Conduction by Small Polarons in Large Electric Fields

P. R. Emtage
Westinghouse Research Laboratories, Pittsburgh, Pennsylvania 15235
(Received 11 September 1970)

The bandwidth of small polarons at low temperatures is so narrow [$\sim 10^{-3}$ (optical phonon energy)] that an electric field of the order of 10^3 V/cm produces a potential difference between neighboring atoms that is comparable with the bandwidth. It is shown that the electron states are then localized and that charge transport proceeds by hopping from site to site, the conductivity being much less than the low-field conductivity by nonlocal electrons. The field-induced transition from band to hopping modes of transport is marked by a region of constant current or negative resistance; at higher fields conduction remains Ohmic until fields of the order of 10^6V/cm are reached.

INTRODUCTION

This account makes extensive use of the description of small polarons that was given by Holstein[1] in 1959 (hereafter referred to as I). A sketch of the description follows: Holstein considered a row of atoms to which the ground electron states are tightly bound, there being a strong interaction between the electron energy and the atomic positions; i.e., if the nth atom is displaced a distance x_n from equilibrium, then the energy of the ground state on it is $E_0 - Ax_n$, where A is a field. When one takes account of the interaction energy J between electron states on neighboring sites, then normally a band of width $2J$ is formed. However, when A is large, the band shifts in energy and becomes extremely compressed, the width being $2Je^{-S}$, in which S is of the order of 5–10 at low temperatures. The approximations hold when J is small ($\lesssim \hbar\omega$, ω being an optical mode frequency) and S is large; detailed limits are given in I. It was found that at low temperatures, $T \lesssim 0.4\Theta$ ($\Theta = \hbar\omega/k$), a band picture of conduction was valid; at high temperatures, $T > 0.5\Theta$, conduction proceeded by hopping between neighboring sites, the uncertainty in energy due to scattering having become greater than the bandwidth.

Three regions of electric field will be considered. In low fields a band picture is valid, so long as a band would ordinarily exist. In moderate fields an electron moving one lattice spacing gains from the field an energy comparable with the bandwidth. In view of the extreme narrowness of the band, such fields can be of the order of thousands of V/cm. In high fields the electron gains an energy comparable with kT per lattice spacing; such fields are of the order of millions of V/cm.

In the presence of an electric field F the energy of the atomic state on the nth atom is

$$E_n = E_0 - Ax_n - elnF, \tag{1}$$

where l is the interatomic distance.[2] Setting this in I and carrying through the argument, one finds that if C_p is the amplitude at the pth atom of the eigenstate of energy E, then the equation following Eq. (22) of I becomes

$$(E - elFp)\,C_p = -Je^{-S_T}(C_{p+1} + C_{p-1}), \tag{2}$$

in which

$$S_T = \frac{1}{\pi}\int_0^{\pi} \frac{A^2}{2M\omega_k^2\hbar\omega_k}(1-\cos k)(1+2N_k)\,dk,$$

N_k being the occupation number of the optical phonon of wave number k. When $F = 0$, the solutions of (2) give a band of width $4Je^{-S_T}$, possessing the striking feature that the bandwidth decreases with rising temperature. For finite values of F, the solutions of (2) that do not diverge at large distances are approximately

$$C_p = (2/\mathfrak{N})^{1/2}\cos\phi_p, \quad p_0 - \tfrac{1}{2}\mathfrak{N} < p < p_0 + \tfrac{1}{2}\mathfrak{N}$$
$$= 0 \quad \text{elsewhere},$$
$$E = elFp_0, \tag{3}$$

in which

$$\phi_p = \sum_{n=p_0}^{p}\sigma_n, \quad \cos\sigma_p = 2elF(p_0 - p)/W$$

and

$$\mathfrak{N} = W/elF, \quad W = 4Je^{-S_T}.$$

These equations describe a localized wave function centered on p_0 and of length $\mathfrak{N}$ equal to the number of sites in which the electron gains from the field an energy equal to the bandwidth W. Each basis state contains all real wave numbers σ and terminates when the local energy, measured from the center of the band, goes outside the local band edge. A more accurate set involving a WKB-like variation in amplitude and some complex wave numbers can be set down, but these (though they fail in orthogonality) are sufficient for our needs (see the Appendix).

It is clear that the basis states must be well localized for large fields, since motion between the atoms is not possible within the band. Solving Eq. (2) for the pth wave function under the condition $eFl \gg W$, one finds

$$E_p = elFp, \quad C_p \simeq 1, \quad C_{p\pm n} = \frac{1}{n!}\left(\mp \frac{W}{4elF}\right)^n .$$

This state is localized on one atom, so at moderate and high fields conduction proceeds by hopping between neighboring atoms even when the temperature is low.

HOPPING CONDUCTION

The development precisely follows that of I, using Eq. (18) of I for the matrix element V connecting neighboring sites; the only difference is that a definite energy difference $\pm eFl$ appears in the condition of energy conservation. The argument that follows Eq. (54) of I is thereby simplified, for the difficulties that derive from transitions conserving all phonon occupation numbers ("diagonal" transitions) cannot arise because it is no longer possible to conserve electron energy during a hop. Equation (65) for the transition probability then becomes, with the aid of (68) and (77),

$$P_\pm = \frac{J^2}{\hbar^2} e^{-G} \int_{-\infty}^{\infty} \exp[-\lambda(t' - i\hbar/2kT)^2 \mp ieFlt'/\hbar]\, dt'$$

$$= \tfrac{1}{2} P_0 \exp\left(\pm \frac{eFl}{2kT} + \frac{e^2F^2l^2}{4\hbar^2\lambda}\right) , \qquad (4)$$

where P_0 is the total low-field transition probability,

$$P_0 = 2\left(\frac{\pi}{\lambda}\right)^{1/2} \frac{J^2}{\hbar^2} e^{-G} . \qquad (5)$$

Here

$$G = \pi^{-1} \int_0^\pi dk\, \frac{A^2}{M\hbar\omega_k^3} (1-\cos k) \tanh\left(\frac{\hbar\omega_k}{4kT}\right)$$

$$\to E_a/kT \quad \text{at high } T,$$

and λ is another sum over phonon wave numbers that tends to $4E_a kT/\hbar^2$ for high T. A closer investigation of this derivation shows that further terms arise at fields so high that eFl is comparable with $S\hbar\omega$, S being given after Eq. (2) and of the order of 10. The term $e^2F^2l^2/4\hbar^2\lambda$ in Eq. (4) is of the same order of magnitude and will be dropped. The current density j per electron is

$$j = eP_0 l \sinh(eFl/2kT). \qquad (6)$$

This is an intuitively obvious form for the high-field behavior and shows no marked departure from linearity until the field is in excess of 10^6V/cm. At such fields J might be altered, resulting in a Zener-like breakdown. We note that the material should have an unusually high dielectric strength, dissipative processes being so strong that an avalanche initiated by polarons is impossible.

BAND CONDUCTION AND INTERMEDIATE FIELDS

Consider band electrons centered on sites p_1 and p_2 and of length $\mathfrak{N}$. Contributions to the square of the total matrix element connecting continuum states come mainly from the sum of squares of individual matrix elements between neighbors, since phase factors lead to the cancellation of other terms [see Eq. (92) of I]. The basis states overlap on $\mathfrak{N} - |p_1 - p_2|$ atoms ($|p_1 - p_2| < \mathfrak{N}$), so at low fields the transition probability is proportional to $[1 - |p_1 - p_2|/\mathfrak{N}]/\mathfrak{N}$, the last factor coming from the normalization. For the time being, we omit the exponential field dependence of P to obtain

$$P(\Delta p) = \frac{(1 - |\Delta p|/\mathfrak{N})P_0}{\mathfrak{N}}, \quad |\Delta p| < \mathfrak{N}$$

$$= 0, \quad \text{elsewhere} \qquad (7)$$

$$\sum_{\Delta p \neq 0} P(\Delta p) = P_0 ;$$

i.e., the total transition probability is the same as in hopping. The conductivity at low fields is proportional to $\langle v^2\tau \rangle = \langle v^2 P_0^{-1} \rangle$, and at low temperatures is much greater than the hopping conductivity of Eq. (6), l^2P_0.

We now seek an approximate description of the intermediate field region in which the low-temperature conductivity changes from a band to a hopping mode of transport. We note that $\mathfrak{N}$ cannot be less than 1, since an electron always occupies at least one site, and that when $\mathfrak{N} = 1$, $P(\pm 1) = \frac{1}{2} P_0 e^{\pm eFl/2kT}$. Equation (7) can therefore be modified to include hopping by replacing it with

$$\mathfrak{N} = 1 + W/eFl,$$

$$P(\Delta p) = \left(1 - \frac{|\Delta p|}{\mathfrak{N}+1}\right) \frac{P_0}{\mathfrak{N}} e^{\Delta peFl/2kT} .$$

The current per electron is then

$$j = \sum_{\Delta p \neq 0} el\Delta p P(\Delta p), \quad -\mathfrak{N} \leq \Delta p \leq \mathfrak{N}$$

$$= \sum_{\Delta p > 0} \frac{e^2l^2F}{kT} (\Delta p)^2 \left(1 - \frac{|\Delta p|}{\mathfrak{N}+1}\right) \frac{P_0}{\mathfrak{N}} . \qquad (8)$$

This description can only apply when $\mathfrak{N}$ is small compared with a mean free path; we make allowance for the improbability of long hops by substituting

$$P_0/\mathfrak{N} \to Be^{-|\Delta p|/L},$$

in which L is the actual mean free path expressed in units of l; B is so chosen that the total transition probability remains P_0,

$$B \sum_{\Delta p \neq 0}^{\mathfrak{N}} \left(1 - \frac{|\Delta p|}{\mathfrak{N}+1}\right) e^{-|\Delta p|/L} = P_0 .$$

After some rearrangement Eq. (8) becomes

$$j=\frac{e^2l^2P_0F}{2kT}\left[\left(1-\frac{1}{\mathfrak{N}+1}\right)S_0+\left(2-\frac{3}{\mathfrak{N}+1}\right)S_1+\left(1-\frac{3}{\mathfrak{N}+1}\right)S_2-\frac{1}{\mathfrak{N}+1}S_3\right]\Big/\left[\left(1-\frac{1}{\mathfrak{N}+1}\right)S_0-\frac{1}{\mathfrak{N}+1}S_1\right], \qquad (9)$$

in which

$$S_\alpha=\sum_{n=0}^{\mathfrak{N}-1}n^\alpha e^{-n/L}.$$

It is readily checked that the low- and high-field limits for the conductivity are $e^2l^2L^2P_0/kT$ and $e^2l^2P_0/2kT$, respectively; since $l^2L^2P_0=\langle v^2\tau\rangle$, these are both correct and the description given by (9) should give a fair estimate of the behavior of the intermediate field region.[3] The results of computation are shown in Fig. 1 for various values of L. Table I shows the appropriate values of T/Θ for each L, and gives the electric field scale in V/cm and current density in arbitrary units; the parameters were taken to be $J=\hbar\omega_0=300k$, $S_0=7$, $l=2$ Å, and phonon dispersion small.

CORRECTIONS TO PERTURBATION THEORY

Major corrections to Holstein's transport theory have been introduced by Lang and Firsov.[4,5] These authors find that the probability of scattering in the band description P_B is generally greater than the probability of hopping P_H. Holstein's conclusion that $P_B=P_H$ holds only for small values of J, and at any temperature the conductivity can be expressed as a sum of band and Holstein-like hopping terms, $\sigma=\sigma_B+\sigma_H$. The probability P_H in the hopping conductivity is correctly given by the P_0 of Eq. (4), apart from higher-order terms that do not alter its form. The main part of the probability P_B in the band conductivity is usually given by

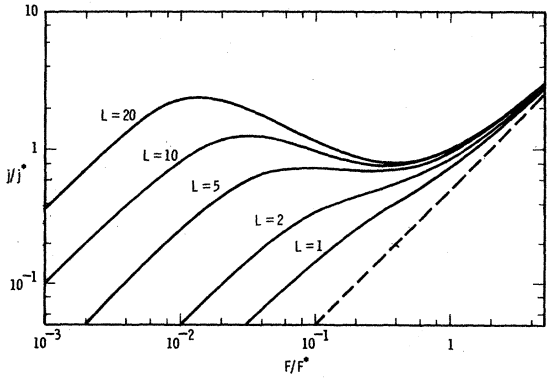

FIG. 1. Variation of current with field at moderate fields and low temperature for various values of the mean free path L; the dotted line is the high-field asymptote. Temperatures and scales are given in Table I.

TABLE I. Temperatures T/Θ, $\Theta=\hbar\omega_0/k$, for the mean free paths L of Fig. 1, and scales of field F^* (V/cm) and current j^* (arbitrary units) found from the parameters in the text.

L	1	2	5	10	20
T/Θ	0.38	0.35	0.31(5)	0.29	0.26(5)
F^* (V/cm)	1600	2000	2500	2900	3300
j^*	1.0	0.77	0.51	0.35	0.25

$$P_B=\frac{(J/S\hbar\omega_0)^4\omega_0^2}{\Delta\omega_0\sinh^2(\hbar\omega_0/2kT)},$$

where $\Delta\omega_0$ is the width of the optical phonon spectrum. This quantity is of higher order in J than is P_0, but it contains a much smaller activation energy and is therefore generally greater than P_0.

The band conduction is derived from the part of the current operator that is diagonal in the phonon occupation numbers:

$$\langle k|I_d|k'\rangle=\delta_{kk'}\,e\hbar^{-1}\frac{\partial E_k}{\partial k}.$$

In moderate electric fields the translational invariance of the system is lost; the basis states are then localized by the field and are real. The matrix elements of I_d for states centered on the sites p and p' are

$$\langle p|I_d|p'\rangle=\tfrac{1}{4}ieWl\hbar^{-1}(\delta_{p,p'+1}-\delta_{p,p'-1});$$

i.e., only nondiagonal terms in p appear even when the number of sites $\mathfrak{N}$ occupied by each basis state is large. Note that in very low fields a momentum representation is appropriate; a Fourier transform of the δ functions in I_d then recovers the form used by Lang and Firsov. In high fields, the fact that all zero-phonon-energy denominators are large makes the effect of I_d small. The high-field hopping current is therefore given adequately by Holstein's account, even at low temperatures.

In sum, the effects described by Lang and Firsov are such as to decrease the conductivity in the band description ($\sigma_B\sim 1/P_B$), but to leave the high-field hopping current unaltered. It follows that the results given in Fig. 1 generally pertain to lower temperatures than those quoted in Table I.

Note added in proof. Those solutions of Eq. (2) that do not diverge at large distances are

$$E=elFp_1,\quad C_{p_1,p}=J_{p_1-p}(\tfrac{1}{2}\mathfrak{N}),$$

in which p_1 is an integer and $J_n(x)$ is the Bessel function of the first kind with argument x and of order n. Use of these wave functions does not alter the results given above.

ACKNOWLEDGMENT

The author is grateful to Dr. R. C. Miller for drawing attention to this subject and for his en-

couragement in the investigation.

APPENDIX

The following remarks give (i) the classical motion of a polaron localized by an electric field, (ii) a better approximation to the wave functions than is given in Eqs. (3), and (iii) a discussion of the errors in the "triangular" weighting function

$$\rho(\Delta p, \mathfrak{N}) \sim 1 - |\Delta p|/\mathfrak{N}$$

that was introduced in Eq. (7) to describe the *a priori* probability that a polaron in a state of length $\mathfrak{N}$ will undertake a hop of length Δp. Units $e = l = \hbar = 1$ are used throughout this Appendix.

When there is no scattering, it is easy to show that a current carrier in an electric field executes a harmonic motion. Consider a polaron with energy equal to that of the center of the band on site p_1. Its energy E relative to the local band center, velocity v, and wave number σ at the pth site are given by [see Eqs. (2) and (3)]

$$\begin{aligned} 2E_p/W &= 2(p_1 - p)/\mathfrak{N} = \cos\sigma_{p_1,p}\ , \\ v_p &= \frac{dp}{dt} = \frac{dE_p}{d\sigma_p} = -\tfrac{1}{2} W \sin\sigma_{p_1,p}\ , \end{aligned} \tag{A1}$$

where $\mathfrak{N} = W/F$. It follows readily that

$$p = p_1 + \tfrac{1}{2}\mathfrak{N}\cos(Ft + \psi)\ , \tag{A2}$$

ψ being a phase angle.

In the harmonic motion described by (A2), the electron spends more time near the band edges than near the center, so the zero-order WKB solutions in Eq. (3) (amplitude equals constant in allowed energy region) are clearly inadequate. It is not difficult to show from Eq. (2) that a better set of solutions is of the WKB-like form,

$$\begin{aligned} C_{p_1,p} &= 2(\mathfrak{N}\pi \sin\sigma_{p_1,p})^{-1/2}\cos\phi_{p_1,p}, \\ &\qquad p_1 - \tfrac{1}{2}\mathfrak{N} \le p \le p_1 + \tfrac{1}{2}\mathfrak{N} \\ &= 0\ , \qquad \text{elsewhere}\ . \end{aligned} \tag{A3}$$

Tails beyond the band edge are here ignored ($\mathfrak{N}$ is large), and the local wave number σ is restricted to the range $0 \le \sigma \le \pi$. The normalization of the functions in (A3) is readily checked provided that one replaces $\cos^2\phi_{p_1,p}$ with $\frac{1}{2}$, a practice that will be followed throughout.

To find the probability of the transition $p_1 \to p_2$ one can (for large $\mathfrak{N}$) neglect the difference between the polaron amplitude on an atom and that on its nearest neighbor. For $p_1 > p_2$ the states centered on p_1 and p_2 both exist within the range $-\frac{1}{2}\mathfrak{N} + p_1 \le p \le \frac{1}{2}\mathfrak{N} + p_2$, so one has

$$\begin{aligned} \rho(\Delta p, \mathfrak{N}) &= \sum_p C^2_{p_1,p}\, C^2_{p_2,p} \\ &\to \frac{1}{(\mathfrak{N}\pi)^2}\int_{-\mathfrak{N}/2+p_1}^{\mathfrak{N}/2+p_2} dp\ (\sin\sigma_{p_1,p}\sin\sigma_{p_2,p})^{-1} \end{aligned}$$

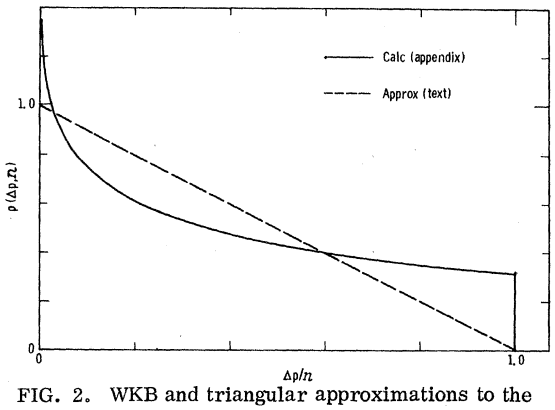

FIG. 2. WKB and triangular approximations to the weighting function $\rho(\Delta p, \mathfrak{N})$ that describes the *a priori* probability of a hop of length Δp between states of length $\mathfrak{N}$.

$$= 2(\pi^2\mathfrak{N})^{-1} F(\pi/2,\ [1 - (\Delta p/\mathfrak{N})^2])\ , \tag{A4}$$

in which $F(\pi/2, k)$ is the complete elliptic integral of the first kind with argument k. Here the summation has been replaced with an integral which was evaluated using Eq. (A1) to express the denominator as a product of four terms of the type $(a - n)^{1/2}$. It can be checked that the total area under ρ, from $-\mathfrak{N} \le \Delta p \le \mathfrak{N}$, is unity.

The function ρ in Eq. (A4) is drawn in Fig. 2 together with the triangular approximation of the text. It is clear that the discontinuities in ρ at $\Delta p = 0$, $\mathfrak{N}$ come from the infinite amplitude of the wave functions in (A3) at the band edge, and that these features can only appear when $\mathfrak{N}$ is very great since the wave functions at the band edge (best described as a WKB turning point) will not otherwise be large. For smaller values of $\mathfrak{N}$ one then expects reasonable accuracy from a triangular approximation to ρ, the greatest errors being incurred when $\mathfrak{N}$ is large; these we now estimate. From Fig. 2 we see that a "square" approximation to ρ ($\rho = \frac{1}{2}$, $0 < |\Delta p| < \mathfrak{N}$; $\rho = 0$, elsewhere) should be about as much in error as the "triangular" approximation, though in the opposite direction. The guess is confirmed by noting that when scattering is small, the current is found from the mean value of $(\Delta p)^2\rho(\Delta p)$, and that these means are 0.17, 0.25, and 0.33 for the triangular, WKB, and square approximations, respectively. A sample computation for the case $L = 10$ was undertaken with the square weighting function, allowance for scattering being made as in the text. It was found that the current was increased by 70% at the current maximum and 50% at the current minimum, fields at these points being increased to a somewhat lesser extent. Errors from the form assumed for ρ are therefore unlikely to be as great as 40%, and do not alter the appearance of the current-voltage curve.

[1]T. Holstein, Ann. Phys. (N.Y.) 8, 343 (1959).

[2]An additional term ϵeFx, ϵ being a local dielectric constant, has been dropped from the Hamiltonian since it has very little effect for fields less than 10^7 V/cm. Hed and Freud [A. Z. Hed and P. J. Freud, J. Non-Crystalline Solids 2, 484 (1970)] suggest that this term can result in switching to normal conduction at high fields because it replaces A with $A_{eff} = A - \epsilon eF$, and a large value of F will then decrease A_{eff} until the criteria for formation of a polaron band are no longer satisfied. Note that ϵ should be close to unity, since the field in a spherical cavity is never more than 50% greater than that in the surrounding dielectric. Hed and Freud take ϵ between 10 and 100, and therefore get large changes at fields of 10^5–10^6 V/cm. Further, these authors failed to note that negative values of F would increase A, so no breakdown occurs in the back direction. Symmetry under field reversal is expected in most materials (including theirs, which were cubic) and can be restored by supposing that A changes sign from site to site. One then finds that A_{eff} is given by

$$A_{eff}^2 = \tfrac{1}{2}(A - \epsilon eF)^2 + \tfrac{1}{2}(-A - \epsilon eF)^2 = A^2 + \epsilon^2 e^2 F^2;$$

i.e., an applied field can generally be expected to augment A and therefore increase the stability of a polaron band.

[3]Lebwohl and Tsu [P. A. Lebwohl and R. Tsu, J. Appl. Phys. 141, 2664 (1970)] have recently used the Boltzmann transport equation to calculate the current density due to carriers with degenerate statistics in a "miniband" arising from a one-dimensional superlattice. They find that the current is inversely proportional to the field at large fields, as is found above in the negative resistance region when L is large. Their current maximum is sharper than mine and occurs at a field that is about 50% greater for a given bandwidth. Both calculations deserve criticism; mine principally because the correlation between scatterings has been included in a quite arbitrary manner; theirs because local momentum eigenfunctions and a relaxation time are not appropriate to the problem of a narrow band in a large electric field. Lebwohl and Tsu's method is, however, likely to be the more accurate of the two, though not readily applied when the discrete nature of electron sites must be allowed for.

[4]I. G. Lang and Yu A. Firsov, Zh. Eksperim. i Teor. Fiz. 43, 1843 (1962) [Soviet Phys. JETP 16, 1301 (1963)].

[5]I. G. Lang and Yu A. Firsov, Fiz. Tverd. Tela 5, 2799 (1963) [Soviet Phys. Solid State 5, 2049 (1964)].

PHYSICAL REVIEW B VOLUME 3, NUMBER 8 15 APRIL 1971

Mobilities of Electrons and Holes in CaF_2

J. Dresner and P. M. Heyman
RCA Laboratories, Princeton, New Jersey 08540
(Received 1 September 1970)

The Hall mobility of electrons and holes in CaF_2 (0.005% Tb) was measured under illumination by near-uv radiation (3100–4000 Å) in the temperature range 25–200 °C. At 334 °K $= \frac{1}{2}\theta_{LO}$, the maximum measured mobilities were 236 and 100 $cm^2V^{-1}sec^{-1}$, respectively, for electrons and holes. These magnitudes are consistent with those predicted by wide-band optical-phonon theories (large polaron). Upper limits of $m^* \leq 0.12m$ and $m^* \leq 0.21m$ were estimated respectively for the electron and hole band masses. Because of two-carrier effects, it was not possible to determine whether the charge transfer during the photochromic process occurs via the conduction or valence bands.

INTRODUCTION

This paper reports the first Hall-mobility measurements for electrons and holes in CaF_2. The samples were doped with 0.005 mole% Tb (and additively colored) to provide some photoconductivity in the region 3100–4000 Å, where a weak absorption band is created.[1]

Some of the properties of CaF_2 relevant to this work are listed: optical band gap[2]-10 eV; dielectric constants at[3] 300 °K—static $\epsilon_0 = 6.63$, optical $\epsilon_\infty = 2.04$; energy of LO phonons[3]—463 $cm^{-1} = 668$ °K $= \theta_{LO}$.

These measurements were motivated by two problems.

(a) Unlike the weakly doped samples measured in this work, CaF_2 containing higher concentrations of rare earths (e.g., 0.1% Tb, La, Ce, and Gd), which has been colored either by ionizing radiation or by chemical reduction exhibits photochromism when irradiated by light in the range 3100–4000 Å (hereafter referred to as switching light). Evidence obtained from optical and electron-paramagnetic-resonance (EPR) data[4,5] shows that the broad absorption band at $\lambda_{max} \approx 4000$ Å originates in a complex center consisting of a RE^{+2} and a nearest-neighbor F center. Upon application of switching light, the complex center loses an electron while a RE (rare earth) located in a state of cubic symmetry, which is trivalent in the stable state, becomes divalent. There is circumstantial evidence[1,4,5] that the charge transfer is effected by electrons through the conduction band. However the alternate process through hole motion cannot be ruled out.[1] The two processes should be distinguishable by the Hall effect.

(b) It is likely that in an ionic crystal such as CaF_2, the electron-lattice interaction should be