Spin resonance of inversion-layer electrons in silicon

W. J. Wallace and R. H. Silsbee
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853-2501
(Received 15 February 1991)

Conduction-electron spin resonance at a silicon–silicon-dioxide inversion layer is reported in the thermally activated conduction regime. The sample consisted of a microstrip patch resonator and gate with contacts to the inversion layer in a van der Pauw arrangement. The major results include the absence of measurable anisotropy in the Landé splitting factor and low-temperature saturation of the susceptibility of activated carriers. Since Hall-effect measurements suggest that the sample is inhomogeneous, the results must be viewed as preliminary, however.

I. INTRODUCTION

Spin-dependent effects in silicon–silicon-dioxide inversion layers and other two-dimensional systems have been studied in the past by conductivity measurements,[1–3] Shubnikov–de Haas oscillation, cyclotron resonance,[4] and optical methods.[5] *Localized* interface defect states have been studied by direct detection of absorbed microwave power during electron-spin resonance (ESR), but only in relatively large-area samples (18 cm^2).[6]

This work reports direct detection of conduction-electron-spin resonance (CESR) in silicon–silicon-dioxide inversion layers in samples of much smaller areas (0.05–0.5 cm^2) than previously possible. Use of microstrip resonators gives higher sensitivity than conventional ESR techniques while using smaller samples and allowing a natural way to "gate" the device. Inclusion of ion-implanted van der Pauw contacts provides for conductivity, Hall-effect, and Shubnikov–de Haas measurements at the same time as ESR measurements. Observation of the CESR reported here is preliminary, and no definitive model can be presented.

II. EXPERIMENTAL PROCEDURES

Use of the microstrip resonator and accompanying instrumentation is described elsewhere.[7] The samples consisted of 0.04-cm-thick, [100] float-zone silicon wafers boron doped to 5×10^{14} cm^{-3}. The wafers were oxidized in HCl at about 1000 °C to an oxide thickness of 250 Å. Capacitance-voltage profiling using the "high-frequency–low-frequency" method determined the interface-state density, oxide fixed charge, and flatband voltage (see Table I). A large-area field-effect transistor (FET) with a square gate and resonator was then fabricated with four ion-implanted Ohmic contacts to the interface which could be used for van der Pauw measurements of the conductivity and Hall constant. These contacts were situated at the microwave voltage node on each side of the gate and resonator. Spreading resistance associated with the contacts was about 6 times the surface resistance of the inversion layer. The dimensions of the microstrip resonator and gate were 6.83×6.83 mm^2; the resonant frequency was 9.3 HGz and the loaded Q value was about 1100. An outline of the sample processing is indicated in the Appendix.

TABLE I. Quantities derived from capacitance-voltage profiling.

Oxide charge (charges/cm^2)	Interface states ($cm^{-2}\,eV^{-1}$)	Flatband voltage (V)
1.9×10^{11}	$<1\times10^{11}$	-1.1

Within the cryostat the sample was oriented such that a line passing from corner to corner of the square resonator was vertical. The magnetic field was oriented in the horizontal plane and could be rotated from perpendicular to the sample plane to parallel. A wedge-bonded wire at the center of the aluminum resonator and gate provided the dc contact necessary for gating the interface.

Spectrometer sensitivity was calibrated using a small known frequency modulation of the klystron. The silicon microstrip resonator at the doping density mentioned above is useful only below about 24 K since the conductivity of the silicon rapidly increases at higher temperatures, causing excessive loading of the resonator. As a result of the aluminum dc contacts being laid out directly on the surface of the silicon wafer, the metal-oxide-semiconductor FET (MOSFET) device could not be used at higher temperatures because the conductivity of the substrate interfered with the measurement of inversion-layer conductivity.

Pumping the helium bath gave the lowest temperature obtainable of 2 K. Only slight retuning of the cavity was required with change of temperature or gate voltage over the experimental range. No phase shift in the CESR line was observed with change of temperature or gate voltage. Thermometry was provided by a germanium thermometer mounted on the waveguide containing the sample.

Absorption CESR was detected with derivative recording, being careful not to broaden the line by overmodulation. Signal averaging of from 4 to 32 sweeps was necessary to obtain sufficient signal to noise. Dispersion CESR signals were not detected because of the FM noise of the klystron. CESR signals detected with the magnetic field parallel to the wafer suffer a signal loss of a factor of 2 compared to the signal with the magnetic field perpendicular to the wafer as a result of the microwave resonant mode used; most data were taken with the field perpendicular to the wafer. Proton NMR in water provided field calibration. A wafer heavily ion implanted with phosphorus provided confirmation of the NMR probe calibration. No saturation of the CESR was noted at the power level employed (≈ 0.5 mW) for susceptibility, width, and g-factor measurements. Higher-power measurements ($H_1 \approx 2$ G) were made to characterize the saturation behavior of the CESR.

III. RESULTS

The sample was characterized by conventional electrical measurements in addition to the basic CESR experiment. The voltage measured in zero magnetic field between the Hall probes was about 25% of the longitudinal value. This is much too high a value given the high symmetry of the sample. Fowler, Fang, and Hochberg measured the Hall constant[8] on a van der Pauw sample of area 2×10^{-3} mm^2 and reported that at low inversion-layer charge the samples very often showed large asymmetry. They concluded that such samples were spatially inhomogeneous. Unfortunately, Shubnikov–de Haas measurements were not made on the present sample; if oscillations had been noted, it would have been evident that there exist some regions of high mobility. Assuming the present sample is inhomogeneous, results of conductivity and Hall measurements must be interpreted carefully.

Conductivity as measured in the van der Pauw geometry showed activated behavior at all accessible gate biases. The maximum voltage which could be applied to the gate of the sample was about $+0.55$ V as a result of electrical breakdown of the oxide. This was not sufficient to obtain enough interface charge for metallic conductivity, although the activation energy at $+0.55$ eV is nearly zero. At 4.2 K the conductivity was not yet linear with gate voltage over the experimental range, and the conduction threshold was about $+0.35$ V.

Use of the oxide capacitance and flatband voltage obtained in room-temperature capacitance-voltage profiles, -1.1 V, gives a number density of interface charge of 1.3×10^{12} cm^{-2} at $+0.55$ V gate bias. For the bulk doping density quoted above, the depletion layer charge is 1×10^{11} cm^{-2} and the interface-state density is about 1×10^{11} cm^{-2}eV^{-1}. The free charge at low temperature could then be expected to be 1.1×10^{12} cm^{-2}, assuming that all the interface states are filled and localized, which should be sufficient for metallic conduction. Capacitance-voltage profiling of the interface-state density is insensitive to fast traps near the conduction band, however. Fang and Fowler[9] found a large density of trap states ($\approx 10^{12}$ cm^{-2}) near the conduction band for some oxide growth and anneal conditions. These same workers also observed large positive threshold voltage shifts below 77 K, similar to shifts measured in the present sample, which they explained by assuming that the trap states near the conduction band must be filled. The results of the electrical measurements in the present case indicate a sample with substantial inhomogeneity and with a high density of traps near the conduction-band edge, which, together with the limited breakdown strength of the large-area oxide, prevented unambiguous measurement in the metallic regime.

The inset to Fig. 1 is a representative CESR signal taken at a temperature of 8 K and a gate voltage of 0.55 V. The CESR was characterized over a temperature range of 2–20 K and for gate potentials from 0.0 to $+0.55$ V. No anisotropy of the Landé splitting factor g was observed for different orientations of the magnetic-field direction under any experimental condition to within ±0.0004. All measurements of g were, within experimental precision, equal to the conduction-electron value observed in the heavily phosphous-doped sample $g=1.9988$, although there are small changes with temperature and gate voltage. The error in the absolute determination of g is ±0.0004. Figure 1 shows relative g-factor measurements, uncertain to ±0.0001, under different sample conditions. The g factor is seen to increase with decreasing temperature and increasing gate voltage, although most of the change with gate voltage is between 0.0 and 0.35 V.

This lack of anisotropy to the g factor is a surprising result. At the interface-charge densities encountered in this work ($<2\times10^{12}$ cm^{-2}), all electrons occupy two of the six conduction-band valleys in an inversion layer. The evidence for this is varied and convincing.[4] Since the orbital angular momentum is anisotropic for a single valley, the spin-orbit coupling should produce an anisotropic Landé splitting factor g.[10] Wilson and Feher[11] measured this anisotropy in bulk silicon by applying stress and found that $g_{\parallel}-g_{\perp}=1.04\times10^{-3}$, where $g_{\parallel}$ refers to the long-valley axis. It is expected then that the CESR of the inversion layer would show g anisotropy as the magnetic field is rotated between being parallel and perpen-

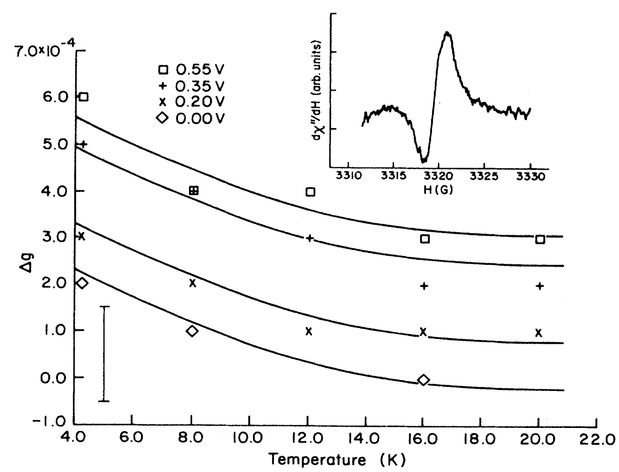

FIG. 1. Landé splitting factor g, relative to the conduction-band value, vs temperature as a function of gate voltage. The zero of this plot is $g=1.9988\pm0.0001$. All data are for a magnetic field perpendicular to the wafer. The error bar shows the typical error in each point relative to the mean. In all graphs lines are guides to the eye. The inset is a CESR signal taken at $T=8$ K, gate bias $=$ 0.55 V, and averaged over 16 sweeps.

dicular to the interface plane.

Figure 2 illustrates the variation of sample susceptibility with gate voltage and temperature. The measured values may be compared with the susceptibility, easily calculated for the two-dimensional gas of noninteracting electrons with a Fermi temperature T_F:

$$\chi=\frac{\mu_B^2 D}{t}(1-e^{-T_F/T}) . \tag{1}$$

The susceptibility is expressed as an average susceptibility for the full depth t of the sample and is dimensionless. The susceptibility of the two-dimensional layer χ_{2D} should be given as $\chi_{2D}=\chi t=0.04\chi$ cm for our sample. In Eq. (1), μ_B is the Bohr magneton and D is the two-dimensional density of states, including both spin states and with a twofold valley degeneracy appropriate to occupation of only the lowest subband of the two-dimensional gas; its value is taken as 1.0×10^{26} $\text{cm}^{-2}\,\text{erg}^{-1}$ (corresponding to 1.6×10^{14} $\text{cm}^{-2}\,\text{eV}^{-1}$).

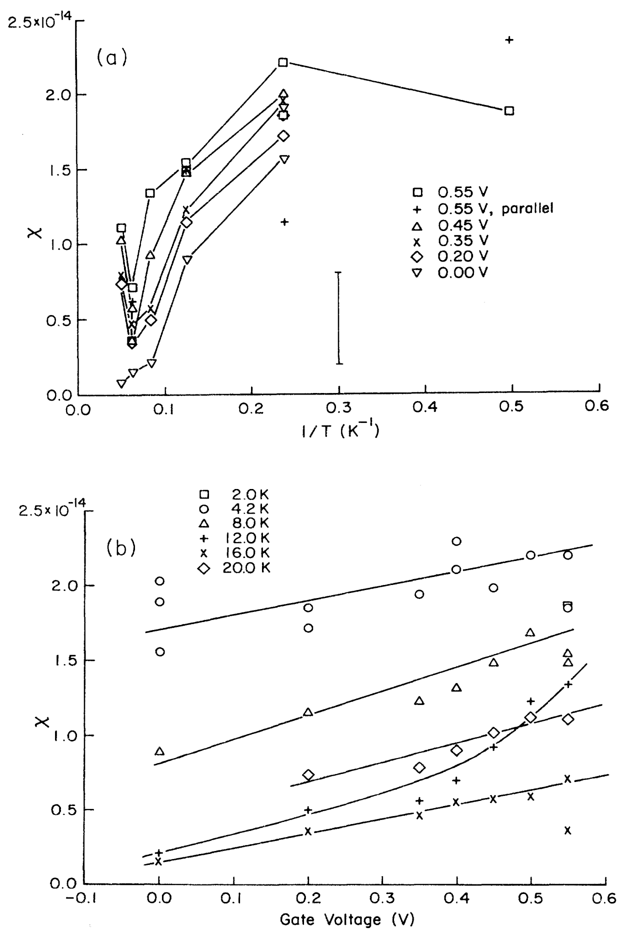

FIG. 2. (a) Susceptibility vs inverse temperature as a function of gate bias. All data are for a magnetic field perpendicular to the wafer unless specified. (b) Susceptibility vs gate voltage as a function of temperature. Error bars are given by the scatter in the 4.2-K data.

At 0 K the Pauli susceptibility calculated from Eq. (1) for the orbital twofold degenerate ground state is 2×10^{-13}. As seen in Fig. 2(a), the measured susceptibility appears to be saturating at about 2×10^{-14}. Note also, in Fig. 2(b), the sharp drop in susceptibility at 16 K and the subsequent rise at 20 K. This behavior is not understood, but it is not believed to be a systematic error.

Below a gate voltage of +0.35 V, the conductivity of the sample drops rapidly and the CESR signal drifts with time if the gate voltage is changed at low temperature, presumably because the interface can no longer reach electrostatic equilibrium. If the sample is cooled from room temperature with zero gate bias, a CESR signal is observed, but if cooled with a gate bias of -1.0 V, no CESR is observed. The sample should be close to electrostatic equilibrium when cooled from room temperature if the gate bias is not changed. Gate biases between these two values were not explored in cooling from room temperature. There is no difference between the measured susceptibility with the magnetic field perpendicular and parallel to the interface to experimental accuracy.

Figure 3 shows the behavior of the full width at half maximum of the Lorentzian absorption lines as a function of temperature. There appears to be a broad minimum around 16 K. The line is broadening rapidly below 4 K, and there may be a difference in linewidth between the magnetic field perpendicular and parallel to the interface at the lowest temperature. Evident as well is the increase of width with increasing gate voltage.

Spin relaxation rates can be estimated from Elliott's[12] theory to be $1/T_1\approx(\Delta g)^2/\tau$, where T_1 is the longitudinal relaxation time, Δg is the shift from the free-electron value, and τ is the momentum relaxation time. For $\Delta g=10^{-3}$ and an interface mobility of 10^2 $\text{cm}^2\,\text{V}^{-1}\,\text{s}^{-1}$, this predicts a linewidth of 1 G. Plotting the CESR signal amplitude versus microwave amplitude into the saturated regime gave an estimate of the ratio of the longitudinal-to-transverse relaxation times T_1/T_2 of approximately 1. The maximum saturation factor obtainable before the onset of excess sample heating and spec-

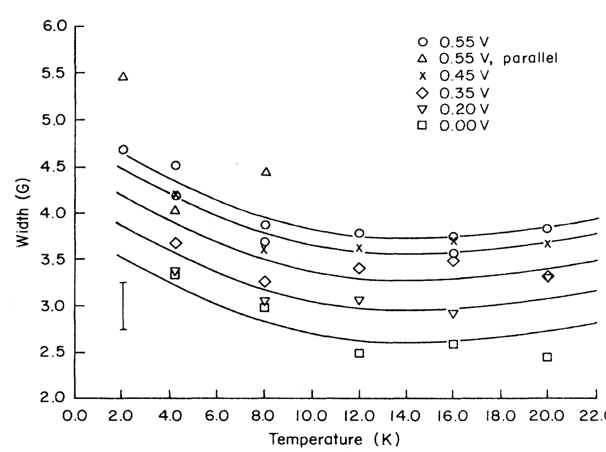

FIG. 3. Width vs temperature as a function of gate voltage with the magnetic field perpendicular to the wafer.

trometer drift was only about 0.8, and so this is only an order-of-magnitude estimate. Although detection of the dispersion signal was difficult because of the klystron frequency noise, when detected it had about the same amplitude as the absorption signal. All of the detected CESR lines appeared to be pure Lorentzian to within experimental accuracy. These results suggest that the line is homogeneously broadened.

IV. DISCUSSION

Several pieces of evidence support the identification of the line detected in the present experiment with the CESR of inversion-layer charge. At low temperature localized defects in silicon and silicon dioxide have very long longitudinal relaxation times and consequently show heavy saturation effects and would only be detectable under present experimental conditions in the dispersion phase of the spectrometer. Thermal radiation from room-temperature sources was excluded from the sample holder since it might shorten relaxation times. Since the detected line is close to homogeneously broadened, its source is probably not localized defects. Finally, most localized states in silicon have g's greater than 2. For instance, the g value of the P_b interface state is 2.0012–2.0081 (Ref. 13) and is outside of the experimental error for the present measured g. Since the g factor is close to the value expected for conduction electrons in silicon, the g factor also suggests that the ESR is from itinerant states.

The implanted arsenic n^+-type contacts to the inversion layer used for conductivity measurements might show a homogeneously broadened line, but the width would be much larger (>10 G) than the measured width.[14] Also, samples made without contacts, but cooled with the gate shorted to the substrate, show the same signal as obtained from samples fabricated with contacts at 0 V gate bias. Calculation also implies that the signal from the implanted contacts would be too small to detect.

ESR from sources at the interface of the aluminum back plane with the silicon is excluded by the absence of signal in samples fabricated without an oxide layer.

Finally, the dependence of the ESR on the interface electric field implies that the signal originates there. In Figs. 2(b) and 3, it is seen that the ESR signal begins to change rapidly at gate voltages above about $+0.35$ V, the threshold voltage deduced from the conductivity measurements at 4.2 K. As mentioned above, when cooled with a gate bias of -1.0 V, at which bias the charge at the interface should be close to zero, no signal is detected.

No explanation of the isotropic g factor in this system can be presented at this point beyond the suggestion that localization produces strong mixing of all six conduction-band minima. Lack of anisotropy in the g factor would seem to be strong evidence that the ESR signal detected is not from conduction electrons in an inversion layer, but it is difficult to explain the data any other way.

The small magnitude of the saturation susceptibility is probably best explained by recourse to the spatial inhomogeneity of the sample; the detected CESR signal comes from only a fraction of the sample interface area. If the CESR signal originates in regions of activated conductivity, then any model must explain simultaneously a saturating susceptibility and activated conductivity. In regions of metallic conductivity, the susceptibility would saturate at the Pauli value, but in regions of activated conductivity at zero temperature there are no carriers and the saturation must have another explanation. It is not known whether regions of metallic conductivity exist in the present sample, but if they did, it would be even more surprising if CESR of these regions showed no g anisotropy. Recourse to localization would be less plausible in these regions. Because the conductivity is activated and the g is isotropic, the tentative conclusion is that the CESR originates in regions of activated conductivity and not pockets of metallic conductivity.

Some insight into the possibility of activated conduction with a saturating susceptibility can be had by analogy to the bulk-silicon case just below the metal-insulator transition. In bulk silicon in the vicinity of the metal-insulator transition, no saturation of the susceptibility is seen at low temperature.[15] In the model of Bhatt and Lee[16] this is a result of a very broad distribution of nearest-neighbor antiferromagnetic exchange interactions arising from the random distribution of donors.

If, on the other hand, there is an upper limit to the strength of the exchange, then for a temperature above this limit the system will exhibit Curie-Weiss susceptibility. In this case the susceptibility saturates as the Néel temperature is approached, while the conductivity remains activated.

Width-versus-temperature data may also be compared with the results of spin resonance in n-type silicon doped below the metal-insulator transition. In the bulk case an increase in relaxation rate is seen at low temperature[15] qualitatively similar to what is seen in the present work. Important differences between the bulk and inversion-layer cases include the small density of (acceptor) impurities and a much smaller Fermi temperature at the metal-insulator transition in the inversion layer.

V. CONCLUSION

The conduction-electron-spin resonance of inversion-layer charge has been detected, but no g anisotropy was observed. The suggestion was made that localization produces strong mixing of all six conduction valleys, resulting in an isotropic g factor. Because of spatial inhomogeneity in the sample, it was not possible to determine unambiguously interface-charge density and mobility. Also, the breakdown field of the oxide layer was too small to allow biasing the inversion layer into the metallic regime. Both of these difficulties are an indirect result of the relatively large area of the gated region. Tentative conclusions are that the detected spin resonance is from thermally activated carriers and that the susceptibility of these carriers is saturating at low temperature.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research, Contract No. NOO014-85-K-0404, and by the National Science Foundation, Grant No. DMR88-18558. Many thanks are due the staff of the National Nanofabrication Facility for help with sample preparation.

APPENDIX: VAN DER PAUW SAMPLE FABRICATION

I. Oxidation of wafer: (a) Grow 250 Å of oxide in HCl tube on two wafers at approximately 1000 °C; (b) check oxide thickness with ellipsometer; and (c) obtain *C-V* profile of test wafer to verify low fixed oxide charge ($\approx 10^{11}$ cm^{-2}), low interface-state density ($< 10^{11}$ cm^{-2} eV^{-1}), and flatband voltage (≈ -1.1 V).

II. Implant-mask fabrication: (a) Thermally deposit ≈ 1 μm of aluminum onto sample wafer and (b) pattern holes in aluminum mask for implanted contacts.

III. Implantation of arsenic: (a) Implant arsenic at 90 keV at a dose of 1×10^{15} cm^{-2} and (b) strip off aluminum implant mask.

IV. Implant anneal: (a) Clean wafer for activation anneal using ammonium hydroxide, hydrogen peroxide, and weak hydrogen fluoride. (Note that oxide is only 250 Å thick so that care must be taken not to overetch with the hydrogen fluoride.) (b) Anneal at 950 °C for 60 min in nitrogen ambient. (Peak density is then 1×10^{20} cm^{-3} with a depth of about 1500 Å. This temperature and time give 100% activation and a surface resistance of ≈ 67 $\Omega/\square$. Ellipsometry after cleaning and annealing shows a loss of about 20 Å of oxide as a result of cleaning.)

V. Patterning of the resonator: (a) Thermally evaporate approximately 2 μm of aluminum on top of wafer (deposition rate was about 100 Å s^{-1}) and (b) etch resonator metallization in a phosphoric acid–acetic-acid–nitric-acid solution at 50 °C (pattern includes *C-V* profiling dots).

VI. Removal of oxide around resonator: (a) Dab photoresist onto *C-V* dots in order to retain oxide around them and bake and etch oxide off in 6:1 buffered HF using aluminum and applied photoresist as mask (thickness of aluminum metallization after etching of oxide around resonator is about 1.3 μm).

VII. Patterning of the source contacts: (a) Thermally evaporate 2000 Å of aluminum into sample (at a RRR value of 40, this gives a surface resistivity of ≈ 1 $\Omega/\square$ for the dc aluminum traces) and (b) pattern and etch contact traces.

VIII. Ground-plane metallization: (a) Thermally evaporate ≈ 2 μm of aluminum onto back, again at about 100 Å s^{-1}, and (b) anneal sample at 400 °C for 10 min in rapid thermal annealer in a 10% hydrogen ambient.

IX. Dicing the wafer.

X. Checking the *C-V* on dots on resonator wafer to verify that processing did not affect the interface.

[1] R. L. Vranch, B. Henderson, and M. Pepper, Appl. Phys. Lett. **52**, 1161 (1988).

[2] A. Berg, M. Dobers, R. R. Gerhardts, and K. v. Klitzing, Phys. Rev. Lett. **64**, 2563 (1990).

[3] Michael Dobers, Surf. Sci. **229**, 126 (1990).

[4] T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).

[5] H. W. van Kesteren *et al.*, Phys. Rev. Lett. **61**, 129 (1988).

[6] K. L. Brower, Appl. Phys. Lett. **43**, 1111 (1983).

[7] W. J. Wallace, Ph.D. thesis, Cornell University, 1990; W. J. Wallace and R. H. Silsbee, Rev. Sci. Instrum. **62**, 1754 (1991).

[8] A. B. Fowler, F. F. Fang, and F. Hochberg, IBM J. Res. Dev. **8**, 427 (1964).

[9] F. F. Fang and A. B. Fowler, Phys. Rev. **169**, 619 (1968).

[10] Y. Yafet, in *Solid State Physics*, edited by Frederich Seitz and David Turnbull (Academic, New York, 1963), Vol. 14.

[11] D. K. Wilson and G. Feher, Phys. Rev. **124**, 1068 (1961).

[12] R. J. Elliott, Phys. Rev. **96**, 116 (1954).

[13] Philip J. Caplan, Edward H. Poindexter, Bruce E. Deal, and Reda R. Razouk, J. Appl. Phys. **50**, 5847 (1979).

[14] Hiroshi Kodera, J. Phys. Soc. Jpn. **19**, 915 (1964); Hiromoto Ue and Shigeru Maekawa, Phys. Rev. B **3**, 4232 (1971).

[15] C. T. Murayama, W. G. Clark, and J. Sanny, Phys. Rev. B **29**, 6063 (1984).

[16] R. N. Bhatt, Physica B **146**, 99 (1987).