Bound Exciton and Hole: An Exactly Solvable Three-Body Model in Any Number of Dimensions

R. Schilling[a] and D. C. Mattis
Department of Physics, University of Utah, Salt Lake City, Utah 84112
(Received 17 May 1982)

A three-body problem, concerning two holes in a nondegenerate valence band and a single electron in a conduction band, with strong short-range interactions, is solved exactly in any number of dimensions. The binding depends nontrivially on the ratio of the valence to conduction bandwidths (i.e., on the inverse ratio of their effective masses) and always vanishes when this ratio exceeds 1. The effective mass of the "trion" bound state also depends sensitively on this ratio and on dimensionality.

PACS numbers: 71.35.+z

There exist so few exactly solvable three-body problems in quantum mechanics that the extremely simplified model we have recently solved may be of general interest in the physics community. It consists of a very tightly bound exciton interacting with one extra hole, on a lattice in any number d of dimensions. Generalizations to the four-, five-,... body problems also seem possible in some instances.

The special role of excitons in solid state and in the optical properties of matter has been recognized in an extensive literature. Recent reviews by Rice[1] and Hensel, Phillips, and Thomas[2] deal with the Wannier exciton (the "large" exciton, consisting of a bound electron-hole pair) suitable for a number of semiconductors, whereas the Frenkel exciton (the "small" exciton, localized within an atomic distance or so) is more appropriate in other semiconductors and in organic materials. The distinction is drawn in Knox's text.[3] In the present model, a parameter U enables us to proceed continuously from the one to the other; the *three-body problem* is explicitly solvable, however, *only* in the Frenkel limit $U=\infty$. It was Lampert who was probably the first to extend the concept of an excitonic bound state to more than two particles, which he denoted "effective-mass particle complexes."[4] Despite the approximate nature of the effective-mass approximation (EMA) (which limits it to binding energies small compared with band gaps, bandwidths, etc.), and its equivalence to the continuum limit of the lattice (which limits it to length scales large compared with an atomic distance), it maps onto ordinary molecular physics of few-body systems and has allowed interesting results to be obtained on the existence of bound states for the exciton plus hole,[5] the exciton plus electron,[6,7] and higher complexes.[1] The exciton plus hole (or electron) was given the name "trion" by Thomas and Rice[8] in 1977, and shortly thereafter experimental evidence for this complex was produced, e.g., in the work of Stébé *et al.*[9] Related theories exist in other fields of physics, concerning the existence of positronium molecular ions $(2e^+,e^-)$,[10,11] the negative hydrogen ion[12-14] found by Hill[13,14] to have but a single bound state, and the lack of a positron-hydrogen bound state[15] for an infinitely massive proton as well as for a proton of the usual mass.[16] But when the EMA breaks down, as it does in our solvable limit, these theories are clearly of limited use.

We start with the two-band Hubbard model first discussed by Anderson,[17] and applied by Falicov and Kimball,[18] Ramirez, Falicov, and Kimball,[19] and Doniach, Roulet, and Fisher[20] to the study of excitons. Egri[21,22] subsequently used the same model in one dimension (1D) to unify the picture of Wannier versus Frenkel excitons; in this work he included an exciton-hopping matrix element, which is a two-center, four-orbital Coulomb integral. In the present Letter, we shall neglect this matrix element, and thus our Frenkel excitons are immobile in the $U=\infty$ limit, as we shall see. We also neglect the spin of the particles. In this way, we achieve a solvable model not limited to 1D. The model Hamiltonian is

$$H=H_C+H_V+H_U+H_B\,, \tag{1}$$

where

$$H_C=-C\sum c_i{}^*c_{i+\delta}\,, \tag{1a}$$

$$H_V=V\sum v_i{}^*v_{i+\delta}\,, \tag{1b}$$

$$H_U=U\sum v_i{}^*v_i\,c_i{}^*c_i\,, \tag{1c}$$

$$H_B=\tfrac{1}{2}B\sum(c_i{}^*c_i-v_i{}^*v_i)\,, \tag{1d}$$

in which i, $i+\delta$ are nearest-neighbor sites, C and V are the conduction- and valence-bandwidth parameters [in EMA, the conduction-band effective mass is $m_e^* = (2Ca^2)^{-1}$ and the valence-band hole effective mass is $m_h^* = (2Va^2)^{-1}$ in units where $\hbar = 1$ and a is the lattice parameter], U is the electronic repulsion parameter that we shall take to the limit $U = \infty$, and B is related to the energy gap as follows: $E_{\text{gap}} = B - z(C+V)$, with z the coordination number of the lattice ($z = 2d$ for a d-dimensional simple cubic).

If, as we shall assume, B is sufficiently large, the valence band is fully occupied with (spinless) electrons (anticommuting creation and destruction operators v_i^* and v_i) and the conduction band (c_i^* and c_i) is empty in the ground state. Starting from this, we introduce n electrons into the conduction band and p holes into the valence band (creation and destruction operators v_i and v_i^* for the holes), setting $n = 1$ and $p = 2$ for the problem of current interest.

In the limit of $U = \infty$, an electron in the conduction band *must* be on the same site as a hole in the valence band; this is also known as the Frenkel limit. In the absence of explicit exciton hopping matrix elements discussed above, the Frenkel exciton is immobile, as proved in Fig. 1. Two or more excitons cannot occupy the same site, although they can be neighbors without incurring any interactions. Therefore, the many-exciton states are trivial in the present model. The introduction of a single extra hole changes the situation dramatically. Even when the exciton is immobile, the bound *trion* can acquire a finite mobility related to the connectivity of the lattice. Although in 1D the trion remains immobile, in 2D or 3D we shall see that it acquires an effective mass which depends on the ratio V/C. In fact, the very existence of the trion bound state depends on this ratio, which we shall denote $|V/C| \equiv \nu$ for convenience. In 1D and 2D there is a bound state whenever $\nu < 1$, whereas in 3D $\nu \lesssim \frac{1}{3}$ is required (amounting to $m_e^*/m_h^* \lesssim \frac{1}{3}$ in the EMA). We shall give the explanation following Eq. (12). In any event, such a dependence on dimensionality and band structure has not been

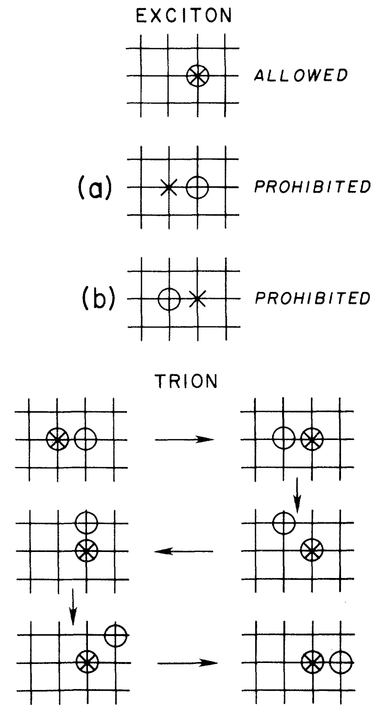

FIG. 1. Illustrating the immobility of the exciton [configurations such as (a) and (b) are prohibited when $U = \infty$] and the motion of the trion in two or higher dimensions under the same rules. An electron in the conduction band is indicated by a cross, and a hole by a circle.

noted in the literature heretofore, and lies outside the realm of the EMA. Nor has it been noted that a hole can even bind *two* excitons—a five-body problem that we have explicitly solved in 1D and which is capable of being solved in 2D or higher, as we show in the expanded version of the present paper.[23]

We introduce the vector distance $\vec{r}$ between the extra hole and the exciton, and the wave functions $F_1(\vec{r})$ and $F_2(\vec{r})$ to distinguish when the electron is on top of hole 1 or 2. Later, we ensure that the wave function is antisymmetric under the exchange of the two holes, to satisfy the Pauli principle. We find the Schrödinger equation for $F(\vec{r})$:

$$-V\sum_{\vec{\delta}} \exp(\tfrac{1}{2} i\vec{K}\cdot\vec{\delta})F_{1,2}(\vec{r}+\vec{\delta}) - C\sum_{\vec{\delta}} F_{2,1}(-\vec{\delta})\delta_{\vec{r},\vec{\delta}} + \hat{U}F_{1,2}(\vec{r})\delta_{\vec{r},\vec{0}} = EF_{1,2}(\vec{r}), \tag{2}$$

where E_0 is the ground-state energy, and $E = E_{\text{trion}} - E_0 - \frac{3}{2}B$ is the energy eigenvalue. $\hat{U}$ is a projection operator, not to be confused with U. Taking the $\hat{U} \to \infty$ limit causes the unphysical amplitude $F_{1,2}(0)$ to vanish. The lowest (bonding) bound state is always given by $F_1(\vec{r}) = +F_2(\vec{r}) = F(\vec{r})$. The antibonding case $F_1(\vec{r}) = -F_2(\vec{r})$ produces a second set of bound states which can, however, be related to the former at the same value of total momentum $\vec{K}$ by taking $C \to -C$ in the various formulas.

By using a straightforward Green's-function method, we reduce Eq. (2) to a set of $2d+1$ equations:

$$-[\hat{U}G(0,0)+1]F(0) + C\sum_{\vec{\delta}'} G(0,-\vec{\delta}')F(\vec{\delta}') = 0, \tag{3a}$$

$$-\hat{U}G(\vec{\delta},0)F(0) + C\sum_{\vec{\delta}'} G(\vec{\delta},-\vec{\delta}')F(\vec{\delta}') - F(\vec{\delta}) = 0, \tag{3b}$$

with $\vec{\delta}$ taking on any of the 2D nearest-neighbor values. While $F(0)=0$ in the $\hat{U}\to\infty$ limit, the product $\hat{U}F(0)$ is finite and must be retained. The Green's functions which appear in (3) are the familiar lattice functions of a cubic lattice:

$$G(\vec{r},\vec{r}') = (2\pi)^{-d}\int d^d q \frac{\exp[i\vec{q}\cdot(\vec{r}-\vec{r}')]}{\epsilon_{\vec{K}}(\vec{q})-E}$$

with

$$\epsilon_{\vec{K}}(\vec{q}) = -V\sum_{\vec{\delta}}\cos[\tfrac{1}{2}(\vec{K}+\vec{q})\cdot\vec{\delta}].$$

Equations (3) must be solved numerically in general, although along the axis of high symmetry $\vec{K}=K(1,1,\ldots)$ they simplify considerably:

$$f(E)+g(E)=0 \text{ (for } d=1 \text{ only)}, \qquad (4)$$

or

$$f(E)\cos K + g(E)=0 \ (d\geq 2), \qquad (5)$$

where $f(E)$ and $g(E)$ are simple combinations of the G's,

$$f(E)=\frac{1}{C}\left[G_0(G_0-G_2)+\frac{1}{2dV^2}(1+EG_0)\right], \qquad (6)$$

$$g(E)=\frac{1}{2dV^2}(G_0-G_2)(1+EG_0)+\frac{1}{C^2}G_0, \qquad (7)$$

with

$$G_m \equiv G_m(E) = (2\pi)^{-d}\int d^d q\, d^{-1}\sum_{j=1}^{d}\cos mq_j/[\epsilon_0(\vec{q})-E] \qquad (8)$$

for $m=0,2$.

Depending on ν, Eqs. (4) and (5) may yield bound states for the various $\vec{K}$ in the Brillouin zone. We investigate three points of high symmetry.

(1) $K=0$.—At $K=0$, (4) and (5) factor:

$$\left[G_0-G_2+\frac{1}{C}\right]\left[G_0+\frac{C}{2dV^2}(1+EG_0)\right]=0. \qquad (9)$$

Because the first factor is nonzero for $C>0$, we can divide it out to obtain

$$G_0(E)+\frac{1}{E+2dV^2/C}=0, \qquad (10)$$

with G_0 given in (8). There is a bound state below the continuum (that is, there is a solution $E<-2dV$) only if

$$\nu=|V/C|<\nu_c, \qquad (11)$$

where the critical ν_c is obtained by setting E at the threshold value $-2dV$ in (10). $G_0(-2dV)$ is related to Watson's integral[24] W_d, and in fact (10) yields

$$\nu_c = 1-1/W_d, \qquad (12)$$

where

$$W_d=(2\pi)^{-d}\int d^d q\left(1-d^{-1}\sum_{j=1}^{d}\cos q_j\right)^{-1}$$

is infinite for $d=1,2$ and finite for $d\geq 3$. $W_3=1.516\ldots$, and W_d monotonically decreases to $W_d\to 1$ as $d\to\infty$. As a consequence, $\nu_c\to 0$ as $d\to\infty$.

These results can be understood qualitatively by analogy with ordinary potential wells in continuum quantum mechanics. In binding the extra hole, the electron lowers its motional energy by $O(C)$, at a cost $O(V)$ required to localize the hole. Thus the potential well is attractive only if $V\leq C$. While an arbitrarily weak potential well binds in 1D, and 2D is marginal, in 3D or higher the well must become deeper with increasing d if a bound state is to be retained; hence ν_c must decrease to 0 as $d\to\infty$. [Note that the range of the forces is a constant $O(a)$ for all d.]

Regrettably there is no trion bound state in 1D for $V=C$. In that limit, our model reduces to the Hubbard model solved by means of a Bethe's *Ansatz* by Lieb and Wu,[25] and although we are able to obtain solutions for arbitrary values of n and p, none of them can be bound states! Nonetheless, for $\nu<1$ we have solved some cases of small n and p and found the bound states. [The interesting case of $n=2$, $p=3$ (an extra hole binding two excitons) is reported elsewhere.[23]]

(2) $K=\frac{1}{2}\pi$.—At midzone, (5) yields

$$\frac{1}{2dV^2}(G_0-G_2)(1+EG_0)+\frac{G_0}{C^2}=0. \qquad (13)$$

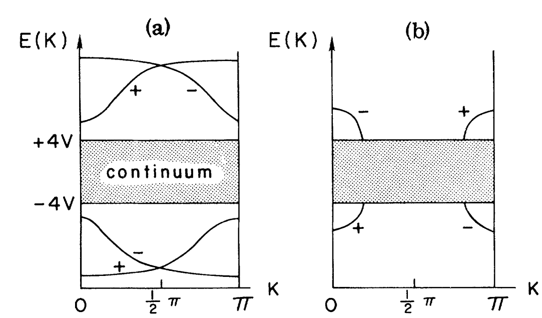

FIG. 2. Bound states of the trion and their dispersion in 2D. Energy is relative to that of an exciton; continuum is that of a free hole. (a) Narrow valence bandwidth $|V/C|=\nu<\nu_c''$, with + labeling bonding and − antibonding branches. (b) $\nu_c'<\nu<1$. The momentum K is along the [1,1] direction.

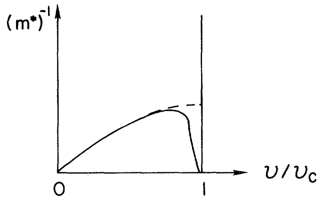

FIG. 3. Qualitative dependence of inverse trion effective mass $1/m^*$ at $K=0$ as a function of ν in 2D, 3D, and 4D (solid curve) and $d \geqslant 5$ (dashed curve). In 1D, where E is independent of K, $1/m^*$ is always zero.

This has a bound state for $\nu < \nu_c'$, with the new threshold ν_c' (smaller than ν_c) given by

$$\nu_c' = [d^{-1}S_d(1 - 1/W_d)]^{1/2} \qquad (14)$$

with S_d a variant of Watson's integral,

$$S_d = (2\pi)^{-d}\int d^d q \left(d^{-1}\sum_{j=1}^{d} \sin^2 q_j\right) \Big/ \left(1 - d^{-1}\sum_{j=1}^{d} \cos q_j\right). \qquad (15)$$

(3) $K=\pi$.—Equation (5) again factors as did (9), with C replaced by $-C$. The second factor is now nonzero, but the first factor can vanish, yielding the bound-state energy. The threshold is

$$\nu_c'' = d^{-1}S_d \qquad (16)$$

and is, in turn, smaller than ν_c'.

The dispersion (variation of E with K) is shown in Fig. 2 for all the bound states in 2D at two different values of ν. Expanding $E(\vec{K}) = E(0) + \vec{K}^2/2m^* + \ldots$ we can define m^*.

The effective mass m^* of the trion is calculable at $K=0$ from a simple formula,

$$m^* = \frac{1}{f(E)}\frac{d}{dE}[f(E) + g(E)]_{K=0}. \qquad (17)$$

When the numerator and denominator both diverge, the ratio must be taken carefully. The result is plotted in Fig. 3. The derivation, and further results, will be justified at length in a later publication.[23]

The support of the U. S. Air Force Office of Scientific Research and the National Science Foundation under Grant No. DMR81-06223 is gratefully acknowledged.

[(a)]On leave from Institute of Physics, University of Basel, Basel, Switzerland.

[1]T. M. Rice, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1977).

[2]J. C. Hensel, T. G. Phillips, and G. A. Thomas, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1977).

[3]R. S. Knox, *Theory of Excitons* (Academic, New York, 1963).

[4]M. A. Lampert, Phys. Rev. Lett. 1, 450 (1958).

[5]B. Stébé and G. Munschy, Solid State Commun. 17, 1051 (1975).

[6]B. Gerlach, Phys. Status Solidi (b) 63, 459 (1974).

[7]G. Munschy and B. Stébé, Phys. Status Solidi (b) 64, 213 (1974).

[8]G. A. Thomas and T. M. Rice, Solid State Commun. 23, 359 (1977).

[9]B. Stébé, T. Sauder, M. Certier, and C. Comte, Solid State Commun. 26, 637 (1978).

[10]J. A. Wheeler, Ann. N. Y. Acad. Sci. 48, 219 (1946).

[11]E. A. Hylleraas, Phys. Rev. 71, 491 (1947).

[12]E. A. Hylleraas, Z. Phys. 60, 624 (1930).

[13]R. N. Hill, Phys. Rev. Lett. 38, 643 (1977).

[14]R. N. Hill, J. Math. Phys. (N.Y.) 18, 2316 (1977).

[15]I. Aronson, C. J. Kleinmann, and L. Spruch, Phys. Rev. A 4, 841 (1971).

[16]E. A. G. Armour, Phys. Rev. Lett. 48, 1578 (1982).

[17]P. W. Anderson, *Concepts in Solids* (Benjamin, New York, 1967).

[18]L. M. Falicov and J. C. Kimball, Phys. Rev. Lett. 22, 997 (1969).

[19]R. Ramirez, L. M. Falicov, and J. C. Kimball, Phys. Rev. B 2, 3383 (1970).

[20]S. Doniach, B. J. Roulet, and M. E. Fisher, Phys. Rev. Lett. 27, 262 (1971).

[21]I. Egri, J. Phys. C 12, 1843 (1979).

[22]I. Egri, J. Phys. C 15, L461 (1982).

[23]R. Schilling and D. C. Mattis, to be published.

[24]G. N. Watson, Quart. J. Math. 10, 266 (1939).

[25]E. H. Lieb and F. Y. Wu, Phys. Rev. Lett. 20, 1445 (1968).