

Spin delocalization in phosphorus donor pairs in silicon

J. H. Pifer

Department of Physics and Astronomy, Rutgers University, New Brunswick, New Jersey 08903

(Received 31 August 1984; revised manuscript received 14 June 1985)

In dilute samples of silicon doped with phosphorus, the spins of exchange-coupled donor electrons are Anderson localized on the P sites by random fields arising from the hyperfine interaction with nearby ^{29}Si nuclei, and by the random up-down orientation of the ^{31}P hyperfine field which is an order of magnitude larger than the ^{29}Si fields. In this paper we exactly solve the problem of an exchange-coupled donor pair subject to random hyperfine fields. We find that spin delocalization occurs in a continuous manner as the exchange constant J becomes comparable to the random field. Delocalization occurs for a much smaller J for pairs where the ^{31}P nuclear spins are parallel than in the case of nonparallel pairs. In the absence of random fields, delocalization due to J is limited by spin-lattice relaxation. We calculate how the electron-resonance spectrum of the pair changes as J increases and compare the exact results with Anderson's predictions from the theory of exchange narrowing.

I. INTRODUCTION

The outermost electron on a phosphorus impurity in silicon is loosely bound with a binding energy of 0.044 eV and an effective Bohr radius $\approx 17 \text{ \AA}$.¹ When two donors are close enough, there is appreciable electron overlap and a spin-exchange interaction $J\mathbf{S}_1 \cdot \mathbf{S}_2$ results. This exchange should produce delocalization of the spin for any nonzero value of J , even in the presence of a large external field ($\sim 3000 \text{ Oe}$ for a typical electron-resonance experiment) since a mutual spin flip costs no energy. Anderson² has shown, however, that spin localization persists when random fields are present since a mutual flip can only occur for spins for which there is an energy match in the random fields. For small enough J the probability of finding a match drops off so rapidly with distance that spin diffusion cannot occur.

In Si:P random fields are produced by the hyperfine coupling of the electron to both the ^{31}P nucleus ($I = \frac{1}{2}$, 100% abundant) and the ^{29}Si nucleus ($I = \frac{1}{2}$, 4.7% abundant). The P hyperfine field has the same magnitude ($A/2 = 21 \text{ Oe}$) for all donor electrons but is randomly up or down relative to the external field. The field due to the ^{29}Si is the sum of the hyperfine fields of ^{29}Si nuclei in various neighboring lattice sites; its root-mean-square value is $\Delta H_{\text{rms}} = 1.4 \text{ Oe}$.³ The Anderson criterion for the onset of localization is that there will be no spin diffusion when J is less than a critical J_c whose magnitude is comparable to the width of the distribution of random fields. Since in Si:P there are two random fields which differ by a factor of 10 in magnitude, we expect a more complicated process of delocalization.

Spin delocalization is independent of charge delocalization, which occurs near the onset of the metal-insulator transition at $3.8 \times 10^{18} \text{ P atoms/cm}^3$ in Si:P. It requires a much smaller critical exchange and can occur at very dilute P concentrations, where the charge is totally localized.

Electron paramagnetic resonance (EPR) is the most

direct way to study spin delocalization. In this paper we discuss the onset of spin delocalization and its effect on the EPR signal for the simplest possible case, a donor pair. We exactly solve the problem of a pair of exchange-coupled electrons in the presence of random fields in order to more precisely determine the manner in which spin delocalization occurs for $J \approx J_c$. The role of spin-lattice-relaxation processes (T_1) in delocalization is discussed. We then calculate the effect of delocalization on the EPR signal of the pair and discuss the applicability of exchange-narrowing theories. In a future paper we will extend these ideas to larger exchange-coupled clusters.

II. EXCHANGE-COUPLED PAIRS

A. Exact solution

We consider two electrons localized on nearby P ions and coupled by an exchange interaction arising from direct overlap of their wave functions. Each electron experiences a net hyperfine interaction (B_1 or B_2) due to the ^{29}Si nuclei randomly occupying various neighboring Si sites. Since the nuclear relaxation times are very long, we assume that B_1 and B_2 , as well as the ^{31}P nuclear hyperfine field A , are static quantities and make the secular approximation, keeping only the diagonal part of the hyperfine interactions. This is valid as long as $J < \gamma H$. The two-spin Hamiltonian is then

$$\begin{aligned} h^{-1}\mathcal{H} = & \gamma H(S_{1z} + S_{2z}) + A(I_{1z}S_{1z} + I_{2z}S_{2z}) \\ & + B_1S_{1z} + B_2S_{2z} + J\mathbf{S}_1 \cdot \mathbf{S}_2. \end{aligned} \quad (1)$$

The nuclear quantum numbers m_1 and m_2 are good quantum numbers for \mathcal{H} ; the electron numbers M_1 and M_2 are good only if $J = 0$. We have exactly diagonalized Eq. (1) and obtained the wave functions ϕ_n as linear combinations of the product spin wave functions $|M_1M_2m_1m_2\rangle$; see the Appendix. The eigenfrequencies ω_n are given in Fig. 1. We follow as closely as possible

STATE NO.	FREQUENCY
1	$\gamma H + \frac{J}{4} + \frac{1}{2} (A+B_1+B_2)$
2,3	$\gamma H + \frac{J}{4} + \frac{1}{2} (B_1+B_2)$
4	$\gamma H + \frac{J}{4} + \frac{1}{2} (A-B_1-B_2)$
5,9	$-\frac{J}{4} + \frac{1}{2} \sqrt{J^2 + (B_1-B_2)^2}$
7	$-\frac{J}{4} + \frac{1}{2} \sqrt{J^2 + (A-B_1+B_2)^2}$
8	$-\frac{J}{4} + \frac{1}{2} \sqrt{J^2 + (A+B_1-B_2)^2}$
16	$-\gamma H + \frac{J}{4} + \frac{1}{2} (A-B_1-B_2)$
14,15	$-\gamma H + \frac{J}{4} - \frac{1}{2} (B_1+B_2)$
13	$-\gamma H + \frac{J}{4} - \frac{1}{2} (A+B_1+B_2)$
6,10	$-\frac{J}{4} - \frac{1}{2} \sqrt{J^2 + (B_1-B_2)^2}$
11	$-\frac{J}{4} - \frac{1}{2} \sqrt{J^2 + (A+B_1-B_2)^2}$
12	$-\frac{J}{4} - \frac{1}{2} \sqrt{J^2 + (A-B_1+B_2)^2}$

FIG. 1. Eigenfrequency diagram of an exchange-coupled phosphorus donor pair subject to random fields B_1 and B_2 . The level-numbering scheme follows that of Cullis and Marko (Ref. 4).

the labeling scheme Cullis and Marko⁴ used in their solution of this problem without the random fields B_1 and B_2 .

B. Spin delocalization

If there are a large number of interacting spins, delocalization is defined by placing one up spin localized on a site with all other spins down. Delocalization occurs if a nonzero probability develops for finding the up spin at an arbitrarily distant site.

In the present case we have only two sites. We start the system at $t=0$ in a well-defined spin state $\psi_0 = | + - m_1 m_2 \rangle$, which is a linear combination of the ϕ_n , $\psi_0 = \sum_n a_n \phi_n$. The time evolution of the initial wave function is then

$$\psi(t) = \sum_n a_n \phi_n e^{-i\omega_n t}. \quad (2)$$

Rather than focus on the "distant" site, we calculate the spin expectation value on the starting site,

$$\langle S_{1z} \rangle = \langle \psi(t) | S_{1z} \psi(t) \rangle.$$

Since this oscillates in time, the quantity of interest is its time average $\langle S_{1z} \rangle$, which equals $\frac{1}{2}$ for localization on site 1 and 0 for complete delocalization, where the up and down spins spend equal time on site 1. We consider two cases:

(1) $m_1 = m_2$. For $\psi_0 = | + - + + \rangle$,

$$\langle S_{1z} \rangle = \frac{1}{2} \left[\frac{y^2}{1+y^2} + \frac{1}{1+y^2} \cos[2(\omega_5 - \omega_6)t] \right], \quad (3)$$

where $y = (B_1 - B_2)/J$ and, from Fig. 1,

$$\omega_5 - \omega_6 = [J^2 + (B_1 - B_2)^2]^{1/2}.$$

The time average,

$$\langle S_{1z} \rangle = y^2/2(1+y^2),$$

is $\frac{1}{2}$ for $J=0$ and 0 for $J=\infty$ or $B_1 - B_2 = 0$. 50% delocalization ($\langle S_{1z} \rangle = 0.25$) occurs for $J = |B_1 - B_2|$.

(2) $m_1 \neq m_2$. For $\psi_0 = | + - - + \rangle$,

$$\langle S_{1z} \rangle = \frac{1}{2} \left[\frac{z^2}{1+z^2} + \frac{1}{1+z^2} \cos[2(\omega_7 - \omega_{11})t] \right], \quad (4)$$

where $z = (A + B_2 - B_1)/J$ and

$$2(\omega_7 - \omega_{11}) = [J^2 + (A + B_1 - B_2)^2]^{1/2} + [J^2 + (A - B_1 + B_2)^2]^{1/2}.$$

Again, $\langle S_{1z} \rangle$ is $\frac{1}{2}$ for $J=0$ and 0 for $J=\infty$ or $B_1 - B_2 = 0$, but 50% delocalization occurs for

$$J = A + B_2 - B_1 = (B_2 + A/2) - (B_1 - A/2).$$

There are several important points to note about these results. First, delocalization is a continuous process; $\langle S_{1z} \rangle$ varies continuously between $\frac{1}{2}$ and 0 as J increases. The value of $\langle S_{1z} \rangle$ for which localization effects are important depend on the physical quantity one is considering and might, for example, be very different for an EPR experiment than for a NMR. Second, the nature of the delocalization for the two above cases is really the same, with the general rule being that delocalization occurs when J becomes comparable to the difference in (random) fields at the two sites. This is simply Anderson's criterion.² Note also that if the fields at the two sites are identical [$B_1 = B_2$ in case (1)], then complete delocalization occurs for any nonzero value of J ; localization does not occur (for infinite T_1) without random fields to shift the energy of the exchange-coupled sites. The third point to note is that the spin density at a given site oscillates at a frequency equal to the difference in the eigenfrequencies of the eigenstates required to form ψ_0 . In the limit of complete delocalization this frequency is $2J$.

C. Effect of spin-lattice relaxation

Anderson's localization theory does not include energy interchange with the surroundings. Likewise we have not included spin-lattice relaxation in our calculation of the delocalization. It is important to do so since a T_1 process changes the energy and destroys the phase coherence of the wave function. The time evaluation of ψ , Eq. (2), is terminated by a spin flip and it makes no sense to talk of delocalization at longer times. We can include T_1 by taking our time average of $\langle S_{1z} \rangle$ only over a time interval equal to T_1 .

An immediate consequence of including T_1 is that spin

TABLE I. Donor-concentration dependence of delocalization parameters in the absence of random fields.

n ($\times 10^{16}$ P atoms/cm ³)	$1/T_1$ (Hz) ^a	J_{10} (Hz)	r_{10} (Å)	r_{MP} (Å)	N_{10}
0.2	3×10^{-4}	6.8×10^{-5}	405	492	0.6
0.5	3×10^{-4}	6.8×10^{-5}	405	363	1.4
1	3×10^{-4}	6.8×10^{-5}	405	288	2.8
2	6×10^{-4}	1.4×10^{-4}	398	229	5.3
5	3×10^{-3}	6.8×10^{-4}	382	168	11.7
10	1	2.3×10^{-1}	324	134	14.2
20	3×10^3	6.8×10^2	244	106	12.2

^aExperimental data for Si:P from G. Feher and E. A. Gere [Phys. Rev. 114, 1245 (1959)].

delocalization cannot occur for arbitrarily small J , even without random fields. To see this we set $B_1 = B_2 = 0$ in Eq. (3),

$$\langle S_{1z} \rangle = \frac{1}{2} \cos(2Jt).$$

As J becomes small, the delocalization rate ($2J$) decreases and the up-spin density cannot be transferred to the second site before spin-lattice relaxation occurs. Let J_{10} be the minimum value of J for which at least 10% of the initial spin density on site 1 is transferred to site 2 before $t = T_1$. Then $J_{10} = 0.23/T_1$. For dilute Si:P, $1/T_1 = 3 \times 10^{-4}$ Hz at 4.2 K and $J_{10} = 6.8 \times 10^{-5}$ Hz. Since J is an exponentially decreasing function of the separation of the donor pair, we can convert J_{10} into a maximum donor separation, r_{10} , by using the values of $J(r)$ calculated numerically by Cullis.^{4,5} For large separations his results can be fitted by⁶

$$J(r) = 2.69 \times 10^7 e^{-r/10} \text{ MHz}, \quad (5)$$

where r is in Å. Substituting J_{10} for $J(r)$, we get $r_{10} = 405$ Å. In Table I we compare this number with the most probable donor separation, $r_{MP} = (\frac{3}{4}\pi n)^{1/3}$, where n is the donor concentration and a Poisson distribution is assumed. We also give $N_{10} = 3/(4\pi n r_{10}^3)$, the average number of donors that will lie inside a sphere of radius r_{10} . We see that for a donor concentration less than 5×10^{16} cm⁻³, $r_{MP} > r_{10}$ and $N_{10} < 1$, which means that for these concentrations most donors are not close enough to a second donor for even 10% delocalization to occur within a time T_1 . That is, T_1 relaxation prevents delocalization for very low donor concentrations even in the absence of random fields. Also note that for larger concentrations the rapid increase in the spin-lattice-relaxation rate limits N_{10} . Although when $N_{10} > 1$ interactions between triples, etc. must be considered, these results suggest that T_1 limits the size of the cluster on which spin diffusion can occur.

The situation is very different when random fields are added. If we set $B_1 - B_2 = \Delta H_{rms} = 3.92$ MHz in Eq. (3), we find $J_{10} = 1.3$ MHz. This gives a spin-delocalization rate that is so rapid compared to the spin-lattice-relaxation rates given in Table I that T_1 is simply not a consideration. The corresponding separation r_{10} is 168 Å. The spin rapidly reaches 10% delocalization in a time short compared to T_1 , but then merely oscillates about this time-averaged value; for a given separation the degree of delocalization is determined solely by the difference in

local fields. For $B_1 - B_2 = 117.6$ MHz, $J_{10} = 3.92$ MHz and $r_{10} = 134$ Å. In this case N_{10} is only 2.0 for $n = 2 \times 10^{17}$ cm⁻³. The local field greatly shrinks the size of the "cluster."

III. ELECTRON PARAMAGNETIC RESONANCE

There have been several calculations of the EPR signal of exchange-coupled pairs.^{4,7-11} These calculations have not included the ²⁹Si hyperfine broadening, except to assume a constant Gaussian or Lorentzian broadening of the signal. We can easily use the results of the preceding section to properly include the effects of delocalization on the EPR signal. The transverse microwave magnetic field H_1 will act on the eigenstates ϕ_n as a perturbation,

$$h^{-1} \mathcal{H}_1 = \gamma H_1 (S_{1x} + S_{2x}) \sin(\omega t). \quad (6)$$

Transitions will be induced between eigenstates for which the matrix element $\langle \phi_n | \mathcal{H}_1 \phi_m \rangle \neq 0$ with an intensity proportional to the square of the matrix element. The allowed transitions and relative intensities are given in the Appendix.

In Figs. 2 and 3 we plot $\Delta\nu$, the frequency deviation from $g=2$ of the transitions, in units of A as a function of J/A for specific silicon hyperfine fields, $B_1 = A/40$ and $B_2 = -A/80$. Figure 2 is for a coupled pair with un-

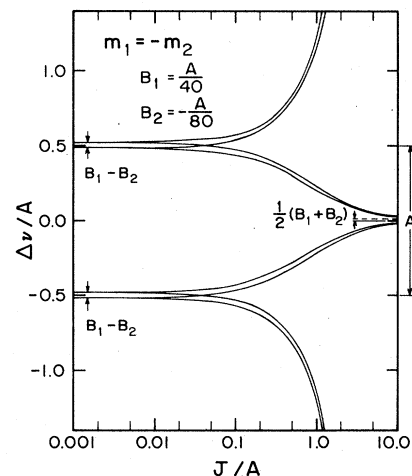


FIG. 2. Dependence on J/A of the position of the EPR absorption of an exchange-coupled phosphorus donor pair subject to random fields B_1 and B_2 for antiparallel P nuclear spins.

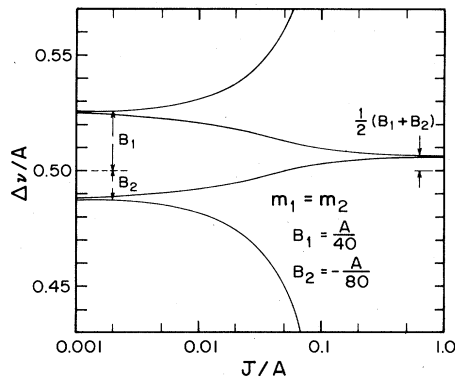


FIG. 3. Dependence on J/A of the position of the EPR absorption of an exchange-coupled phosphorus donor pair subject to random fields B_1 and B_2 for parallel P nuclear spins. The EPR pattern is symmetric about $\Delta\nu=0$.

like orientation of the P nuclear spins ($m_1 \neq m_2$). Each line of the initial ($J=0$) four-line pattern is exchange-split into two lines. One line of each pair moves down to form a single line near $g=2$. In the limit of very large J/A this line is displaced from $g=2$ by $(B_1+B_2)/2$, which is just the average hyperfine field seen by an electron totally delocalized over the two sites. The intensity of this line equals the total intensity of the four unsplit lines. The remaining four exchange-split lines rapidly move away from $g=2$ and decrease to zero intensity. However, the decrease is slow enough that, although the second moment of the eight-line pattern is independent of J , the fourth moment diverges as J^2 .

Figure 3 gives the corresponding plot for one of the two $m_1=m_2$ lines on an enlarged scale. The overall pattern is the same as in Fig. 2, with the merged central line being

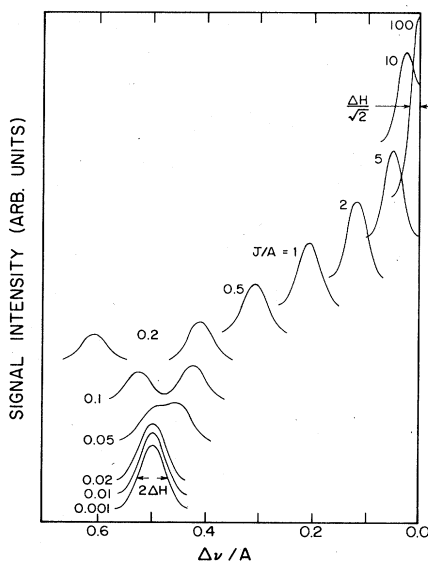


FIG. 4. Dependence on J/A of the EPR absorption line shape of exchange-coupled phosphorus donor pairs subject to a Gaussian distribution of random local fields for antiparallel P nuclear spins. The pattern is symmetric about $\Delta\nu=0$.

displaced from $A/2$ by $(B_1+B_2)/2$. The merger, however, occurs for J/A an order of magnitude smaller in Fig. 3 than in Fig. 2, as expected since, as we saw above, delocalization occurs sooner for $m_1=m_2$ than for $m_1 \neq m_2$.

In an actual Si:P sample the electrons at the various P sites will experience different Si hyperfine fields. Feher³ has shown that this distribution of fields is closely Gaussian. In Fig. 4 we plot the calculated line shape for a Gaussian distribution assuming all electrons experience the same J . (A real sample would have a distribution in J values as well as the superimposed spectra of triples and larger exchange-coupled clusters.) Only half the spectrum is shown since it is symmetric about $\Delta\nu=0$. We can see an initial exchange broadening of the line for small J due to the exchange splitting shown in Fig. 2. This is followed by exchange narrowing and an increase in intensity for large J . The large increase in intensity for $J/A > 10$ is due to the merger of the symmetric line for $\Delta\nu < 0$ as well as to the exchange narrowing. Note that the final width is not $\frac{1}{2}\Delta H$ as we might expect from Fig. 2, but rather $(\sqrt{2}\Delta H)/2$ since we are adding fields which are randomly positive or negative so that the average phase difference is 90° . The line shape for large J is intermediate between a Gaussian and a Lorentzian, with the Gaussian being a better approximation. These effects should be observable for very dilute samples.

We note that there will be a similar narrowing of the $m_1=m_2$ EPR lines (see Fig. 3). This would be extremely hard to see experimentally since there is also a broadening in the range $0.01A < J < 0.1A$, where it might be visible due to the initial splitting of the overlapping $m_1=-m_2$ lines (see Fig. 5). For large values of J the effect of larger clusters predominates and would be hard to separate out. The outside line in Fig. 4 diverges monotonically as J/A

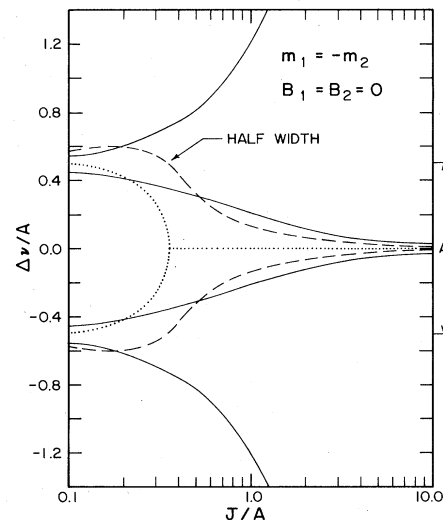


FIG. 5. Comparison of the J/A dependence of the position of the EPR absorption lines of an exchange-coupled pair of phosphorus donor spins with antiparallel nuclear moments for the exact solutions (solid lines) and the exchange-narrowing theory (dotted lines) from Eq. (7). The exact solution gives unbroadened lines, while Eq. (7) predicts first a broadening and then a narrowing (dashed lines).

increases and has not been observed since the distribution in J in an actual sample smears the line out to produce a smooth tail.

IV. COMPARISON WITH THE THEORY OF EXCHANGE NARROWING

Anderson and Weiss¹² and Kubo and Tomita¹³ have developed theories of the line shape of exchange-narrowed EPR lines which make equivalent assumptions and arrive at identical predictions. The crucial assumption in the case of the Anderson-Weiss theory is that the electron absorbs energy at a frequency determined by the local dipolar or hyperfine fields it experiences, but that this frequency is randomly modulated as it varies over the distribution of local fields at a rate determined by the exchange interactions. This assumption is also appropriate for electrons hopping from site to site, so there is an equivalence of exchange and motional narrowing.

Anderson¹⁴ has applied his theory to the simple case of two exchange-coupled hyperfine lines, which we have solved exactly in Sec. IA. It is interesting to make a comparison in order to understand the limitations on the exchange-narrowing theories. Anderson obtains a result, earlier obtained by Archer,¹⁵ in which the absorption intensity is proportional to the function:

$$I(\Delta\nu) = \left\{ 16\Delta\nu^4 + 8\Delta\nu^2 A^2 \left[8 \left(\frac{J}{A} \right)^2 - 1 \right] + A^4 \right\}^{-1}. \quad (7)$$

This function gives two sharp lines at $\Delta\nu = \pm A/2$ for small J which broaden and move together until they merge at $J=0.35A$ to give a single broad line at $\Delta\nu=0$. This line exchange-narrows without shifting as J increases. Anderson has given a plot of Eq. (7) (Fig. 2 of Ref. 14). In Fig. 5 we compare the exact results for $\Delta\nu$ (from the Appendix with $B_1=B_2=0$) with the position of the peaks and half width calculated from Eq. (7). We see that, although there is some qualitative similarity, the agreement is poor except for large J/A . In particular, the two lines which diverge (while losing intensity) are not predicted. In addition, the exact theory predicts exchange *splitting* while the exchange-narrowing theory predicts a *broadening* (neither theory includes an intrinsic linewidth). The continuous distribution of absorption frequencies in Eq. (7) (in contrast to the four discrete absorption frequencies of the exact theory) arises from the assumption of random frequency modulation, which introduces a wide Fourier spectrum of frequencies. The exact theory shows that the only frequency involved is the difference in eigenfrequencies of the two exchange-coupled states, Eqs. (3) and (4). Thus Anderson's application of the exchange-narrowing theory to two hyperfine lines is inappropriate. We expect the exchange theory to be useful only when a large number of eigenstates are exchange-coupled, so that one can replace the distribution of eigenfrequencies with the Fourier spectrum of frequencies introduced by assuming random frequency modulation. The theory should be useful in Si:P for larger values of n ($> 10^{18}$ P atoms/cm³), where large clusters of P impurities are exchange-coupled.

V. CONCLUSION

Our calculation has shown that a spin localized on a donor subject to random fields will partially spread out to a nearby exchange-coupled donor in a time short compared to T_1 . The amplitude of the transferred spin density depends on the difference in random fields at the two donors and on their separation. It drops rapidly as either variable becomes large.

In order to treat spin diffusion in a real sample it is necessary to consider clusters of exchange-coupled donors. As we have seen, there is no unique criterion for when delocalization has occurred. So likewise there is no unique definition of a cluster; whether two donors are in the same cluster will depend on the phenomenon one is considering. Our calculations do suggest that for donor concentrations below 2×10^{17} cm⁻³, the clusters—whatever the definition—will not be very large since the hyperfine random fields are rather large and the exchange interaction drops exponentially with distance. However, it is not likely to be fruitful to attempt to exactly solve for cluster wave functions since even a triad of donors is algebraically formidable. A statistical approach seems much more promising.

ACKNOWLEDGMENTS

I am grateful for useful discussions with Dr. J. Dos Santos and Professor E. Abrahams.

APPENDIX

The exact eigenstates of the Hamiltonian in Eq. (1) are listed below in terms of product wave functions $|M_1 M_2 m_1 m_2\rangle$ (these states do not all have a well-defined parity for $B_1=B_2$):

$$\begin{aligned} \phi_1 &= |++++\rangle, \\ \phi_2 &= \frac{1}{\sqrt{2}}(|+++-\rangle + |++-+\rangle), \\ \phi_3 &= \frac{1}{\sqrt{2}}(|+++-\rangle - |++-+\rangle), \\ \phi_4 &= |++--\rangle, \\ \phi_5 &= K_+(|+-++\rangle + \Omega_+| -+++\rangle), \\ \phi_6 &= K_-(|+-++\rangle + \Omega_-| -+++\rangle), \\ \phi_7 &= L_{2+}(|+---\rangle + \Gamma_{2+}| -+++\rangle), \\ \phi_8 &= L_{1+}(|-++-\rangle + \Gamma_{1+}| +--+-\rangle), \\ \phi_9 &= K_+(|+----\rangle + \Omega_+| -+----\rangle), \\ \phi_{10} &= K_-(|+----\rangle + \Omega_-| -+----\rangle), \\ \phi_{11} &= L_{2-}(|+---\rangle + \Gamma_{2-}| -+++\rangle), \\ \phi_{12} &= L_{1-}(|-++-\rangle + \Gamma_{1-}| +--+-\rangle), \\ \phi_{13} &= |--++\rangle, \\ \phi_{14} &= \frac{1}{\sqrt{2}}(|----\rangle + |----+\rangle), \\ \phi_{15} &= \frac{1}{\sqrt{2}}(|----\rangle - |----+\rangle), \\ \phi_{16} &= |-----\rangle, \end{aligned}$$

where

$$\begin{aligned} K_{\pm} &= (1 + \Omega_{\pm}^2)^{1/2}, \\ \Omega_{\pm} &= y \pm (1 + y^2)^{1/2}, \\ y &= (B_2 - B_1)/J, \\ L_{n\pm} &= (1 + \Gamma_{n\pm}^2)^{1/2}, \\ \Gamma_{n\pm} &= z_n \pm (1 + z_n^2)^{1/2}, \end{aligned}$$

$$z_1 = (A + B_1 - B_2)/J,$$

$$z_2 = (A - B_1 + B_2)/J.$$

There are sixteen allowed transitions for which the matrix elements of Eq. (6) are nonzero. The transition probability is proportional to the square of the matrix elements of Eq. (6). We neglect the Boltzmann factors giving the population differences between levels.

Transition	Relative transition probability	Transition frequency
$T_1: (16) \longleftrightarrow (9)$	$K_+^2 (1 + \Omega_+)^2$	$\gamma H - \frac{J}{2} - \frac{1}{2}(A - B_1 - B_2) + \frac{1}{2}[J^2 + (B_1 - B_2)^2]^{1/2}$
$T_2: (16) \longleftrightarrow (10)$	$K_-^2 (1 + \Omega_-)^2$	$\gamma H - \frac{J}{2} - \frac{1}{2}(A - B_1 - B_2) - \frac{1}{2}[J^2 + (B_1 - B_2)^2]^{1/2}$
$T_3: (13) \longleftrightarrow (5)$	$K_+^2 (1 + \Omega_+)^2$	$\gamma H - \frac{J}{2} + \frac{1}{2}(A + B_1 + B_2) + \frac{1}{2}[J^2 + (B_1 - B_2)^2]^{1/2}$
$T_4: (13) \longleftrightarrow (6)$	$K_-^2 (1 + \Omega_-)^2$	$\gamma H - \frac{J}{2} + \frac{1}{2}(A + B_1 + B_2) - \frac{1}{2}[J^2 + (B_1 - B_2)^2]^{1/2}$
$T_5: (9) \longleftrightarrow (4)$	$K_+^2 (1 + \Omega_+)^2$	$\gamma H + \frac{J}{2} - \frac{1}{2}(A - B_1 - B_2) - \frac{1}{2}[J^2 + (B_1 - B_2)^2]^{1/2}$
$T_6: (10) \longleftrightarrow (4)$	$K_-^2 (1 + \Omega_-)^2$	$\gamma H + \frac{J}{2} - \frac{1}{2}(A - B_1 - B_2) + \frac{1}{2}[J^2 + (B_1 - B_2)^2]^{1/2}$
$T_7: (5) \longleftrightarrow (1)$	$K_+^2 (1 + \Omega_+)^2$	$\gamma H + \frac{J}{2} + \frac{1}{2}(A + B_1 + B_2) - \frac{1}{2}[J^2 + (B_1 - B_2)^2]^{1/2}$
$T_8: (6) \longleftrightarrow (1)$	$K_-^2 (1 + \Omega_-)^2$	$\gamma H + \frac{J}{2} + \frac{1}{2}(A + B_1 + B_2) + \frac{1}{2}[J^2 + (B_1 - B_2)^2]^{1/2}$
$T_9: (15) \longleftrightarrow (8)$	$\frac{1}{2}L_{1+}^2 (1 + \Gamma_{1+})^2$	$\gamma H - \frac{J}{2} + \frac{1}{2}(B_1 + B_2) + \frac{1}{2}[J^2 + (A + B_1 - B_2)^2]^{1/2}$
$(14) \longleftrightarrow (8)$		
$T_{10}: (14) \longleftrightarrow (7)$	$\frac{1}{2}L_{2+}^2 (1 + \Gamma_{2+})^2$	$\gamma H - \frac{J}{2} + \frac{1}{2}(B_1 + B_2) + \frac{1}{2}[J^2 + (A - B_1 + B_2)^2]^{1/2}$
$(15) \longleftrightarrow (7)$		
$T_{11}: (8) \longleftrightarrow (3)$	$\frac{1}{2}L_{1+}^2 (1 + \Gamma_{1+})^2$	$\gamma H + \frac{J}{2} + \frac{1}{2}(B_1 + B_2) - \frac{1}{2}[J^2 + (A + B_1 - B_2)^2]^{1/2}$
$(8) \longleftrightarrow (2)$		
$T_{12}: (7) \longleftrightarrow (2)$	$\frac{1}{2}L_{2+}^2 (1 + \Gamma_{2+})^2$	$\gamma H + \frac{J}{2} + \frac{1}{2}(B_1 + B_2) - \frac{1}{2}[J^2 + (A - B_1 + B_2)^2]^{1/2}$
$(7) \longleftrightarrow (3)$		
$T_{13}: (15) \longleftrightarrow (11)$	$\frac{1}{2}L_{2-}^2 (1 + \Gamma_{2-})^2$	$\gamma H - \frac{J}{2} + \frac{1}{2}(B_1 + B_2) - \frac{1}{2}[J^2 + (A + B_1 - B_2)^2]^{1/2}$
$(14) \longleftrightarrow (11)$		
$T_{14}: (14) \longleftrightarrow (12)$	$\frac{1}{2}L_{1-}^2 (1 + \Gamma_{1-})^2$	$\gamma H - \frac{J}{2} + \frac{1}{2}(B_1 + B_2) - \frac{1}{2}[J^2 + (A - B_1 + B_2)^2]^{1/2}$
$(15) \longleftrightarrow (12)$		
$T_{15}: (12) \longleftrightarrow (2)$	$\frac{1}{2}L_{1-}^2 (1 + \Gamma_{1-})^2$	$\gamma H + \frac{J}{2} + \frac{1}{2}(B_1 + B_2) + \frac{1}{2}[J^2 + (A - B_1 + B_2)^2]^{1/2}$
$(12) \longleftrightarrow (3)$		
$T_{16}: (11) \longleftrightarrow (3)$	$\frac{1}{2}L_{2-}^2 (1 + \Gamma_{2-})^2$	$\gamma H + \frac{J}{2} + \frac{1}{2}(B_1 + B_2) + \frac{1}{2}[J^2 + (A + B_1 - B_2)^2]^{1/2}$
$(11) \longleftrightarrow (2)$		

¹W. Kohn, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1957), Vol. 5, p. 257.

²P. W. Anderson, *Phys. Rev.* **109**, 1492 (1958).

³G. Feher, *Phys. Rev.* **114**, 1219 (1959).

⁴P. R. Cullis and J. R. Marko, *Phys. Rev. B* **1**, 632 (1970).

⁵J. R. Marko and J. D. Quirt, *Phys. Status Solidi B* **64**, 325 (1974).

⁶ $J(r)$ does not decrease like e^{-2r/a_B} since there is an additional r dependence arising from the interference of the phase factors for the six conduction-band valleys in Si. The degree of interference changes as r increases, giving a more rapid exponential decrease. The constant in front of the exponential is $\sim \frac{1}{10}$ the single-valley value due to this cancellation.

⁷C. P. Slichter, *Phys. Rev.* **99**, 479 (1955).

- ⁸D. Jerome and J. M. Winter, *Phys. Rev.* **134A**, 1001 (1964).
⁹T. Shimizu, *J. Phys. Soc. Jpn.* **25**, 1021 (1968).
¹⁰M. Rosso, *J. Phys. Soc. Jpn.* **38**, 780 (1968).
¹¹D. New and T. G. Castner, *Phys. Rev. B* **29**, 2077 (1984).
¹²P. W. Anderson and P. R. Weiss, *Rev. Mod. Phys.* **25**, 269 (1953).
¹³R. Kubo and K. Tomita, *J. Phys. Soc. Jpn.* **9**, 888 (1954).
¹⁴P. W. Anderson, *J. Phys. Soc. Jpn.* **9**, 316 (1954).
¹⁵J. E. Archer, Ph.D. thesis, Harvard University, 1953.