Bose–Einstein condensation of excitons in Cu$_2$O: progress over 30 years

David Snoke$^1$ and G M Kavoulakis$^2$

$^1$ Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15260
$^2$ Technological Educational Institute of Crete, PO Box 1939, GR-71004, Heraklion, Greece

Received 20 October 2011, revised 3 February 2014
Accepted for publication 28 July 2014
Published 24 October 2014

Invited by G Baym

Abstract
Experiments on Bose–Einstein condensation (BEC) of excitons in the semiconductor Cu$_2$O started over 30 years ago, as one of the first serious attempts at exciton BEC. Early claims were based on spectroscopic signatures and transport data which have since been reinterpreted, in large part because the Auger recombination process for excitons was not well understood. Understanding of the Auger process has advanced, and recent experiments have made significant progress toward exciton BEC. We review the history of experiments on exciton BEC in Cu$_2$O, the Auger recombination process, and the prospects for observing exciton BEC in this system in the near future.

Keywords: BEC, excitons, Bose–Einstein condensation

(1. Introduction

BEC of excitons. Bose–Einstein condensation (BEC) of excitons was predicted in the early 1960s, by Moskalenko [1], Blatt et al [2], and Casella [3], in independent efforts. The concept is simple: just as two electrons with half-integer spin can be paired to make a Cooper pair which acts as an integer-spin boson, and just as an even number of fermionic electrons, protons, and neutrons are bound together to make up every bosonic atom, so too an electron and hole in a semiconductor, each with half-integer spin, can be bound together in the complex known as an exciton, which will also act as a boson. Therefore one can expect BEC of excitons under two conditions: (1) when the number of excitons is approximately conserved; in particular, their lifetime is much longer than the time it takes for them to thermalize to a well-defined temperature, and (2) when the density of the excitons is not too high, because competing phases occur at high density such as electron–hole plasma (EHP), in which the binding of the electrons and holes into excitons breaks down.

Under these conditions, it is straightforward to describe the exciton gas by weakly interacting Bose gas theory. The general condition for quantum coherent effects to be important is that the thermal deBroglie wavelength $\lambda_{\text{DB}}$ of the particles be comparable to or larger than the average distance $r_s$ between the particles. The order of magnitude of the deBroglie wavelength is found by equating

$$\frac{\hbar^2 k^2}{2m} = \frac{\hbar^2 (2\pi)^2}{2m\lambda_{\text{DB}}^2} \sim k_B T,$$

which implies

$$\lambda_{\text{DB}} \sim \frac{2\pi \hbar}{\sqrt{2mk_B T}}.$$

In three dimensions, the average distance $r_s$ scales as $r_s \sim n^{-1/3}$, where $n$ is the density of the particles. Setting $\lambda_{\text{DB}} \sim r_s$ gives

$$n \sim \frac{2^{3/2}}{(2\pi)^{3/2}} \frac{(mk_B T)^{3/2}}{\hbar^3},$$

or

$$T \sim \frac{(2\pi \hbar)^2}{2mk_B n^{2/3}}.$$

In equilibrium, the standard calculation of statistical mechanics [4] gives the critical density for BEC,

$$T_c = 0.17 \frac{(2\pi \hbar)^2}{2mk_B n^{2/3}}.$$.  


0034-4885/14/116501+17$88.00 1 © 2014 IOP Publishing Ltd Printed in the UK
Figure 1. Generic phase diagram of excitons in a three-dimensional bulk semiconductor. FE = free exciton gas. BEC = exciton BEC. BCS (EI) = excitonic insulator state, which is a BCS state; there is a BEC-BCS crossover at low temperature as density increases. EHP = electron–hole plasma, i.e., ionized excitons. There are two regimes for this, a non-degenerate and a degenerate plasma. The general condition for degeneracy is \( r_i \sim a \), that is, the average distance between the particles be comparable to the exciton Bohr radius.

Even away from perfect equilibrium, however, the relations (3) and (4) will still apply as the conditions for Bose–Einstein statistical effects, which include a peaking of the distribution of the particles near the ground state, also known as a ‘quasicondensate’.

Note that these relations give a critical temperature inversely proportional to the mass \( m \). This means that light mass implies Bose–Einstein effects at higher temperature. Since excitons have mass of the order of two electron masses, this led to the proposal that exciton condensates could be the first room temperature condensate. The excitons in many semiconductors are stable at room temperature, e.g., Cu₂O, CuCl, CdSe, ZnO, and GaN. To have a Bose–Einstein condensate, one must ensure that the exciton lifetime is also long enough to allow thermalization, as discussed above. Several of these semiconductors seemed to satisfy this requirement, and experiments on them began in the 1970s, as discussed below.

The relations (3) and (4) also imply that the critical temperature for Bose–Einstein effects depends on the density. In liquid helium and in BCS superconductors, the density is essentially a fixed quantity. By contrast, it is quite easy to vary the density of excitons to very low density by simply changing the intensity of a light source which generates excitons. This can be done, for example, by changing the intensity of a light source which generates excitons to very low density by simply changing the intensity of a light source which generates excitons.

**Competing phases.** There are three possible competing phases for electrons and holes besides the exciton Bose–Einstein condensate (EBEC). One of these is the ‘excitonic insulator’ (EI) state which is a Bardeen–Cooper–Schrieffer (BCS) state of the electrons and holes instead of two spin states of the electrons. This state arises under the general condition \( r_i < a \), where \( a \) is the exciton Bohr radius, i.e., the pair correlation length. This is equivalent to the condition \( na^3 > 1 \). This condition corresponds to the case that a Fermi level of the electrons (and also of the holes, assuming they have comparable mass) exists in the system which is large compared to the intrinsic binding energy of the excitons. We can see this by noting first that the exciton binding energy \( R_{\text{Ex}} \) and the excitonic Bohr radius \( a \) are related by the formula (in MKS units)

\[
R_{\text{Ex}} = \frac{e^2}{8\pi e a},
\]

which is the same as for a hydrogen atom except that the electric force has been renormalized by the screening in the medium, \( e^2 \rightarrow e^2/\varepsilon \), where \( \varepsilon \) is the dielectric constant of the medium. The hydrogenic formula for \( a \) is

\[
a = \frac{4\pi e^2 h^2}{e^2 m}.
\]

On the other hand, the Fermi energy of an electron gas in the low-temperature limit is given by [4]

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

where \( n \) is the density. Setting this larger than the binding energy gives

\[
\frac{3\pi^2 h^3}{\sqrt{2m^3/2}n} > \left( \frac{e^2}{8\pi e a} \right)^{3/2}
\]

\[
\sim \frac{4\pi e^2 h^2}{e^2 m}^{3/2} n > \frac{1}{a^3},
\]

which is equivalent to \( na^3 > 1 \), using (7) for the excitonic Bohr radius.

In the EI limit the gas still is coherent, like a BCS superconductor, below a critical temperature. There should be a crossover from exciton BEC to the EI state at low temperature, as density is increased. Indeed, some of the earliest work on the theory of BEC-BCS crossover was done in the context of thinking about exciton systems, notably foundational theory by Keldysh and Kozlov [5, 6], Hanamura and Haug [7–9], and Comte and Nozières [10, 11]; for a general review of the early theory of excitonic condensates, see [12]. These works showed that excitonic condensation in the low-density BEC limit was a sound concept, despite the fermionic nature of the underlying electrons and holes, and there is a smooth transition to the BCS-like EI phase at high density. The critical temperature for this high-density EI phase will decrease with increasing density, however, so that raising the density is no longer advantageous. In practical terms, this means that the density of the particles should be kept below \( 1/a^3 \). Excitonic condensation can be expected at room temperature only if \( n \) is small enough that this limit can be satisfied even for densities given by (3) at room temperature. The condition that the
critical density at room temperature be less than \(1/a^3\) turns out not to be an issue for most of the semiconductors listed above with excitons which exist at room temperature, since large binding energy of the excitons also corresponds to small Bohr radius.

Another competing phase is ionization to an incoherent EHP. This is analogous to the classical ionization of an atomic gas at high temperature and high density due to three-body collisions; it is sometimes called the excitonic ‘Mott’ transition because it is a conductor–insulator transition, although the mechanism is quite different from the Mott transition in cold-atom lattices. The theory of this transition is actually quite complicated (for a review, see [13]; see also recent work by Manzke et al [14, 15]). Self-consistent theory of this transition involves the dynamic screening of the electron–hole interaction, which in turn depends on the number of ionized electrons and holes. Figure 1 illustrates the basic shape of the boundary for this phase. The position of the phase boundary depends on the binding energy of the excitons; in general we can say that higher binding energy pushes this curve to higher temperature. Therefore having deeply bound excitons also helps avoid this competing phase.

Another competing phase is the electron–hole liquid (EHL), sometimes also called ‘condensation’. This state received a significant amount of attention in the 1970s and 1980s [16, 17] following its observation in two bulk semiconductors, Ge and Si. This state is analogous to liquid mercury: the electrons and holes generated by optical excitation are not bound as excitons or biexcitons (bound states of two excitons) but instead form two interpenetrating Fermi gases with the properties of a classical liquid, with a surface tension. The EHL state is a conductor, while the EBEC state is an insulator, because excitons are charge neutral. If the EHL state does exist, it will prevent EBEC at any temperature and density, since its phase boundary also scales as \(T \sim n^{2/3}\) in three dimensions.

The EHL state only occurs in indirect-gap semiconductors like Si and Ge with multiple degenerate valleys in the conduction band. This degeneracy allows the density of the excited electrons to be high while still keeping the average kinetic energy of the carriers low. This means that the mutual attraction of the electrons and holes overcomes the kinetic energy cost of the Fermi level of free carriers, leading to a net free energy savings to enter the EHL state at high excited carrier density [17]. In a nice experiment, Timofeev et al [18] used a stress geometry to lift this degeneracy in a germanium crystal, and magnetic field to prevent bie exciton formation, and found that the EHL state did not occur, leaving only free excitons. At high density they saw evidence for Bose–Einstein statistics of the excitons (see [12], section 1.4.2), although not for EBEC.

As illustrated in figure 1, it is expected that at low enough temperature and density, excitons can undergo BEC, and if the excitons have large binding energy and small Bohr radius this could even occur at room temperature. Therefore experimental research on BEC of excitons began in the 1970s with a focus on deeply bound excitons in bulk semiconductors. As we will see, however, the second competing effect mentioned above, namely density-dependent recombination, has turned out to be a major issue.

**Deeply bound excitons and \(\text{Cu}_2\text{O}\).** The earliest experiments on excitonic condensation in bulk semiconductors were done with the bulk semiconductors CdSe and CuCl [19–22]. The most complete study on CuCl was done in the early 1980s [23, 24]. In CuCl, there is a tightly bound bie exciton state (i.e., an excitonic molecule like \(\text{H}_2\)), which, of course, is also an integer-spin boson which can in principle undergo BEC. The bie exciton binding energy in CuCl is 26 meV, which is comparable to \(k_BT\) at room temperature. The experiments with CuCl ran into interpretive difficulties, however, because CuCl has a strong polariton effect, in which photons in the medium couple strongly to the exciton or bie exciton states, leading to a very short radiative lifetime for low-momentum excitons. At high density, the bie excitons in CuCl showed a distinct peak in low-energy states, but were far from equilibrium. The enhancement of the population in low-energy states was probably related to the boson nature of the particles through the process of stimulated scattering, which drives the onset of BEC in all bosonic gases [25, 26]. In recent years, several sophisticated models [27–29] have been developed which describe the nonequilibrium behavior of polaritonic gases and show clearly the role of the bosonic stimulated scattering. It would be quite interesting to apply these modern theoretical tools to the early CuCl experiments to estimate to what degree the occupation number of the bie excitons was enhanced by their bosonic nature.

Around the same time, the semiconductor \(\text{Cu}_2\text{O}\) was identified as a good candidate for excitonic BEC. The excitons in \(\text{Cu}_2\text{O}\) have binding energy of 150 meV [30]; the bie exciton state is either nonexistent or only very weakly bound, on the order of 1 meV or less [31, 32], so that one can generally ignore the existence of bie excitons (although the existence of bie exciton states has been invoked to explain nonradiative recombination in \(\text{Cu}_2\text{O}\), as discussed below).

The lowest conduction band in \(\text{Cu}_2\text{O}\) comes from the 4s-orbitals of the oxygen atoms, while the highest valence band comes from the 3d-orbitals of the copper atoms. The crystal field in the cubic lattice splits the five d-orbitals into a triplet and a doublet, with the triplet higher; when the electron spin is taken into account, the spin-orbit interaction splits this triplet into a higher twofold-degenerate state and a lower fourfold-degenerate state. The lowest-energy excitons are formed from holes in these two highest states of the valence band and electrons in the two s-like states of the lowest conduction band. The electron–hole exchange splits these four states into a triplet and a singlet, which are called the orthoexciton and the paraexciton, respectively. These are split by 12 meV, with the paraexciton the lower state. In general, these two exciton states, which are the lowest energy exciton states, are the only important ones unless higher-lying exciton states are being excited directly.

The symmetry group of unstressed \(\text{Cu}_2\text{O}\) is \(\text{O}_h\), which is cubic with inversion symmetry. This high symmetry leads to the effect that the lowest exciton state in \(\text{Cu}_2\text{O}\), the singlet paraexciton, has zero oscillator strength for interaction with photons, due to the selection rules which arise from the symmetry of the crystal [30, 33]. This implies that

---

3 See [12], Appendix A.
there is no polariton effect for the paraexcitons. The triplet orthoexciton state, by contrast, has a symmetry-allowed radiative recombination process. (The singlet-triplet terminology of Cu$_2$O sometimes confuses those who work with organic light emitters, in which the state with forbidden recombination is the triplet state and the state with allowed radiative recombination is the singlet state. The difference in the terminology is because in the case of organics, the transition is assumed to have an $s$–$p$ dipole-allowed matrix element, and the spin singlet state corresponds to no spin flip of the excited electron, and thus an allowed process, while the spin triplet state corresponds to a spin flip. In the case of Cu$_2$O, neither the orthoexciton nor the paraexciton is a pure spin state; the triplet orthoexciton state is actually a superposition of singlet and triplet pure electron spin states, while the singlet paraexciton is a superposition of only pure spin triplet states.)

In principle, the fact that the lowest exciton state has a forbidden transition means that these excitons have infinite lifetime, but a weak phonon-assisted recombination process is allowed [34], and also recombination at impurities can occur, so that the lifetime of paraexcitons in Cu$_2$O is typically of the order of hundreds of nanoseconds to microseconds [35–37]. This is very long compared to excitons in most semiconductors, and very long compared to the thermalization times of the excitons via exciton–phonon interaction, which is tens of picoseconds [38]. Because there has never been a large technological effort to produce pure Cu$_2$O, most of the purest samples come from rock collectors who obtain them from copper mines.

A very important step was accomplished in the 1980s in the trapping of excitons in Cu$_2$O. Inhomogeneous stress was used to create a harmonic potential minimum for the excitons inside the crystal [39, 40] very similar in geometry to the magneto-optical traps used to hold cold atoms for BEC experiments. As discussed below, some early experiments on EBEC used surface excitation of the crystal, which had the advantage of achieving high carrier density due to the short absorption length of the excitation light, but generated a highly nonequilibrium situation with an expanding gas that was difficult to analyze. The stress used in the exciton traps has the side effect that it changes the symmetry of the crystal, so that the paraexciton state no longer has a forbidden radiative recombination transition [41], but the lifetime of the paraexcitons in a stressed crystal is still very long compared to the exciton–phonon interaction time.

**Phonon-assisted luminescence spectroscopy.** The phonon-assisted recombination process in Cu$_2$O provides nice data to indicate the behavior of the excitons and whether they undergo BEC. Figure 2 shows the phonon-assisted recombination process of excitons in a direct-gap semiconductor like Cu$_2$O. Only excitons with low momentum can recombine via the direct recombination process, while an exciton at any momentum can recombine via the phonon-assisted process, with the phonon taking up any excess momentum. The phonon-assisted process can occur for both the paraexciton and the orthoexciton state. The energy of the emitted photon is equal to the total energy of the exciton minus the energy of the optical phonon, which is nearly constant. The energy spectrum of the phonon-assisted luminescence therefore gives the kinetic energy distribution of the excitons directly. If we take the matrix element as nearly independent of the exciton momentum, which is the case in Cu$_2$O, then the intensity of the light emitted at a given energy is directly proportional to the number of excitons with the corresponding kinetic energy $\hbar^2 K^2/2m$.

Figure 3 shows typical spectra of phonon-assisted luminescence from orthoexcitons in Cu$_2$O at various temperatures up to room temperature. At high temperature, an optical phonon can be both emitted or absorbed when an exciton recombinates. Therefore there are two lines, one at the exciton energy minus the optical phonon energy, and one at the exciton energy plus the optical phonon energy. The theory fits to each peak in the data in figure 3 are given simply by $I(E+E_0) \propto D(E)f(E)$, where $E$ is the exciton kinetic energy, $D(E)$ is the density of states of the excitons, proportional to $E^{-1/2}$ in three dimensions, and $f(E)$ is the occupation number of the excitons, proportional to $e^{-E/kT}$ at low density. The relative height of the two lines depends on the temperature of the lattice. At low temperature, the phonon-emission line dominates, since there are no phonons to absorb, while at high temperature, the two lines become comparable.

When many-body interactions are taken into account, the analysis of the phonon-assisted luminescence is slightly more complicated. Each momentum state emits with an energy profile given by the spectral function $A(k,\omega)$, which is approximately equal to $\delta(\omega - \omega_g)$ at low density and low temperature, but can be substantially broadened when there are strong particle–particle interactions [42–44]. This brings up a crucial difference between exciton BEC experiments and experiments on cold atoms. The fact that excitons can recombine and turn into photons allows us to have direct access to the spectral function. (This property has been used widely in the exciton-polariton BEC experiments, e.g. to show spectral narrowing due to coherence [45, 46].) In cold atom experiments, the spectral function is a theoretical construct that is never directly observed.

![Figure 2. Recombination processes of Wannier excitons in a direct band-gap semiconductor.](image)
Direct access to the spectral function means that we can see absolute shifts of the ground-state energy of the excitons directly as energy shifts of the emitted photons, and that we can see the effect of broadening of the spectral function as spectral broadening of the photon emission. To lowest order, the former can be identified as the real part of the exciton self-energy, and the latter as the imaginary part of the self-energy. (See [42], section 8.4). The exciton–phonon interaction increases as the temperature increases, leading to a red shift of the luminescence lines in figure 3 due to renormalization of the real self-energy as well as broadening due to increasing imaginary self-energy. The imaginary self-energy is, to first order, just given by $\Gamma = \hbar/\tau$, where $\tau$ is the average scattering time of the excitons. This scattering time is an average of all scattering processes, including exciton–phonon scattering, which increases as temperature increases, exciton-impurity scattering, which also increases with increasing temperature, and exciton–exciton scattering.

**Exciton–exciton scattering.** This line broadening effect was used in the first attempt to estimate the absolute density of the excitons independently from the assumed statistics of the excitons in the Cu$_2$O experiments [47], as discussed below. Given an estimate of the exciton–exciton scattering cross section, the scattering time $\tau$ for a given density can be calculated for the case when the lattice is cold, near 1 K, so that exciton–phonon interactions are negligible, and this can be compared to the observed broadening as a function of exciton density.

The cross section for exciton–exciton scattering is actually a problematic theoretical topic for exciton physics, as compared to cold atom physics, where atom–atom cross sections are calculated to three decimal places or better. The difficulty of calculating the exciton–exciton scattering cross section has nothing to do with details of the band structures of the crystals, which are well understood. The primary difficulty is the fact that the electron and hole have comparable mass, so that all the following exchange effects are important and cannot be neglected: electron–electron exchange, hole–hole exchange, and electron–hole exchange. In atoms, there is no electron–nucleus exchange, and nucleon–nucleon exchange can typically be ignored (although M. Combescot has argued that under the conditions of BEC the nucleon–nucleon exchange may be important, since the wavelength, by definition, is comparable to the distance between the particles; see [48] and references therein for the general approach of accounting for these exchange effects).

The most accurate calculations for the exciton–exciton scattering cross section in Cu$_2$O have been done by Shumway and Ceperley [49], and give a cross section of 5 Å, which is comparable to the Bohr radius of the excitons. In that work, the spin of the electrons and holes was assumed to be aligned. The spin structure of the excitons can lead to significant differences in the scattering cross sections between different spin states [50], although these calculations have not been worked out with the same accuracy as the Ceperley and Shumway work.

2. Early experiments on exciton BEC in Cu$_2$O

2.1. Spectral analysis experiments

The first experiments to show promise for exciton BEC in Cu$_2$O used excitation with a green laser which was absorbed within 5 $\mu$m of the surface [51, 52]. This created a thin ‘pancake’ of excitons which were not trapped, and could flow away from the surface, but on time scales of a few nanoseconds remained at high density at the point of creation.

The phonon-assisted luminescence spectrum of the orthoexcitons (the upper, triplet state of the 1s excitons is...
in Cu$_2$O) could be well fit by the Bose–Einstein distribution

$$ I(E) \propto D(E)f(E) \propto E^{1/2}\frac{1}{e^{(E-\mu)/k_B T} - 1}. \quad (10) $$

with $\alpha = |\mu/k_B T| \sim 0.1$ over a wide range of densities. In other words, the gas remained near the phase boundary for BEC but did not cross it. This was termed ‘Bose saturation’. Figure 4 shows that this saturation corresponds to an energy scaling such that the shape of the photoluminescence remains the same. This scaling was found to apply over more than an order of magnitude of density variation.

The fits to the Bose–Einstein distribution imply an absolute density of the particles according to

$$ N = \int_0^\infty \frac{1}{e^{(E-\mu)/k_B T} - 1}D(E) \, dE. \quad (11) $$

The densities from the fits were compared to data for the relative change of the density, found by taking the total luminescence intensity and dividing the measured exciton cloud volume at each point in time, and the variation in the density deduced from the fits agreed with the variation in density from the intensity data within a factor of two over a wide range of density [51].

Later experiments looked at the paraexciton phonon-assisted luminescence under very similar experimental conditions and concluded that the paraexciton density exceeded the density for BEC predicted by the ideal gas theory [53]. Line shape analysis of the paraexciton phonon-assisted emission was harder, because the emission intensity of this line is 500 times weaker than the orthoexciton line, and it lies near another, brighter emission line. The relative intensity of the paraexciton line seemed consistent with the conclusion that the paraexcitons exceeded the critical density.

The first indication that this analysis could not be the whole story came from noting that the many-body collision broadening of the lines (discussed above) was very low, and not consistent with the densities implied by the BEC fits [47]. Although the exciton–exciton scattering cross section is not known to better than a factor of two, as discussed above, if the experiments were truly to have densities in the range of $10^{19} \text{ cm}^{-3}$, as implied by the fits of equation (10) to data like that shown in figure 4, along with the observed spectral broadening of less than 0.1 meV, then these two observations together would require a scattering cross section with radius much less than the Bohr radius of the excitons. The crucial evidence came in comparing the exciton luminescence spectrum in the case of a cold (2 K) lattice, when the exciton distribution fit a Bose–Einstein distribution at 70 K, to the luminescence spectrum with a 70 K lattice under the same excitation conditions [54]. The total number of photons emitted in each case was nearly the same, indicating that the densities were comparable, but the shapes of the spectra were quite different. The deviation from a Maxwell–Boltzmann distribution in the cold-lattice case could therefore not be from Bose statistics at high density if the same density gave a Maxwell–Boltzmann distribution when the system was closer to equilibrium with the lattice. This pointed to spatial inhomogeneity in a nonequilibrium system as the cause of the non-Maxwellian distribution.

The main ambiguity of these experiments comes from the fact that the luminescence from the exciton gas must be integrated over at least one dimension, since the photons are collected by imaging a three-dimensional crystal in two dimensions. To fit the Bose–Einstein distribution to the data, some model of the spatial variation of the density and temperature must be used. This was recognized in the early work [52], and it was found that nearly all models which assumed that the temperature and density of the orthoexcitons decreased monotonically in the direction away from the surface gave spatially integrated spectra that looked very similar to the simple homogeneous spectrum of equation (10). (The difficulty of interpreting the three-dimensional data integrated over one dimension was one motivation for pursuing experiments with excitons, or exciton–polaritons, in two dimensions; for reviews see, e.g., [55] and [56].)

Later modeling [57] found that a reasonable fit to the spatially-integrated orthoexciton luminescence spectra could be obtained by assuming a large cloud of cold orthoexcitons far from the surface (see figure 5). A simple analysis would seem to indicate that orthoexcitons do not travel very far within their lifetime due to down-conversion into paraexcitons [68–71], and therefore one should not expect so many orthoexcitons away from the surface. However, the model of [57] showed that picking a high enough rate for an exciton Auger recombination process could generate orthoexcitons at long distances from the surface due to up-conversion from paraexcitons.

**Auger recombination.** This brings up the Auger recombination process in Cu$_2$O, which was established to occur in Cu$_2$O in the 1980s [39] and became an increasingly important topic of study in the 1990s. The Auger process occurs when two excitons collide, and one of them recombines,
Figure 5. Fits of a theoretical model for nonequilibrium flow of excitons to the phonon-assisted photoluminescence spectrum of Cu$_2$O with high-intensity surface excitation for three excitation powers under conditions similar to those of figure 4, at two different times after a short laser pulse. Reproduced with permission from [57]. © 2000 The American Physical Society.

Figure 6. The Auger recombination process. In (a), the exciton on the left recombines, giving its energy to the exciton on the right, which leads to the final state (b) of a single ionized exciton.

but instead of emitting a photon, the energy of the recombining exciton is given to ionizing the second exciton, as shown in figure 6. The hot electron and hole thus produced can then lose energy by phonon emission or by collisions with other carriers, and finally form into an exciton again.

The Auger process is a well-known decay mechanism in the physics of semiconductors, and is also present in EHPs, in which the recombination of an electron and a hole excites either an electron high in the conduction band or a hole deep in the valence band. This process, which occurs in many semiconductors, has long been studied both theoretically [58–61] and experimentally [62, 63]. A similar decay process has also been observed in experiments with cold metastable He atoms, where the term ‘Penning collisions’ is used [64]. It also occurs with Frenkel excitons in organic materials, where it can be quite strong [65, 66]. There may, in fact, be a general principle that excitons which have large binding energy, and therefore are Frenkel-like, intrinsically will have strong Auger recombination. This is seen in the phonon-assisted Auger theory discussed below, which as we will see in equation (16), gives a rate proportional to $1/a^2$, where $a$ is the exciton Bohr radius.

The Auger recombination process is density dependent, since it is proportional to the rate of excitons colliding with each other. The Auger rate is therefore parametrized by a constant typically called $A$, such that the rate of the process is

$$\frac{1}{\tau} = An,$$  \hspace{1cm} (12)

where $n$ is the density. For a single population with an intrinsic radiation lifetime $\tau_r$, this gives the rate equation

$$\frac{dn}{dt} = G(t) - \frac{n}{\tau_r} - \frac{1}{2}An^2,$$ \hspace{1cm} (13)

where $G(t)$ is the generation rate of the excitons. A factor of exactly $1/2$ occurs because half the excitons are returned to the population after ionization. If we keep track of two sub-populations such as the orthoexcitons and paraexcitons in Cu$_2$O, we must write two coupled rate equations,

$$\frac{dn_o}{dt} = G_o(t) - \frac{n_o}{\tau_o} - An_o^2 - An_o n_p \left(1 - \frac{1}{2}A(n_o + n_p)^2\right)$$

$$\frac{dn_p}{dt} = G_p(t) - \frac{n_p}{\tau_p} - An_p^2 - An_p n_o + \frac{1}{2}(1 - \alpha)A(n_o + n_p)^2$$ \hspace{1cm} (14)

to take into account the collisions between different species.

The last term in each equation gives the effect of the return of the ionized electrons and holes; a new parameter $\alpha$ gives the fraction of these which return to each species. It is typically assumed that $\alpha = 3/4$, since the orthoexciton population has three degenerate states while the paraexciton is a singlet. Most generally, the constant $A$ can be different for collisions of different types of excitons, for example due to the effects of indistinguishability on the different spin states [67]. Writing down down three parameters, $A_{oo}$, $A_{op}$, and $A_{pp}$, the rate...
equations (14) become

\[
\frac{\partial n_o}{\partial t} = G_o(t) - \frac{n_o}{\tau_{oa}} - A_{oo} n_o^2 - A_{op} n_o n_p
\]

\[
+ \frac{1}{2} \alpha (A_{oo} n_o^2 + 2 A_{op} n_o n_p + A_{pp} n_p^2)
\]

\[
\frac{\partial n_p}{\partial t} = G_p(t) - \frac{n_p}{\tau_{pa}} - A_{pp} n_p^2 - A_{op} n_o n_p
\]

\[
+ \frac{1}{2} (1 - \alpha) (A_{oo} n_o^2 + 2 A_{op} n_o n_p + A_{pp} n_p^2).
\]

Additional modifications of these equations can be made to take into account direct conversion of one species into the other by phonon emission and absorption [68–71], and direct conversion by collision of two excitons, e.g., two orthoexcitons with opposite spin can collide and turn into two paraexcitons, conserving total angular momentum [72, 73].

The Auger process (a) reduces the lifetime of the excitons dramatically at high density, from the radiative lifetime of milliseconds down to less than a nanosecond, (b) heats the exciton gas due to the gap energy given to the ionized electron and hole, which then interact with the rest of the system, and (c) gives a population of orthoexcitons in excess of that expected from the Boltzmann occupation factor \( e^{-\Delta/k_B T} \), where \( \Delta \) is the paraexciton–orthoexciton splitting. Such an imbalance implies that the system is not in full equilibrium, which can occur when the lifetime due to Auger annihilation of the paraexcitons is shorter than the intrinsic ortho-to-para conversion time.

Throughout the 1990s and early 2000s, several groups tried to measure the rate of Auger recombination in Cu2O. This work is reviewed in section 3. The measurements are made difficult by the fact that there is also a two-body conversion process by which two orthoexcitons can collide and turn into two paraexcitons. Since this process has the same density dependence as the Auger process it can be difficult to distinguish the two processes in many experiments. In recent experiments, Kuwata-Gonokami’s group used infrared absorption measurements to excite the 1s–2p exciton transition [74–76], which is analogous to the 1s–2p transition in alkali atoms.

Since this transition only occurs when the excitons exist, the absolute absorption cross section of the transition gives an absolute measurement of the number of both species of excitons. Therefore the spin flip process, which conserves total number of excitons, could be distinguished from the Auger process, which annihilates excitons.

Although there has been a range of values found for the Auger rate, the Auger measurements generally imply that the density of the excitons could not be as high as predicted by the fits of the surface-excitation data to the Bose–Einstein distribution, because the lifetime would be too short at those densities. Since, as discussed above, there was also inconsistency in the broadening of the spectral lines and the absolute number of photons coming from the phonon-assisted luminescence, it was therefore generally concluded by the end of the 1990s that the early experiments did not show BEC. The Bose saturation effect, which corresponds to an invariance of the luminescence spectrum as temperature is scaled (see figure 4), is therefore surprising, because there are a number of details which go into the classical numerical model of which have to be fortuitously related to give the nearly invariant spectrum which is observed; without these coincidences the spectrum would have different shapes at different times. The model of [57] used choices of the physical parameters which were reasonable based on the experimental data, but even with the optimal choice of realistic parameters, the fit of the model to the data is qualitatively similar in shape but not especially good (see figure 5). The saturation effect corresponds to a scaling law, which we discuss below, in section 3.2.

2.2. Transport experiments

Also during the 1990s, another set of experiments was performed that focused on transport of the excitons rather than spectral signatures [77–80]. The experiments were done as follows: a long-wavelength (red) laser with photon energy nearly resonant with the exciton energy was used to flood a whole three-dimensional crystal of Cu2O with a low-density exciton gas. Then a very intense pulse of green light was used to excite one surface of the crystal, creating a very dense exciton cloud at the surface. Excitons were then detected on the opposite side of the crystal, in some cases over a millimeter distant, by a bimetal detector that ionized the excitons into free electrons and holes which then produced an external current. The signal would only occur if excitons migrated across the crystal into the detector.

The experiments showed that the exciton signal on the back side of the crystal jumped up in a sharp pulse when the green laser pulse intensity exceeded a critical threshold. The arrival time of the pulse corresponded to propagation at the speed of sound across the crystal. This was interpreted by the authors as superfluid motion of the excitons across the crystal when the excitons exceeded a critical density threshold. The high density pulse of excitons created by the green pulse was viewed as being amplified by the large population of cold excitons created by the resonant laser.

There were soon alternative interpretations of these experiments. Reference [81] pointed out that the width of the pulses was orders of magnitude longer than the expected coherence length, and that a classical soliton model could reproduce many of the effects. Tikhodeev and coworkers [82–85] pointed out that the intense green laser pulse also created many hot phonons which could blow as a nonequilibrium ‘phonon wind’ away from the surface, pushing excitons at the speed of sound, and reasonable models of the phonon wind could explain the exciton propagation. The phonon wind was well established (e.g., [86, 87]) in the 1980s as an effect in surface-excited solids; a flux of phonons streaming away from the surface can push excitons and EHL for hundreds of microns. One experimental result in support of this interpretation is that a second pulse was seen at the detector at much later times, corresponding to three passes across the crystal at the speed of sound, i.e., hitting the back surface, bouncing off, and coming back one round trip later [88]. This strongly supports the view that a sound pulse was created by the green laser pulse, and this sound pulse propagated across the crystal and reflected from the surfaces, sweeping excitons from the cold exciton gas as it traveled.
As with the spectral signature experiments, it was widely concluded by the end of the 1990s that these experiments with surface excitation also did not show BEC. Both of these experiments used very intense surface excitation in order to get very high exciton density, but both had the same drawback that the exciton gas thus created was very far from equilibrium and unconstrained in space, so that flow out of the excitation region had to be modeled, including the heat flow and phonon wind. Because of this, attention returned to trapping the excitons using stress.

2.3. A way forward

It became clear by the first International Conference on Spontaneous Coherence in Excitonic Systems (ICSCE) in 2003 [89] that the Auger effect played a dominant role in preventing EBEC in Cu$_2$O by shortening the lifetime of the excitons. To avoid this, the density of the excitons can be dropped, with a corresponding drop in the temperature, since $T_c$ of the BEC transition is proportional to the density to the 2/3 power in three dimensions. At first, one might think that dropping the density would not help, because the collision rate between the particles has the same dependence on density as the Auger rate, and therefore the ratio of the Auger lifetime to the collision time will not change. Excitons in a semiconductor have another means of thermalizing besides collisions, however, which is exciton–phonon scattering. The time scale for exciton–phonon equilibration is hundreds of picoseconds to nanoseconds [38, 90] at low temperature, so that when the gas has low density, the time scale for thermalization via phonons can be much shorter than the lifetime due to Auger recombination.

One issue in some solid-state systems in going to lower temperature and density is that the excitons can become localized in minima of the disorder (e.g., [91]), so that they do not move and do not thermalize. Many Cu$_2$O samples have extremely good quality, however, so that disorder is not a limiting factor. Another issue is that the polariton effect, i.e., mixing of the exciton and photon states, becomes important for excitons at low momentum, i.e., low temperature. There is a polariton effect in Cu$_2$O for both the orthoexciton species and for the paraexciton species in stressed crystals; although the direct recombination process for paraexcitons is forbidden by symmetry in unstressed crystals, if stress is used to create a trap for the excitons, the paraexciton coupling to the photons becomes comparable to that of the orthoexcitons [39]. It is possible, however, to go to temperatures so low that the excitons have momentum which is below the mixing region of the photon and exciton states.

The above considerations imply that the best temperatures to see BEC effects in Cu$_2$O are in the sub-Kelvin range of tens to hundreds of mK. This removes one of the original motivations for studying exciton condensates, which is to see high-temperature condensation; the fact that the excitons in Cu$_2$O are stable up to room temperature turns out not to help in getting them to condense at room temperature. One can still hope to see a true exciton BEC in a three-dimensional system, however. Two experimental groups have pursued EBEC experiments of excitons in traps in Cu$_2$O at millikelvin temperatures in the past 5 years; these are reviewed in section 4.

3. Review of the Auger process

As discussed above, one of the most important of the mechanisms which determine the lifetime of excitons in Cu$_2$O is believed to be a nonradiative two-body decay process, as evidenced by numerous experiments which have shown that the exciton lifetime is density dependent.

In Cu$_2$O, the exact nature of the Auger process is still an open question and is under current investigation. As discussed above, the general picture is that in such a process two excitons collide, the one recombinates, transferring its energy to the other, which ionizes. The electron and the hole which result from the ionization process then form an exciton, after a fraction of their energy is transferred to the lattice, via phonons. The final states of the ionized electron and hole, however, in particular the bands in which they lie, are not entirely clear.

Before we review the quantitative results on this process, we start mention some reasons which make the study of this process in Cu$_2$O difficult:

1. The experiments which have measured this process cover a rather wide range of exciton temperatures $T$ and exciton densities $n$. The Auger rate is likely to depend on the temperature, but the temperature of the excitons may not be the same as the lattice temperature in all cases, making it harder to determine the dependence on temperature accurately.

2. In excitons in Cu$_2$O there are many different processes which take place at the same time (radiative recombination, conversion between the orthoexcitons and the paraxcitons, expansion, etc) These processes obscure the measurement of the Auger decay process alone, since there is strong evidence that the rate of some of the above mechanisms is comparable to the rate of the Auger process.

In particular, the interconversion process of the total angular-momentum triplet state orthoexcitons into the total angular-momentum singlet state paraxcitons is of great importance. It is believed that it takes place via two mechanisms, namely a phonon-assisted process [68–71, 92, 93] and a collisional spin-exchange process [72, 73]. At low temperature and high density, the collisional ortho-para conversion process in many ways mimics the Auger collisional recombination process, making analysis much more difficult.

The rate equations (15) must therefore be modified to include density-independent and density-dependent ortho-para conversion, and the Auger rates are deduced by fitting the full equations to the data. Therefore, to a large extent, the determination of the Auger rate is dependent on a multi-parameter fit of several processes, and this complicates the problem. In many cases, the models must also include parameters for spatial diffusion of the excitons in order to estimate the volume of the exciton cloud.

In some regimes of parameters, different processes can be isolated. For example, experiments can be done in a very low density regime in which collisions between
excitons are negligible. In this case, the phonon-assisted conversion between orthoexcitons and paraexcitons can be determined very accurately [69, 70]. It is also possible to establish the intrinsic lifetime of the excitons by fitting the decay of the luminescence at very low densities, when the Auger recombination rate is negligible. In addition, at temperatures which are high compared to the energy splitting between the ortho and para states, chemical equilibrium between the ortho and para populations can be assumed, so that all ortho-para conversion processes drop out of the rate equations. This fact has been used to extract the total Auger rate of the exciton population in the high-temperature regime [94], although these results still relied on a model for the exciton cloud volume.

(3) The determination of the exciton density has proven to be a rather difficult task. There are basically three approaches that one may follow to extract the density. The first is the spectroscopic approach, in which some specific phonon-assisted recombination line is fitted to a Bose–Einstein distribution and from that the density and the temperature are extracted [51], as discussed in section 2.1 above. As discussed above, the estimates of density based on spectroscopic fits to an equilibrium homogeneous gas can strongly overestimate the density. The second method of extracting the exciton density is based on estimating the volume of the exciton gas and the number of excitons [54, 57, 95]. The third method relies on the absorption spectrum that corresponds to the 1s to 2p radiative excitonic transition. This method does not depend on the strength of the recombination lines and is thus advantageous in that respect. In addition, the absorption spectrum of this process is very sensitive to the degeneracy of the excitons and can serve as a clear and indisputable probe for the transition to a Bose–Einstein condensed phase [96].

(4) The theoretically-calculated decay rate involves several parameters with experimental uncertainty in their numerical value. As a result, the theoretical estimate of the decay rate is also uncertain.

The importance of the Auger process in excitons in Cu2O has been revealed in several experiments over the last three decades [37, 39, 57, 70, 75, 76, 93–95, 97–102]. The decay rate of the excitons shows an approximate proportionality to their density, which is an indication of a two-body collision process. The Auger process is also confirmed by the presence of orthoexcitons in stress-generated traps for times much longer than the orthoexciton lifetime [39, 99]: as discussed above, high-energy electrons and holes produced by the Auger ionization of para excitons lead to re-formation of excitons in essentially random internal angular momentum states, and thus to orthoexcitons in excess of the amount expected from Boltzmann occupation.

The difficulties discussed above have led to a wide scatter in experimental estimates of the Auger decay rate. The decay rate has been extracted to be negligible [101] and also non-negligible, and when non-negligible, to be independent of the temperature [76], to increase linearly with the exciton temperature [94], and to be inversely proportional to the exciton temperature [93]. The stress dependence of the Auger decay rate is also of importance [70]. The process is sensitive to the various symmetries of the crystal, which impose certain selection rules/constraints on the matrix elements and on the decay rate, while stress tends to lift these constraints. It has also been argued [102] that excitons in Cu2O form biexcitons, which then rapidly decay via the Auger process. While there is not yet direct experimental evidence for this process, the possibility of biexcitons in Cu2O needs to be explored further. The process considered in [102] is forbidden by symmetry, but the authors argued that the Auger mechanism is thus associated with breaking the band symmetries, e.g. by impurities, and therefore differences of impurity concentrations could explain the different Auger rates measured.

3.1. Review of Auger theory, constrained by experiments

Turning to the theoretical study of the Auger decay rate in Cu2O, according to [67] this process may be either ‘direct’, or phonon-assisted. In the phonon-assisted process there is a phonon involved, as shown in the diagram in figure 7. While the participation of a phonon suppresses the rate of this process, it lifts the symmetry constraints set by the band structure of Cu2O. We should stress here that there is an intimate connection between the radiative and the Auger decay mechanisms: the radiative lifetime of excitons is rather long because the dipole matrix element between the conduction and the valence bands (which have the same parity) vanishes. Phonon-assisted mechanisms provide an alternative way for the excitons to decay radiatively. The exciton–phonon interaction suppresses the rate of this process, but on the other hand it makes the process dipole allowed. As a result, the
phonon-assisted processes are actually the dominant radiative decay mechanisms. Rather similar results also hold for the direct and the phonon-assisted Auger processes: the theoretical 
calculations indicate that the dominant Auger process is the phonon-mediated process. Although in this case one also 
needs a phonon, which makes it less likely, on the other hand, there is no symmetry constraint, which enhances the rate of 
this process.

The dependence of the matrix element for the Auger 
collisions on the momentum exchange between the colliding 
excitons is of major importance [67]. If this is constant, then 
its only temperature dependence comes from phonon emission 
from [94]. If the matrix element is proportional to the momentum 
exchange, the Auger decay rate is proportional to the exciton 

$$\frac{1}{\tau_{\text{Auger}}} = n_a \frac{128h}{m} \left( \frac{m_e^{3/2}}{m^{1/2}} \right) \left( \frac{2\langle p \rangle^2}{mE_{cv}} \right) \left( \frac{\epsilon}{\epsilon_a} \right)^2 \times \frac{D^2}{\rho E_{\text{opt}} E_{cv}} \left( \int_0^b x^4 \sqrt{b^2 - x^2} \ dx \right)$$

$$\times \left( 1 + \frac{1}{\epsilon_{\text{opt}}/k_B T - 1} \right),$$

with

$$b^2 = \frac{2m_e a^2 E_{cv}}{\hbar^2},$$

where $m_e$ and $m_h$ are the effective mass of the electron and 
hole, respectively, $m$ is the standard electron mass, $\epsilon$ is the 
dielectric permittivity of the medium, $\langle p \rangle$ is the dipole moment 
for optical inter band transitions, $\rho$ is the crystal mass density, 
$D$ is the deformation potential for excitons interacting with 
optical phonons with energy $E_{\text{opt}}$, $a$ is the exciton Bohr radius, 
and $E_{cv}$ and $E_{cv}'$ are the energy differences between a higheryling 
conduction band $c'$ (which is assumed to be involved in 
virtual transitions) and the valence and conduction bands $v$ 
and $c$, respectively. All of the numbers in this formula can be 
estimated using results from other experiments. The optical 
phonon deformation potential is known from time-resolved 
experiments [38] to be on the order of a few times $10^7$ eV cm$^{-1}$; 
the oscillator strength $f = 2\langle p \rangle^2/mE_{cv}$ can be assumed to be 
of order unity [42]; the interband energies are known and are 
all on the order of an eV; the optical phonon energies have 
been measured accurately and are on the order of 10–20 meV; 
the exciton Bohr radius is known to be around 5 Å. (See [94] 
for the experimental references for these numbers.) These 
numbers give $A$ on the order of $10^{-18}$ cm$^3$ ns$^{-1}$ in the 
low-temperature limit. In other words, an exciton gas with density 
$10^{17}$ cm$^{-3}$ is predicted to have a lifetime of around 10 ns due to 
Auger recombination, much shorter than the typical radiative 
recombination times in the low-density limit.

The large number of parameters which go into (16), 
each with experimental uncertainty, especially the oscillator 
strength and the deformation potential, lead to significant 
uncertainty in this prediction. The experimental work of [94] 
found a best fit, shown in figure 8, of

$$A \approx 3.698 \times 10^{-17} \left( 1 + \frac{2}{e^{160/T - 1}} \right)$$

$$+1.298 \times 10^{-17} \left( 1 + \frac{2}{e^{247/T - 1}} \right),$$

for $A$ in units of cm$^3$ ns$^{-1}$ and the temperature $T$ in Kelvin. 
The two exponential factors come from the emission of the 
optical phonons $\Gamma_{12}$ and $\Gamma_{15}$, with energies 13.8 meV and 
18.7 meV, respectively. The temperature dependence of the 
thoretical formula is in good agreement with the experimental 
measurements for a wide range of temperature between 2 and 
235 K. According to equation (18), the rate is independent of 
$T$ up to about 77 K, while for higher temperatures the rate 
increases linearly with $T$ (which is obvious if one expands the 
exponentials). Treating the fit as dependent on essentially two 
fitting parameters, namely the matrix element $\langle p \rangle$ times the 
deformation potential $D$ for each phonon-assisted transition, 
each of which is squared in the calculation, the experimental 
fit is in agreement with the theory (16) if this product of 
parameters is within about a factor of 3 of the value estimated 
above.

References [95] and [100] estimated the exciton number 
and the volume of the gas at relatively low temperatures (and 
thus did not get the density spectroscopically). The values of 
the Auger constant $A$ that they found using this method were 
$7 \times 10^{-17}$ cm$^3$ ns$^{-1}$ and $6 \times 10^{-17}$ cm$^3$ ns$^{-1}$, respectively, which 
is in decent agreement with the result of equation (16) and with 
the experimental results of [94] at low temperature. Reference 
[37] also recently measured the orthoexciton–orthoexciton 
Auger rate to be $5 \times 10^{-17}$ cm$^3$ ns$^{-1}$ at low temperature, very 
close to the low-temperature value of [94].

For temperatures above 100 K, it is worth a close look to 
compare the results of [93, 94]. As noted above, the values 
for the Auger constant of the two groups at low temperature 
fall within the same range, with about a 30% deviation, which
is not surprising given the various experimental uncertainties. However, in the range 100–200 K, the two experiments showed opposite trends with temperature: the Auger rate was found to rise by about a factor of two in [94] and to fall by a little less than a factor of two in [93]. The two experiments used similar methods, with pulsed near-resonant dye-laser excitation. Even apart from any modeling, the raw data of [94] clearly showed that the excitons had shorter lifetime at high temperature, while the raw data of [93] shows the opposite trend.

The primary difference between the experiments was that a much tighter focus was used in the experiments of [93], about 15 µm radius, as compared to 120 µm radius in [94]. This gave higher densities and much greater range of lifetime, but may also have had other effects. One possibility is that the local lattice temperature may have gone well above the background bath temperature due to the concentration of energy, leading to a local ‘hot spot’ (see, e.g., [103, 104].) The exciton temperature may also have gone well above the lattice temperature, if the Auger rate was fast compared to the phonon emission rate. Taking the Auger rate as increasing with temperature as discussed above, if a local exciton cloud is hotter than the rest of the crystal, a higher Auger rate will occur. To actually have a decrease of the Auger rate with increasing lattice temperature, however, would require that the local temperature excursions was larger at lower bath temperatures. It is also possible that in the experiments of [94], at the highest temperatures there was also an excursion of the temperature above the lattice temperature, because the absorption length of the input laser shortened to approach 20 µm (although the lateral size still gave much larger volume for the spot size than the experiments of [93]).

Alternatively, or possibly in conjunction with a local hot spot, the rapid expansion of the exciton cloud seen in [93] may have played a role. In the experiments of [94], volume expansion was a minor factor because the initial size of the cloud was large. With the smaller initial spot size in [93], the expansion of the cloud played a bigger role. The authors of [93] carefully estimated the volume as a function of time \( V(t) \) based on a model of diffusion to get the density \( n(t) = N(t)/V(t) \), but the relative uncertainty of the volume will be larger with a smaller spot size.

At present, [76] and [101] are the outliers. Reference [76] reported a value for the Auger rate of approximately \( 4 \times 10^{-16} \) cm³/s, at low temperature, and [101] reporting a much lower, indeed negligible, rate, of the order of \( 10^{-23} \) cm³/s. The majority of other experimental work [37, 93, 94] gives Auger rates in the range of \( 10^{-17} \) to \( 10^{-16} \) cm³/s, Reference [76] reported uncertainty of a factor of three, which would allow a value in this range by taking the low side of the uncertainty range.

In the case of the work by Gonokami’s group, a multi-step process of analysis and measurements was used to derive the Auger rate. In the first step, the absolute number of excitons is established by far infrared Lyman absorption spectroscopy of the 1s–2p, 1s–3p, etc. transitions for both orthoexcitons and paraexcitons [74–76]. The reason that the Lyman spectroscopy can establish the absolute number of excitons while band-to-band absorption cannot is that the absolute cross section for intraband transitions, while the band-to-band oscillator strength cannot.

Having this method to establish the absolute number of both ortho and paraexcitons, the number must be converted to a density by an estimate of the volume of the exciton cloud. In a stress-generated harmonic trap, this can be done by imaging the exciton cloud from two sides of the crystal, but typically the number of excitons in such a trap is too few to generate a reasonable Lyman absorption measurement. Therefore, Gonokami’s group generated a large exciton cloud in a bulk crystal, without a trap, and a model for the exciton diffusion was used to estimate their volume. This modeling is the greatest source of uncertainty in the measurements, since various effects such as phonon wind, discussed above, may also affect the exciton motion. The authors of [76] used the diffusion constant for paraexcitons reported in [105], which was very high. This was reasonable, given that their sample was cut from the same natural sample as that used in [105], but was not checked directly. If the diffusion constant was estimated too high, then the volume would be estimated as much greater, and therefore the density much lower, which could give systematically high Auger rates.

The work of [101] gave the opposite result of very low Auger rate for both paraexcitons and orthoexcitons. In this experiment, the absolute density of the excitons was estimated by using the effect of screening of excitons at high density on the absorption spectrum [106], based on a theory of the Mott transition. Their theory of the screening led them to estimate initial densities on the order of \( 10^{15} \) cm⁻³, and under these conditions they saw no evidence of an Auger effect. The theory of the excitonic Mott transition is hotly debated and quite complex, however [13–15, 107], and therefore it is difficult to find agreement on a universal theory for the density at which the exciton absorption will be screened out.

The above experiments primarily were sensitive to the ortho-ortho Auger recombination rate. The question of the value of the Auger collisions between paraexcitons is even more subtle. Reference [37] has provided evidence for a very low paraexciton–paraexciton Auger rate, on the order of \( 10^{-18} \) cm³/s. Earlier experiments of [70] in stressed crystals also gave a para-para Auger rate constant for all collisions in their model, so that their fit was not sensitive to a difference between the para-para, ortho-, and ortho-ortho rates.

Theoretically the para-para Auger collisions should be negligible [67] at zero stress (the effect of stress is discussed below). The reason for this is the band structure of Cu₂O, which makes the direct Auger process for para-para collisions forbidden, while the phonon-assisted process is highly suppressed [67]. This is due to the fact that the phonon-assisted Auger process for para-para Auger collisions involves a virtual transition to a deep valence band, as opposed to the ortho-ortho collisions, for which this process involves a conduction band which is close in energy to the lowest conduction band.
3.2. Orthoexciton to paraexciton conversion process and 'saturation' of the orthoexciton gas

As discussed above, at very low densities, when the two-body processes are negligible, it has been observed that orthoexcitons convert into paraexcitons through a phonon-assisted process \[68, 69, 71, 92, 93\]. According to \[69\] its rate scales as \(T^{3/2}\). More recently \[92\] and \[93\] have reported that its rate is constant at very low temperatures and increases linearly with temperature at higher temperatures.

In addition to the phonon-assisted process, orthoexcitons convert into paraexcitons via a spin-exchange process, where two orthoexcitons collide, they exchange their electrons or holes, resulting into two paraexcitons in the final state. This process was studied theoretically in \[73\], while \[75\] has shown experimental evidence for it. The decay rate of this process is proportional to the orthoexciton density \(n_o\).

A simple argument has been proposed at various times over the years (see, e.g., \[108\] and \[109\]) to explain the saturation effect of orthoexcitons, discussed in section 2.1, which corresponds to a power law \(n_o \propto T^{3/2}\). Recall that the exciton gas, which in general has a higher temperature than the lattice temperature, exchanges phonons with the lattice and thus loses energy. From simple deformation potential theory, the rate of loss of energy is proportional to \(T^3\), see \[42\], section 5.1.4. On the other hand, both the Auger process and the collisional conversion process of orthoexcitons to paraexcitons heat the gas. In both cases the rate of energy gain of the gas is \(bn_o\), where \(n_o\) is the orthoexciton density. As a result, we can write a rate balance equation for the entropy of the orthoexcitons divided by the orthoexciton number \(S_o/N_o\):

\[
T \frac{d}{dt} \left( \frac{S_o}{N_o} \right) \approx -aT^{3/2} + bn_o, \tag{19}
\]

which implies that the orthoexcitons move along adiabats, where \(n_o \propto T^{3/2}\), in agreement with the experimental data \[99\]. This above argument applies both in the classical and in the degenerate regime \[109\].

A problem with this argument, however, is that the rate of heat loss for orthoexcitons in Cu\(_2\)O never exactly follows the \(T^{3/2}\) law predicted by elementary deformation potential theory. There are both high-temperature corrections, due to optical phonon emission and high-frequency corrections of the exciton wave function, and low-temperature corrections due to momentum-conservation limitations \[105\], that give deviations from this law. The rate of phonon emission in Cu\(_2\)O has been measured carefully and reported in \[38\]; a summary plot is shown in figure 9.

Reference \[57\] also addressed the saturation effect, using a classical rate equation model of with only acoustic phonon emission for energy relaxation and including the Auger process. They did not obtain a \(n \propto T^{3/2}\) power law but did find a general trend of temperature increasing with density comparable to the experimental results.

3.3. The effect of stress

As we mentioned earlier, the Auger decay process is strongly influenced by stress. In stressed crystals the selection rules which determine the Auger decay rate are no longer valid, and as a result stress enhances the rate of this process.

The most important issue in the case of stressed crystals is the effect of stress on the paraexciton–paraexciton Auger collisions. As discussed above, in non-stressed crystals it is an open question whether the para-para Auger process is allowed, but it is certain that under stress it is allowed. An analogous situation is the radiative decay of paraexcitons, where stress makes the direct radiative recombination process of paraexcitons allowed, and it also enhances the phonon-assisted recombination lines. More specifically, uniaxial stress mixes the \(\Gamma^+\) and \(\Gamma^-\) valence bands with the result that the paraexciton energy increases for small values of the applied stress but decreases for higher values, and the direct radiative recombination increases quadratically with increasing stress \[41, 110–112\].

Indeed, the experiment reported in \[39\] shows clearly that para-para Auger collisions are allowed in stressed crystals. Uniaxial stress makes the direct Auger process allowed as well as the phonon-assisted process for paraexcitons. In both cases the orientation of the uniaxial stress is very important, since it determines the selection rules in the deformed crystal \[41, 110–112\]. These processes are responsible for the presence of orthoexcitons in the stress-generated trap \[39\]. The orthoexcitons are generated from free electrons and holes from Auger-ionized paraexcitons; the free electrons and holes are assumed to enter random angular momentum states.

A more recent experiment \[70\] has found that the Auger recombination rate for paraexciton collisions increases with...
increasing stress, while the rate of conversion of orthoexcitons into paraexcitons decreases. Furthermore, the same study found that for zero stress the paraexciton Auger collisions are negligible. As discussed above, these results are consistent with the theoretical study of this problem.

The fact that the paraexciton Auger process becomes stronger under crystal stress creates a drawback for using stress to create a harmonic trap for the excitons. While the trap can be used to increase the density of the excitons, it also shortens the lifetime. Which process wins out will depend on the temperature, the trap depth, and other experimental details.

4. Recent experiments on exciton BEC in Cu$_2$O

Two currently ongoing experiments are attempting to produce EBEC in Cu$_2$O under the most optimal conditions. In these experiments, stress is used to produce a three-dimensional harmonic trap for paraexcitons. The lattice temperature is taken to the milliKelvin range, so that paraexcitons can equilibrate in states below the polariton mixing region, and also so that very low densities can be used, giving a very slow Auger recombination rate.

As discussed in section 3.1, Gonokami’s group in Japan performed measurements of the Auger recombination rate of the excitons in Cu$_2$O using exciton Lyman spectroscopy to establish the number of excitons and a theoretical model for the diffusion to estimate the volume of the exciton cloud. Based on this method, the Auger cross section was found to be independent of temperature over the range of 4–70 K [76]. The authors suggested that this is similar to the case of inelastic collisions of atoms, but as discussed in section 3.1, this result is also consistent with the phonon-assisted mechanism for Auger recombination presented in [67].

The authors then dropped the temperature into the range of tens to hundreds of mK and performed both spatial imaging and spectroscopy of the exciton photoluminescence. Initially, they found that the exciton cloud size agreed well with the predictions of equilibrium of the exciton cloud with the lattice temperature. The cloud size depends on temperature in a trap because the harmonic potential of the trap leads to the relation $r^2 \propto k_B T$.

As the lattice temperature dropped, the exciton cloud temperature as measured by the cloud size and energy distribution stayed at around 800 mK. Below a lattice temperature of about 400 mK, an ‘explosion’ was found [113], in which the cloud expanded and was very far out of equilibrium (see figure 10). One mechanism which could in principle lead to this is that proposed by Hijmans et al. [114] for atomic hydrogen condensates. In this scenario, in a harmonic trap, the condensate may have such small volume and high density that density-dependent loss processes are strongly enhanced. The size of the condensate cloud is controlled by the repulsive interactions of the particles; if these are weak, then the size of the condensate can be quite small.

Strange nonequilibrium effects of the paraexcitons at high density have been seen before, namely the observation of a strongly non-Maxwellian energy distribution of the paraexcitons, seen in the work done with surface excitation in the late 1980s and early 1990s [52]. It is possible that when the paraexcitons approach the BEC state, some nonequilibrium recombination effect does indeed prevent them from condensing.

The German collaboration of Stolz and coworkers [37] and Fröhlich and coworkers [117] also has seen equilibration of paraexcitons in a trap, including at temperatures in the sub-Kelvin range. No evidence of an ‘explosion’ of the type seen by Gonokami and coworkers was seen at low temperature, but the conditions were not exactly the same. In general, the degree of equilibration will depend on the rate of heat removal from the sample in the cooling system, the excess energy of the pump laser which creates the excitons, the intrinsic lifetime of the excitons, which depends on the impurities in the sample, and the duration of the laser pulse. Stolz’s group has argued [115, 116] that the mean-field renormalization of the exciton self-energy at high density should lead to a blue shift of the
paraexciton line at high density and a change of the condensate spatial profile. Such a shift has not generally been seen in excitons in Cu$_2$O at high density (see, e.g., figure 4). Although the mean-field density dependence is generally expected to lead to a blue shift (see, e.g., [42], chapter 8), and this has been seen with microcavity polariton condensates [118] and with indirect excitons in GaAs quantum double wells [119], in bulk semiconductors like Cu$_2$O the effect of higher-order correlations seems to cancel out most of this blue shift. A blue shift may indicate that much higher effective densities have been reached, although also at higher temperature than needed for condensation.

5. Future prospects

The process of the ‘explosion’ seen by Gonokami’s group is not well understood. It is possible that further theoretical study will allow a full understanding of this effect and its connection to condensation. In the meantime, Gonokami’s group has reported that the explosion effect seems to be less dramatic at lower temperatures, which imply lower densities for condensation, at lower temperatures, and so experiments to go to even lower temperatures, which may lower densities for condensation, may overcome this barrier.

Instead of just looking at the photoluminescence, far infrared absorption can be used to study the excitons. In [96], it was shown that the absorption spectrum of infrared radiation inducing internal transitions of the excitons from the 1s to the 2p level is very sensitive to the degree of quantum degeneracy of the gas. A similar idea has been used in the experiments with atomic hydrogen [120]; however, in this case the mass of the excitons in the 1s and the 2p states is different (due to band structure effects [121]), as opposed to the case of hydrogen, where the mass in the 1s and in the 2p states is the same to a very good approximation. Such experiments have been performed, as described above, e.g., [75, 76, 101].

In such an experiment one can observe the contribution of the orthoexcitons and of the paraexcitons to the absorption separately, with an energy separation of order of the orthoexcitons and of the paraexcitons to the absorption as described above, e.g., [75, 76, 101].

Despite many false trails, research on excitons in Cu$_2$O has progressed and is still active, and there is good reason to believe that progress on establishing BEC of excitons in this well-studied crystal may still happen.

References

[1] Moskalenko S A 1962 Fiz. Tverd. Tela. 4 276
[2] Blatt J M, Böer K W and Brandt W 1962 Phys. Rev. 126 1691
[3] Casella R C 1963 J. Phys. Chem. Solids 24 19
[4] E.g., Huang K 1987 Statistical Mechanics 2nd edn (New York: Wiley)
[5] Keldysh L V and Kozlov A N 1967 Zh. Exp. Teor. Fiz. Pis’ma 5 238
[6] Keldysh L V and Kozlov A N 1968 Zh. Exp. Teor. Fiz. 54 978
[7] Hanamura E and Haug H 1974 Solid State Commun. 15 1567
[8] Hanamura E and Haug H 1975 Phys. Rev. B 11 3317
[9] Hanamura E and Haug H 1977 Phys. Rep. C 33 209
[10] Comte C and Nozières P 1982 J. Phys. 43 1069
[11] Nozières P and Comte C 1982 J. Phys. 43 1083
[12] Moskalenko S A and Smoke D W 2005 Bose–Einstein Condensation of Excitons and Biexcitons (Cambridge: Cambridge University Press)
[13] Smoke D W 2008 Solid State Commun. 146 73
[14] Manzke G, Semkat D, Richter F, Kremp D and Henneberger K 2010 J. Phys.: Conf. Ser. 210 012020
[15] Semkat D, Richter F, Kremp D, Manzke G, Kraeft W-D and Henneberger K 2009 Phys. Rev. B 80 155201
[16] Wolfe J P and Jeffries C D 1987 Electron–Hole Droplets in Semiconductors ed C D Jeffries and L V Keldysh (Amsterdam: North-Holland)
[17] Keldysh L V 1987 Electron–Hole Droplets in Semiconductors ed C D Jeffries and L V Keldysh (Amsterdam: North-Holland)
[18] Timofeev V B, Kulakovskii V D and Kukushkin I V 1983 Physica B+C 117/118 327
[19] Kuroda H, Shionoya S, Saito H and Hanamura E 1973 J. Phys. Soc. Japan 35 534
[20] Goto T, Anzai T and Ueta M 1973 J. Phys. Soc. Japan 35 940
[21] Nagasawa N, Nakata N, Doi Y and Ueta M 1975 J. Phys. Soc. Japan 38 593
[22] Nagasawa N, Nakata N, Doi Y and Ueta M 1975 J. Phys. Soc. Japan 39 987
[23] Chase L L, Peyghambarian L L N, Gryenberg G and Mysyrowicz A 1979 42 1231
[24] Peyghambarian N, Chase L L and Mysyrowicz A 1983 Phys. Rev. B 27 2325
[25] Yakhut Y and Levich E 1980 Phys. Lett. A 80 301
[26] Smoke D W and Wolfe J P 1989 Phys. Rev. B 39 4030
[27] Hartwell V E and Smoke D W 2010 Phys. Rev. B 82 075307
[28] Tassone F and Yamamoto Y 1999 Phys. Rev. B 59 10830
[29] Malpuech G, Kubo Y, Laussy F, Bigenwald P and Kavokin A 2003 Semicond. Sci. Technol. 18 S395
[30] Agekyan V T 1977 Phys. Status Solidi a 43 11
[31] Bassani F and Rovere M 1976 Solid State Commun. 19 887
[113] Yoshioka K, Chae E and Kuwata-Gonokami M 2011 Nature Commun. 2 328
[114] Hijmans T W, Yu Kagan, Shlyapnikov G V and Walraven J T M 1993 Phys. Rev. B 48 12886
[115] Soltz H and Semkat D 2010 Phys. Rev. B 81 081302
[116] Sobkowiak S, Semkat D, Soltz H, Koch T and Fehske H 2010 Phys. Rev. B 82 064505
[117] Sandfort C, Brandt J, Finke C, Fröhlich D and Bayer M 2011 Phys. Rev. B 84 165215

[118] Balili R, Nelsen B, Snoke D W, Pfeiffer L and West K 2009 Phys. Rev. B 79 075319
[119] Vörös Z, Snoke D W, Pfeiffer L and West K 2009 Phys. Rev. Lett. 103 016403
[120] Fried D G, Killian T C, Willmann L, Landhuis D, Moss S C, Kleppner D and Greytak T J 1998 Phys. Rev. Lett. 81 3811
[121] Kavoulakis G M, Chang Y-C and Baym G 1997 Phys. Rev. B 55 7593