Exciton mass and exciton spectrum in the cuprous oxide

A Alvermann and H Fehske

Institut für Physik, Ernst-Moritz-Arndt-Universität, D-17487 Greifswald, Germany

E-mail: alvermann@physik.uni-greifswald.de

Received 29 July 2017, revised 25 November 2017
Accepted for publication 8 December 2017
Published 23 January 2018

Abstract

Excitons with a radius of a few lattice constants can be affected by strong central-cell corrections, leading to significant deviations of the optical spectrum from the hydrogen-like Rydberg series, and also to an enhancement of the exciton mass. We present an approach to this situation based on a lattice model that incorporates the effects of a non-parabolic band structure, short distance corrections to the Coulomb interaction between electrons and holes, spin–orbit and exchange coupling. The lattice model allows for observation of the crossover from large radius Wannier to small-radius Frenkel excitons without invoking a continuum approximation. We apply the lattice model approach especially to the yellow exciton series in the cuprous oxide, for which the optical spectrum and exciton mass enhancement are obtained through adaptation of only a few model parameters to material-specific values. Our results predict a strongly anisotropic ortho-exciton mass.

Keywords: excitons, cuprous oxide, central-cell corrections, exchange interaction

1. Introduction

Excitons in inorganic semiconductors such as silicon, germanium, or gallium arsenide are usually well understood by Wannier theory, which describes excitons in analogy to the hydrogen atom [1]. Especially the exciton energies follow the Rydberg series \( E_n = E_{\text{gap}} - R_X/n^2 \), and the excitonic Rydberg \( R_X \) can be computed from the dielectric constant and the reduced electron–hole mass. In some semiconductors however, with the cuprous oxide \( \text{Cu}_2\text{O} \) as one of the most prominent representatives [2, 3], excitonic properties differ significantly from this simple picture. Prominent features are the deviation of the energies of even exciton states from the Rydberg series, and of the exciton mass from the sum of electron and hole mass. These deviations can be attributed to central-cell corrections that become important whenever excitons are strongly bound states and thus sufficiently small to be affected by such corrections [4].

In the cuprous oxide, odd exciton states follow the Rydberg series rather accurately, with an excitonic Rydberg \( R_X \approx 97 \text{ meV} \). Even exciton states, on the other hand, defy Wannier theory: the binding energy of the 1S (ortho-) exciton state is about 40% larger than \( R_X \), and the exciton mass is about 50% larger than the sum of electron and hole mass (see table 1). Much theoretical effort has been invested into the exploration of this situation and its consequences [5–9].

One aspect of the explanation is hinted at by the non-parabolic dispersion of the upper valence band (see figure 1), which gives rise to the yellow exciton series. We should expect that such a strong violation of the effective mass approximation already at small momentum will lead to strong deviations from Wannier theory.

Recently, a comprehensive treatment of the even exciton spectrum has been presented in [13], with good agreement between theory and experiment. This reference presents a detailed analysis of spin–orbit coupling, exchange interaction, and central-cell corrections to the Coulomb attraction between electron and hole, which are incorporated into an extended Luttinger-type [14–20] Hamiltonian.

In this paper, we approach the problem of excitons with strong central-cell corrections from a different angle. Our goal is to relate two disparate effects, the shift of the even exciton states relative to the Rydberg series and the enhancement of the exciton mass relative to the sum of hole and electron mass, through one theoretical description. We will try to achieve this goal with a microscopic approach that is based...
on a lattice model instead of the continuum approximation. Central-cell corrections occur naturally in the lattice model when excitons become small.

In the remainder of this paper we introduce the exciton lattice model in section 2 and discuss the various terms, when excitons become small. Central-cell corrections occur naturally in the lattice model instead of the continuum approximation.

### 2. Exciton lattice model

The exciton lattice model comprises three valence bands and one conduction band with cubic symmetry. Its construction follows the general reasoning for Luttinger-type exciton models [13–20], but in contrast to previous studies we consider the lattice structure explicitly and do not invoke a continuum or effective mass approximation.

The valence bands are formed out of three orbitals \(|\vec{x}\), \(|\vec{y}\), \(|\vec{z}\rangle\) at each lattice site, which transform according to the \(\Gamma^+_7\) representation of the cubic symmetry group \(O_h\) [22]. We can picture these orbitals as d-type orbitals, associating \(|\vec{x}\rangle \approx |d_x\rangle\) etc. The choice of the \(\Gamma^+_7\) representation is specific to the cuprous oxide [23, 24], other representations would be equally possible. The hole states \(|r_\text{h}, d, s_\text{h}\rangle = |r_\text{h}\rangle \otimes |d\rangle \otimes |s_\text{h}\rangle\) are labelled by the hole position \(r_\text{h}\), orbital index \(d = x, y, z\), and hole spin orientation \(s_\text{h} = \uparrow, \downarrow\).

The conduction band is formed out of a single orbital per lattice site, which transforms according to the \(\Gamma^+_7\) representation of \(O_h\). The electron states \(|r_\text{e}, s_\text{e}\rangle = |r_\text{e}\rangle \otimes |s_\text{e}\rangle\) are labelled by the electron position \(r_\text{e}\) and spin orientation \(s_\text{e} = \uparrow, \downarrow\).

In combination, the hole–electron states \(|r_\text{h}, d, s_\text{h}\rangle \otimes |r_\text{e}, s_\text{e}\rangle\) that appear in the excitonic wave function have five indices. The entries of the hole and electron position vectors \(r_\text{h}, r_\text{e}\) can assume \(3 \times 2 \times 2\) different values.

The five terms in the lattice model Hamiltonian

\[
H_X = -H_\text{hole} - H_\text{so} + H_\text{electron} + H_\text{Coulomb} + H_\text{ex}
\]

(1)

capture the most important contributions to exciton formation: the hole \((H_\text{hole})\) and electron \((H_\text{electron})\) kinetic energy and the Coulomb attraction between electron and hole \((H_\text{Coulomb})\). The spin–orbit coupling for the hole states \((H_\text{so})\) is instrumental for explaining the peculiar shape of the valence bands, and thus of the yellow excitons in the cuprous oxide. The exchange interaction \((H_\text{ex})\), which depends on the relative orientation of electron and hole spin, leads to the splitting of ortho (triplet) and para (singlet) excitons.

For the present study we restrict ourselves to the most important leading order terms, and thus minimise the number of adaptable model parameters. For a fully accurate description of excitonic properties, extension of the model by additional terms might be necessary.

#### 2.1. Hole kinetic energy

The leading term in the hole kinetic energy is the nearest-neighbour hopping term

\[
H_{\text{hole}} = \sum_{d=x,y,z} \left( t_1 (1 - I_d^2) + t_2 I_d^2 \right) (T(e_d) + T(-e_d)).
\]

(2)

In this expression, \(T(e)\) denotes the translation operator in direction \(e\), i.e., \(T(e)r_\text{h} = r_\text{h} + ae\) with the lattice constant \(a\). The vector of nearest neighbour translations is denoted by \(e_d\), i.e., \(e_d = (1, 0, 0)'\) etc.

The operators \(I_d\) for \(d = x, y, z\) act on the hole states according to

\[
\begin{align*}
I_x|\vec{x}\rangle &= 0, & I_x|\vec{y}\rangle &= -i|\vec{z}\rangle, & I_x|\vec{z}\rangle &= -i|\vec{y}\rangle, \\
I_y|\vec{y}\rangle &= 0, & I_y|\vec{z}\rangle &= i|\vec{x}\rangle, & I_y|\vec{x}\rangle &= +i|\vec{y}\rangle, \\
I_z|\vec{z}\rangle &= +i|\vec{x}\rangle, & I_z|\vec{x}\rangle &= +i|\vec{y}\rangle, & I_z|\vec{y}\rangle &= 0.
\end{align*}
\]

(3)
The hole kinetic energy (2) describes three valence bands \( d = x, y, z \) (see figure 2) with dispersion
\[
E_h^{(d)}(k_i) = 2t_i \cos (k_i \cdot e_d) + 2t_d \sum_{i = d} \cos a(k_i \cdot e_t).
\]

On physical grounds, considering the overlap of valence band orbitals with orientations parallel and perpendicular to the direction of hole momentum \( k_h \), we expect \( 0 < t_2 < t_1 \). The corresponding light (\( m_{lh} \)) and heavy (\( m_{hh} \)) hole mass, along the direction of a lattice axis \( k_i || e_d \), is \( m_{lh} = h^2/(2a^2 t_1) \) and \( m_{hh} = h^2/(2a^2 t_2) \). In intermediate directions, the hole mass interpolates between these two values.

### 2.2. Spin–orbit coupling

The spin–orbit coupling term for the hole takes the form
\[
H_{so} = -\frac{2E_{so}}{3\hbar} (1 \cdot S^2),
\]
with the hole spin operator \( S^2 \). It connects hole states within the same unit cell. The minus sign in the spin–orbit coupling term is explained effortlessly by the d-type nature of the valence band orbitals.

At \( k_h = 0 \) the hole states possess full cubic symmetry, which results in a four-fold and a two-fold degenerate state according to the splitting \( \Gamma_3^+ \otimes \Gamma_6^- = \Gamma_1^+ + 2\Gamma_2^+ \) of the orbital-spin representation for valence band states [22]. For general \( k_h \neq 0 \) the symmetry is reduced, and only a two-fold degeneracy remains. The two-fold degeneracy is associated with reflection symmetry of the valence band Hamiltonian \( H_{hole} + H_{so} \) (see appendix B).

Because of the minus sign in equation (5), the two-fold degenerate ‘split-off’ band is shifted to higher energies. Yellow excitons in the cuprous oxide involve holes in this split-off band. At \( k_h = 0 \), the two set of states are separated by the spin–orbit coupling energy \( E_{so} \), as depicted in figure 2.

The full hole dispersion can be obtained through diagonalization of a \( 3 \times 3 \) matrix (see appendix B). For the split-off band in figure 2(b) (right panel) we have the following situation. For small \( |k| \to 0 \), the dispersion along the direction of a lattice axis, e.g., along the \([100]\) direction \( e_x \), is
\[
E(k_x) - E_i = \frac{2}{3}E_{so} + \frac{1}{3}(t_1 + 2t_2)(ak_i)^2 + \frac{8(t_1 - t_2)^2}{36E_{so}}(ak_i)^4 + O((ak_i)^6),
\]
which gives the new hole mass in the split-off band
\[
m_{s-o} = \frac{3h^2}{2a^2(1 + 2t_2)} = \frac{3m_{lh}m_{hh}}{2m_{lh} + m_{hh}},
\]
and a pronounced non-parabolicity from the \( k^4 \) term. Here, \( E_i = 2t_1 + 4t_2 \) is the energy of the valence bands at \( k = 0 \) according to \( H_{hole} \), i.e., essentially the Fermi energy in the absence of spin–orbit coupling. We have \( m_{lh} < m_{s-o} < m_{hh} \).

For larger \( |k| \), in the region of \( E_{so} \lesssim 2(t_1 - t_2)(ak_i)^2 \), the dispersion changes into
\[
E(k_x) - E_i = \frac{1}{3}E_{so} + 2t_2(\cos ak_i - 1),
\]
which is the dispersion of the heavy hole band with mass \( m_{hh} \), shifted by spin–orbit coupling. As a consequence of spin–orbit coupling, the hole mass in the split-off band changes significantly already for small \( |k| \). The pronounced non-parabolicity of the hole dispersion is partly responsible for the peculiar properties of yellow excitons in the cuprous oxide.

### 2.3. Electron kinetic energy

For one conduction band, the leading term of the electron kinetic energy has the simple form
\[
H_{electron} = -t_e \sum_{i=x,y,z} (T(e_i) + T(-e_i)),
\]
where \( T(e) \) is still the translation operator with \( T(e)|r_i⟩ = |r_i + ae⟩ \). The Hamiltonian (9) describes a conduction band with cosine dispersion \( E_c(k_x) = -2t_e \sum \cos a(k_x \cdot e_t) \) and electron mass \( m_e = h^2/(2a^2 t_e) \). Subsequent terms in the electron kinetic energy are restricted by cubic symmetry, and can be assembled in analogy to the hole kinetic energy terms (see appendix A).

### 2.4. Coulomb attraction

Excitons form through the Coulomb attraction between electron and hole. We assume an interaction
\[
H_{Coulomb} = \sum_{r_e,r_h} U(r_e - r_h)|r_h⟩⟨r_h| \otimes |r_e⟩⟨r_e|,
\]
which is diagonal in the spin, orbital, and lattice indices, and choose the expression
\[
U(r) = -e^2 \frac{1}{r} \text{ if } r \neq 0,
\]
\[
\frac{1}{r} \text{ if } r = 0
\]
for the dependence on the electron–hole separation. Here, $e$ is the elementary charge and $\epsilon$ a material-specific dielectric constant. The parameter $\ell_C$ is a characteristic length comparable to the width of the valence and conduction band orbitals.

For large electron–hole separation, the Coulomb potential decays $\propto 1/(r_e - r_h)$, while for $r_e = r_h$ it assumes a finite value $U(0) \propto 1/\ell_C$. Note that a finite local value $U(0)$ does not result from the screening of the Coulomb attraction, but from the finite size of the valence and conduction band orbitals.

The correct choice of the dielectric constant $\epsilon$ is debatable, as screening of the Coulomb attraction depends on distance. Interpolation between the different values of the dielectric constant is possible through modified expressions for the Coulomb potential (see, e.g., [25]). To avoid the introduction of additional parameters, we keep the above simpler expression. Some of the distance dependence of screening can be captured through adaptation of the model parameter $\ell_C$.

2.5. Exchange interaction

Exchange interaction leads to the splitting of exciton states into ortho-(triplet)-excitons and para-(singlet-) excitons. We assume a local exchange interaction

$$\mathcal{H}_{ex} = E_{ex} \sum_{r_e=r_h} \left( \frac{1}{4} - \frac{1}{\hbar^2} S^h \cdot S^e \right) \otimes | r_h \rangle \langle r_h | \otimes | r_e \rangle \langle r_e |,$$

with the hole (electron) spin operator $S^h(S^e)$, which connects electron and hole states in the same unit cell and is diagonal in the orbital indices. In the limit of zero exciton radius, the 1S ortho-exciton state is shifted by $E_{ex}$ relative to the 1S para-exciton state, which is not affected. The level splitting decreases with increasing exciton radius, and is zero for odd exciton states.

3. Excitons in the lattice model

Our computations determine the exciton wave function

$$| \psi_{K} \rangle = \sum_{r_e} e^{i K \cdot r_e} (T(r_e) | \phi_{K} \rangle) \otimes | r_e \rangle,$$

which is the translational invariant extension of the relative hole–electron wave function

$$| \phi_{K} \rangle = \sum_{r_h \in \mathcal{D}} \sum_{d=x,y,z} \sum_{s_h, s_e = \uparrow, \downarrow} \phi_{K}(r_h, d, s_h, s_e) | r_h \rangle \otimes | r_h, d, s_h \rangle \otimes | s_e \rangle.$$

The latter wave function is an eigenstate of the Bloch Hamiltonian $H_{X}(K)$, parametrised by the exciton momentum $K$. In the above expression for $| \phi_{K} \rangle$, we fix the electron position $r_e$ and use the hole position $r_h$ as the relative coordinate. Any other convention is equally possible, and we can swap the role of the electron and hole.

We compute the exciton wave function and energies numerically with standard sparse matrix eigenvalue solvers [26, 27] applied to $H_{X}(K)$. The numerical effort can be reduced considerably by using an adaptive scheme to deal with the increasing radius of excited exciton states [28].

From the exciton wave function at zero momentum $K = 0$, the exciton radius is obtained as the expectation value

$$a_{X} = \frac{2}{3} \langle r \rangle = \frac{2}{3} \langle (x^2 + y^2 + z^2)^{1/2} \rangle.$$

We include the prefactor $2/3$ to recover the Bohr radius, which is defined as the most probable radial distance, in the continuum limit where $\phi_{K}(r) \propto e^{-r/aX}$.

From the exciton energy $E_{K}$ at finite momentum, the exciton mass is obtained as the derivative

$$m_{X}^{-1} = \frac{1}{\hbar^2} \frac{\partial^2 E_{K}(E_{d})}{\partial E_{d}^2} \bigg|_{E_{d} = 0}.$$

In general, the exciton mass is anisotropic, but we will only report its values along a lattice axis $K||E_{d}$.

Note that in order to obtain the exciton energies $E_{d} = E_{gap}$, the computed eigenvalues of $H_{X}$ have to be shifted by a constant $E_{gap} = 2t_{1} + 4t_{2} + 6t_{4} + (2/3)E_{so}$—the groundstate energy of $H_{X}$ at zero Coulomb attraction—that is determined by the above form of the kinetic energy and spin–orbit coupling terms. From the exciton energies, the spectral lines are obtained by adding the experimental band gap value $E_{gap} = 2.172$ eV, which is a free constant that does not enter the Hamiltonian $H_{X}$.

3.1. Symmetry considerations

Due to the coupling of the orbital and spin degrees of freedom and the reduced symmetry in the lattice model in comparison to the continuum description, most simple observables fail to be conserved. In particular, neither the effective orbital spin $I$, hole spin $S^h$, electron spin $S^e$ nor any of their combinations commute with the full Hamiltonian. Already for the hole Hamiltonian $H_{hole} + H_{so}$, the effective orbital-hole spin $I + S^h$ is conserved only at $k_h = 0$ or other points of high symmetry.

The minimal symmetry that is preserved at least for zero exciton momentum $K = 0$ is the cubic symmetry $O_h$. We do not use the entire group $O_h$ to set up the computational problem, but use the following easier construction that results in a comparable reduction of the computational effort.

The Hamiltonian $H_{X}$ and the three operators $R_{d} = (2\ell_{d}^{2} - 1) \otimes (2/3)S^h_{d} \otimes (2/3)S^e_{d}$ for $d = x, y, z$ commute among themselves. This allows us to split the Hamiltonian, with its $3 \times 2 \times 2 = 12$ local degrees of freedom (orbital, hole spin, electron spin), into four operators acting on three local degrees of freedom each. The respective basis states are listed in appendix C.

Geometrically, the operators $R_{d}$ implement a local reflection in orbital and spin space, and thus provide a representation of the Abelian group $Z_{2} \times Z_{2} \times Z_{2} \simeq D_{2h}$. The fact that $H_{X}$ and the $R_{d}$ commute holds true only for the
simple hole kinetic energy expression in equation (2). Extension of the symmetries to a general hole kinetic energy beyond the leading term requires combination of the operators $R_d$ with a reflection in position space, e.g., $(x, y, z) \rightarrow (-x, y, z)$ for $R_z$. Here, we can consider all reflections independently. In this way, the computational effort can be reduced by a factor $1/(4 \times 2^3) = 1/32$.

3.2. Crossover from large to small excitons

Formally, the crossover from large to small excitons can be achieved by scaling of the lattice constant $a$. Physically, the crossover from large to small excitons takes place, for example, when the hole mass increases. In the Wannier theory of continuum excitons, the excitonic Rydberg (i.e., binding energy) $R_X$ and excitonic radius $a_X$ are given by

$$R_X = R_{H} \frac{\mu}{\epsilon}, \quad a_X = a_0 \frac{\epsilon}{\mu},$$  \hspace{1cm} (17)

with the Rydberg energy $R_{H} \approx 13.6 \text{ eV}$ of the hydrogen atom and the Bohr radius $a_0 \approx 0.053 \text{ nm}$. These expressions involve only the reduced electron–hole mass $\mu = (m_e/m_h + m_0/m_e)^{-1}$, with the elementary electron mass $m_e$, and the dielectric constant $\epsilon$. Excitons become large for $m_h \rightarrow 0$, when $a_X \rightarrow \infty$ and $R_X \rightarrow 0$.

In figure 3 we show the exciton energy, radius, and mass as a function of variable hole mass $m_h = m_{hh} \equiv m_h$. The remaining model parameters are $m_e = m_0$, a large spin–orbit coupling $E_{so} = 300 \text{ meV}$ to separate the valence bands (since $m_{hh} = m_{hh}$ the precise value is not relevant), lattice constant $a = 1 \text{ nm}$, dielectric constant $\epsilon = 10$ and Coulomb length $\ell_C = 2a$, and for the sake of simplicity no exchange interaction $E_{ex} = 0$.

As was to be expected, significant deviations from Wannier theory occur as soon as $a_X \approx a$, which happens here for $m_h/m_0 \gtrsim 0.2$. As the exciton becomes small, its binding energy becomes much larger than the Wannier Rydberg $R_X$. We observe the shift of the energy of the even (1S, 2S, 3S) exciton states, and the splitting of the first excited (2S, 2P) state in the reduced (cubic instead of full rotational) symmetry that now applies [29]. Note that the odd exciton states are nearly unaffected by an energy shift (in the inset in figure 3(a) the two curves for the 2P state and the Wannier theory result can barely be distinguished). For $m_h \rightarrow \infty$, the Wannier Rydberg would converge to the value $R_0 m_e/(m_0 \epsilon^2)$, while the true binding energy approaches the (for given parameters larger) value $\epsilon^2/\ell_C + 6\epsilon$ (both values are marked in figure 3). The crossover to small excitons is accompanied by a strong enhancement $m_e/(m_0 + m_e) \gg 1$ of the exciton mass, as the exciton becomes increasingly immobile with decreasing radius. This situation, when kinetic energy is suppressed in favour of Coulomb attraction, is the regime of small Frenkel excitons.

![Figure 3](image-url)

**Figure 3.** Crossover from large to small excitons, as a function of the hole mass $m_h = m_{hh} \equiv m_h$. Top panel (a): energy of the first three even (1S, 2S, 3S) exciton states. The inset gives the energy of the 2S and 2P states. Middle panel (b): exciton radius in the lowest (1S) state. Bottom panel (c): exciton mass $m_n$ in the lowest (1S) state. In all panels, the dashed lines give the corresponding result obtained within the Wannier theory of hydrogen-like excitons.

4. Yellow and green excitons in the cuprous oxide

4.1. Model parameters

The model Hamiltonian (1) depends on eight parameters. Our logic for choosing these parameters is as follows: (i) use the lattice constant $a$ and electron mass $m_e$ known from experiment; (ii) trust the DFT valence bands, which fix the three parameters $m_h, m_{hh}, E_{so} \text{ of } H_{hole} + H_{so}$; (iii) obtain the remaining three parameters $\epsilon, \ell_C, E_{ex}$ from comparison with the experimental energies of the three lowest exciton states.

The successive steps of parameter fitting are documented in figures 4, 5, and lead to the model parameters in table 2. As
seen in figure 4, the DFT valence bands are very accurately reproduced at smaller values of |k| with the chosen parameters, including the non-parabolicity of the split-off band. The DFT bands cannot be reproduced by the simple model kinetic energy at larger values of |k| (not shown), where the interaction with other bands become significant, but as we will see below these parts of the band structure are not relevant for exciton formation. For a consistency check, we note that equation (7) gives $m_{1s\to}\approx 0.44 m_0$ for the hole mass in the split-off band. At the present level of treatment, and with respect to the experimental uncertainties, this is consistent with the experimental hole mass $m_h = 0.58 m_0$ from [12].

In figure 5, we first change $\epsilon$ and fit the energy of the 2P (ortho or para) exciton state, which does not depend on $\ell_C$ or $E_{ex}$, to the experimental value $E_{2P}^{\exp} = -23.6$ meV. Then, we change $\ell_C$ and fit the energy of the 1S para-exciton state, which does not depend on $E_{ex}$, to the experimental value $E_{1S(\text{para})}^{\exp} = -151$ meV. Finally, we change $E_{ex}$ and fit the energy of the 1S ortho-exciton state to the experimental value $E_{1S(\text{ortho})}^{\exp} = -139$ meV. Now, all model parameters have been determined.

### 4.2. Exciton mass and spectrum

With the model parameters from table 2 we can now compute, without further adjustments, the exciton mass and the remaining states in the exciton spectrum. A few numerical values are listed in table 3.

Figure 6 shows the mass in the 1S ortho-exciton state. The ortho-exciton state transforms as $\Gamma_7$ at $\mathbf{K} = 0$, which splits into a one-dimensional and a two-dimensional representation at finite $\mathbf{K}$ parallel to a lattice axis (on a $\Delta$-line). Therefore, the exciton mass can assume two different values $m_{\chi}^\circ$, $m_{\chi}^\circ$. They coincide only in the limit $E_{ex} = 0$, where they give the mass of the 1S para-exciton state (2.06 $m_0$), and evolve in opposite directions as $E_{ex}$ increases. At the value $E_{ex} = 666$ meV from the previous model parameter fit, we obtain $m_{\chi}^\circ = 2.56 \times 10^{-4}$, which is in reasonable agreement with the experimental data (see table 3). Note the significant influence of electron–hole exchange on the exciton mass that can be observed in figure 6: without exchange, the correct exciton mass could not be obtained.

The right panel (b) in figure 6 shows the dispersion of the heavy ($m_{\chi}^\circ$) and light ($m_{\chi}^\circ$) exciton state. We include this panel mainly to demonstrate that the exciton dispersion is parabolic at small $\mathbf{K}$, in contrast to the dispersion of a hole in the split-off valence band out of which this exciton state is formed.

Figure 7 shows the exciton spectrum. We compare the experimental data with our model computation and the
Figure 6. Left panel (a): exciton mass $m^*_X$, $m^*_Y$ in the 1S ortho-exciton state, as a function of the exchange energy $E_{\text{ex}}$. The value $m^*_X = m^*_Y$ in the limit $E_{\text{ex}} = 0$ corresponds to the exciton mass in the 1S para-exciton state. The vertical dashed line is placed at the value $E_{\text{ex}}^\text{fit} = 666$ meV obtained from the previous fit to the spectrum, which results in $m^*_X = 1.83m_0$, $m^*_Y = 2.56m_0$. The horizontal dashed lines give the value of $m_X = m_h + m_e$ that would be obtained within Wannier theory, and the range of experimental values $2.5 < m_X/m_0 < 2.7$. Right panel (b): exciton dispersion $E(k)$ as a function of momentum $k$ parallel to a lattice axis, for the 1S ortho-exciton state.

Table 3. Model results for excitons in the cuprous oxide Cu$_2$O, including binding energies of the lowest even exciton states, in comparison to experimental values.

|          | Experiment | Our model |
|----------|------------|-----------|
| 1S ortho exc. |            |           |
| Mass $m_X/m_0$ | 2.5–2.7    | 2.56 ($m_X^*$) |
| Radius     | —          | 1.62$a$ ≈ 7 Å |
| Even states |            |           |
| Binding energy | 151        | −fitted− |
| $E_{\text{bag}} - E_{\text{a}}$ (meV) | 139        | −fitted− |
|          | 45.1       | 46.62     |
|          | 34.2       | 33.84     |
|          | 17.6       | 17.52     |
|          | 11.73      | 11.53     |
|          | 10.17      | 9.80      |
|          | 9.98       | 9.75      |
|          | 8.97       | 8.90      |

Table 2. Model parameters for the cuprous oxide Cu$_2$O. The lattice constant $a = 0.42696$ nm and electron mass $m_e = 0.99 m_0$ (such that $t_e = 0.21$ eV in $E_{\text{electronic}}$) are taken from the experimental data given in table 1.

|          | DFT bands | Spectrum |
|----------|-----------|----------|
| Bare hole mass $m_{\text{bh}}$    | 0.16 $m_0$ |
| $t_1$    | 1.31 eV |
| Bare hole mass $m_{\text{bh}}$    | 3.10 $m_0$ |
| $t_2$    | 0.067 eV |
| Spin–orbit coup. $E_{\text{so}}$  | 0.128 eV |
| Dielectric constant $\epsilon$    | 6.94 |
| Coulomb length $\ell_C/a$         | 1.75 |
| Exchange energy $E_{\text{ex}}$   | 666 meV |

Figure 7. Exciton spectrum in the cuprous oxide, for even (left and central panels (a), (b)) and odd (rightmost panel (c)) exciton states up to $n \leq 5$. Each panel gives the experimental values (as in figure 1) in comparison to our model computation and the theoretical values from [13]. Note the different scales of the energy axes. The central panel (b) is the magnification of the upper part of the left panel (a). The spectrum from the continuum theory presented in [13]. Agreement with experiment is comparably good for both the lattice model and the continuum theory, although not perfect in either case. The maximal deviation is below 2 meV, which is about 1.5% of the excitonic Rydberg energy. Both the continuum theory and the lattice model correctly reproduce the shift of exciton lines relative to the Rydberg series, and the splitting of exciton states in the cubic symmetry.

Note how these results prove the relevance of the exciton lattice model: the only information about the exciton spectrum put into the model is derived from the energy of the lowest three (two even and one odd) exciton states. With this information, the model allows us to compute the entire spectrum with all deviations from the Rydberg series, and, as a second independent quantity, the exciton mass. The agreement between model computations and experiment is certainly good enough to warrant the conclusion that all relevant aspects of the physics of exciton with strong central-cell corrections are correctly captured by the lattice model.

5. Conclusion

Starting from a microscopic lattice model for exciton formation we study the effects of central-cell corrections on small-radius excitons, using the yellow exciton series in the cuprous oxide as the main example. Our study relates two, at first sight unrelated, quantities: the shift of even exciton states relative to the Rydberg series and the enhancement of the exciton mass.

The success of the lattice model partly stems from the fact that it incorporates central-cell corrections in a natural way, such that specific exciton properties follow without additional effort. Local corrections to the Coulomb attraction and non-parabolic terms in the kinetic energy are an intrinsic part of the model. A crucial ingredient is spin–orbit coupling, which immediately explains the peculiar dispersion of the highest valence band.

Not surprisingly, exchange interaction is equally crucial, as it is responsible for the splitting of para-exciton and ortho-exciton
states. It is probably surprising, however, that exchange interaction is also responsible (i) for pushing the ortho-exciton mass towards the correct experimental value, and (ii) for a strongly anisotropic ortho-exciton mass. In particular, the mass enhancement for ortho-exciton and para-exciton is not the same.

The anisotropy of the ortho-exciton mass in the lattice model does not result from a $k$-dependent exchange interaction—exchange is local in our model—but from the interplay of exchange interaction with the other terms in the Hamiltonian, including spin–orbit coupling. Furthermore, the mass anisotropy is not small: the ‘heavy’ and ‘light’ ortho-exciton mass differs by a factor 1.4. One consequence is that ortho-exciton states are no longer three-fold degenerate at finite $k$.

In this context, note that the para-exciton mass obtained from the lattice model (2.06 $m_0$) is distinctly smaller than recent experimental values (2.61 $m_0$ in [11]), although it somewhat agrees with smaller values cited elsewhere (2.2$m_0$ in [10]). Given the fact that our result for the ortho-exciton mass agrees nicely with the experiment, one should try to identify the origin of this discrepancy both theoretically and experimentally.

The lattice model in the form used here is kept deliberate simple, yet is easily extendable. This poses a natural question: is it possible to perfect the agreement between theory and experiment? The answer is, of course, yes. Reasonable extensions of the model would include additional terms in the hole kinetic energy (most importantly the terms with $e = (0, 1, 1)$ and $e = (1, 1, 1)$ in the notation of appendix A), a refined treatment of screening of the Coulomb attraction, and Coulomb terms that couple different orbitals.

Inclusion of additional terms in the Hamiltonian introduces additional parameters. kinetic energy terms could be fitted to DFT band structure calculations or deduced from comparison to experimental exciton data. The value $U(0)$ of the on-site Coulomb interaction could be estimated from the DFT orbitals, and screening could be included via model potentials. The present agreement between model and experiment is yet good enough to let us prefer the present simpler model over excessive parameter fitting.

As the principal usefulness of an exciton lattice model has now been established, several questions remain to be resolved. For example, we have not computed the oscillator strengths of one- or two-photon absorption, or the splitting of exciton states under mechanical stress [33, 34]. The computation of oscillator strengths requires some additional theoretical work, while the latter issue is immediately treatable by letting the kinetic energy terms become direction dependent. Closer examination of these and related questions should be a worthwhile activity in future investigations.

### Acknowledgements

This work was financed by Deutsche Forschungsgemeinschaft via AL1317/1-2 and SFB 652. AA is grateful for helpful discussions with H Knedlik.

### Table A1. Possible hole kinetic energy terms for the $\Gamma^+_+$ representation.

| $e$ | $|G|_k$ | $H_e$ |
|-----|------|------|
| 0 = $e_i = e_j < e_k$ | 8 | $a$ |
| 0 = $e_i < e_j = e_k$ | 4 | $a$ |
| 0 = $e_i < e_j < e_k$ | 2 | $a$ |
| 0 < $e_i = e_j = e_k$ | 6 | $a b$ |
| 0 < $e_i < e_j = e_k$ | 2 | $a d^*$ |
| 0 < $e_i < e_j < e_k$ | 1 | $a e^*$ |

with a translation vector $e \in \mathbb{Z}^3$ and a $3 \times 3$ matrix $H_e$.

### Appendix A. Higher order terms of the kinetic energy

The hole kinetic energy terms have the form

$$H_{\text{kin}}[e] = \frac{1}{|G|_k} \sum_{g \in O_k} \Gamma(g) H_e \Gamma(g)^{-1} T(g \cdot e),$$

with a translation vector $e \in \mathbb{Z}^3$ and a $3 \times 3$ matrix $H_e$.

The matrices $\Gamma(g)$ are given by the representation $\Gamma^+_+$ for the valence band orbitals, and $g \cdot e$ denotes the action of the group element $g \in O_h$ on a vector $e$. The prefactor $1/|G|_k$ with the isotropy group $G_e = \{ g \in O_h \text{ with } g \cdot e = e \}$ corrects for multiple counting.

Hermiticity of $H_{\text{kin}}[e]$ and the existence of an inversion $g_i \in O_h$ with $g_i \cdot e = -e$ implies $H_e = \Gamma(g_i) H_e^* \Gamma(g_i)^{-1}$. For the $\Gamma^+_+$ representation, $\Gamma(g_i) = 1$, such that the matrix $H_e$ itself must be Hermitian.

Furthermore, we must have $\Gamma(g) H_e \Gamma(g)^{-1} = H_e$ for every $g \in G_e$. Constraints on $H_e$ result from elements $g = 1$ of $G_e$. To enumerate the different cases, note that among the vectors $g \cdot e$ we can always find one vector whose entries are in ascending order $e_j \leq e_i \leq e_k$. Apart from the trivial case $e = 0$, we have to consider the six cases listed in table A1. Note that we do not list the case $0 < e_i = e_j < e_k$, since it is equivalent to the fifth case $(0 < e_i < e_j < e_k)$ up to relabeling of the coordinates.

### Appendix B. Hole dispersion

The hole Hamiltonian $H_{\text{hole}} + H_{\text{so}}$ commutes with each operator $R_d^2 = (2L_d^2 - 1) \otimes (2L_d^2 - 1) S^k_d$, for $d = x, y, z$, such that its eigenstates can be classified according to one of these operators, say, for $d = z$. Note that, in contrast to the operators $R_d$ from section 3, the operators $R_d^2$ do not commute.
among themselves. The corresponding Bloch Hamiltonian, parametrized by the hole momentum $k_h$, splits in two $3 \times 3$ blocks: one in the subspace of states $|\vec{x}, \uparrow\rangle, |\vec{y}, \uparrow\rangle, |\vec{z}, \downarrow\rangle$ to the eigenvalue $+1$ of $R^h$, the other in the subspace of states $|\vec{x}, \downarrow\rangle, |\vec{y}, \downarrow\rangle, |\vec{z}, \uparrow\rangle$ to the eigenvalue $-1$ of $R^h$. Both blocks are related by the unitary transformation induced by $R^h$ (or $R^h$), and have the same eigenvalues. Therefore, the hole bands are two-fold degenerate.

For $k_x = k_y = 0$, diagonalization of one of the $3 \times 3$ matrices gives the eigenvalues

\[
E_1(k_z) = E_v + 2t_2(c_z - 1) + 2\lambda,
E_2(k_z) = E_v + (h + t_2)(c_z - 1) - \lambda - \xi,
E_3(k_z) = E_v + (h + t_2)(c_z - 1) - \lambda + \xi,
\]

with $c_d = \cos \theta(k_h \cdot \mathbf{e}_d)$, $\lambda = -\frac{E_0}{6}$, $E_v = 2t_1 + 4t_2$, and $\xi^2 = (t_1 - t_2)(c_z - 1) + \lambda^2 + 8\lambda E_v$. Equivalent expressions hold for the $y, z$ direction. At $k = 0$, with $E_0 = E_2(0) = E_v - (1/3)E_{10}$, $E_3(0) = E_v + (2/3)E_{10}$ the energy level splitting is $E_{10}$.

### Appendix C. Symmetrized basis states

The local degrees of freedom of the Hamiltonian $H_X$ in equation (1) are labelled by the orbital, hole spin, and electron spin indices as $|\vec{d}, s_h, s_e\rangle$ with $d = x, y, z$ and $s_h, s_e = \uparrow, \downarrow$ (see section 2). Symmetrized basis states with respect to the operators $R_d$ from section 3 are given in table C1. They belong to four different sets depending on the eigenvalues $\pm 1$ of the $R_d$ operators. Since $R_x R_y R_z = 1$ only the listed eigenvalue combinations can occur. The first set gives the para-exciton states, and contains only basis states with non-zero hole–electron spin $S^h + S'$, the two remaining sets give the ortho-exciton states.

Spin–orbit and exchange coupling connect only the three states within each group. For each group, spin–orbit coupling is given by the same $3 \times 3$ matrix

\[
\frac{2}{\hbar} \mathbf{I} \cdot \mathbf{S}^h = \begin{pmatrix}
0 & i & 1 \\
-i & 0 & i \\
1 & -i & 0
\end{pmatrix},
\]

while exchange interaction is given by a diagonal matrix with entries as listed in the last column of table C1.

### ORCID iDs

A Alvermann @ https://orcid.org/0000-0002-0367-2721

### References

[1] Knox R 1963 Theory of Excitons (New York: Academic)
[2] Moskalenko S A and Smoke D W 2000 Bose–Einstein Condensation of Excitons and Bieions (Cambridge: Cambridge University Press)
[3] Kazimierczuk T, Fröhlich D, Scheel S, Stolz H and Bayer M 2014 Nature 514 343
[4] Kavoulakis G M, Chang Y-C and Baym G 1997 Phys. Rev. B 55 7593
[5] Fröhlich D, Kenklines R, Uhlein C and Schwab C 1979 Phys. Rev. Lett. 43 1260
[6] Uhlein C, Fröhlich D and Kenkline S R 1981 Phys. Rev. B 23 2731
[7] Thewes J, Heckötter J, Kazimierczuk T, Altmann M, Fröhlich D, Bayer M, Semina M A and Glazov M M 2015 Phys. Rev. Lett. 115 027402
[8] Schweizer F, Main J, Feldmaier M, Wunner G and Uhlein C 2016 Phys. Rev. B 93 195203
[9] Schöne F, Krüger S-O, Grünwald P, Stolz H, Scheel S, Altmann M, Heckötter J, Thewes J, Fröhlich D and Bayer M 2016 Phys. Rev. B 93 075203
[10] Jörgen M, Fleck T, Klingshirn C and von Baltz R 2005 Phys. Rev. B 71 235210
[11] Brandt J, Fröhlich D, Sandfort C, Bayer M, Stolz H and Naka N 2007 Phys. Rev. Lett. 99 217403
[12] Naka N, Akimoto I, Shirai M and Kan no K O K 2012 Phys. Rev. B 85 035209
[13] Schweiner F, Main J, Wunner G and Uhlein C 2017 Phys. Rev. B 95 195201
[14] Luttinger J M 1956 Phys. Rev. 102 1030
[15] Balderschi A and Lipari N C 1971 Phys. Rev. B 3 439
[16] Balderschi A and Lipari N O 1973 Phys. Rev. B 8 2697
[17] Balderschi A and Lipari N O 1974 Phys. Rev. B 9 1525
[18] Suzuki K and Hensel J C 1974 Phys. Rev. B 9 4184
[19] Lipari N O and Altarelli M 1977 Phys. Rev. B 15 4883
[20] Altarelli M and Lipari N O 1977 Phys. Rev. B 15 4898
[21] French M, Schwartz R, Stolz H and Redmer R 2009 J. Phys.: Condens. Matter 21 015502
[22] Koster G, Dimmock J, Wheeler R and Stutz H 1963 Properties of the Thirty-Two Point Groups (Cambridge, MA: MIT Press)
[23] Elliott J R 1961 Phys. Rev. 124 340
[24] Dahl J and Switendick A 1966 J. Phys. Chem. Solids 27 931
[25] Pollmann J and Büttner H 1977 Phys. Rev. B 16 4480
[26] Saad Y 1992 Numerical Methods for Large Eigenvalue Problems (Manchester: Manchester University Press)
[27] Sorensen D C 2002 Acta Numer. 11 519
[28] Alvermann A, Littlewood P B and Fehske H 2011 Phys. Rev. B 84 035126
[29] Cho K 1976 Phys. Rev. B 14 4463
[30] Dasbach G, Fröhlich D, Stolz H, Klieber R, Suter D and Bayer M 2003 Phys. Rev. Lett. 91 107401
[31] Dasbach G, Fröhlich D, Klieber R, Suter D, Bayer M and Stolz H 2004 Phys. Rev. B 70 045206
[32] Dasbach G, Fröhlich D, Stolz H, Klieber R, Suter D and Bayer M 2005 Phys. Status Solidi c 2 886
[33] Trebin H-R 1977 Phys. Status Solidi b 81 527
[34] Waters R G, Pollak F H, Bruce R H and Cummins H Z 1980 Phys. Rev. B 21 1665