plenum, 1986.
- [11] K. El-Sayed, T. Wicht, H. Haug, L. Banyai, Z. Phys. B 86 (1992) 345.
- [12] H. Haug, S.W. Koch, Quantum Theory of the Optical and Electronic Properties of Semiconductors, 3rd edition, World, 1990.
- [13] D. Snoko, S. Denev, Y. Liu, L. Pfeiffer, K. West, Solid State Comm. 127 (2003) 187.
- [14] M.H. Szymanska, P.B. Littlewood, Phys. Rev. B 67 (2003) 193305.