Coherent transient effects on stepwise modulation of ^{57}Fe Mössbauer γ radiation

E. H. Realo, M. A. Haas, and J. J. Jogi

Institute of Physics, Academy of Sciences of the Latvian SSR, Tartu
(Submitted 7 December 1984)
Zh. Eksp. Teor. Fiz. **88**, 1864–1872 (May 1985)

The time dependence of resonant γ radiation traversing an absorber of arbitrary thickness has been investigated for stepwise modulation of the source. Modulation of the ^{57}Co(Pd) γ source was achieved by exciting rectangular voltage pulses in a piezoelectric quartz transducer (x cut). Coherent transient effects were found in the 14.4 keV Mössbauer γ radiation from ^{57}Fe manifested in the formation of short (< 80 ns) pulses of the intensity transmitted through the absorber. The maximum intensity of the transient pulses exceeds the magnitude of the stationary resonance effect. Expressions describing the transient pulse shape were derived within the framework of classical theory by taking into account the finite duration of the amplitude distribution of the mechanical displacements in the γ source. The shapes of the pulses and the dependence of the intensity on pulse excitation amplitudes and on isomer shift observed experimentally are in good agreement with the theoretical calculations.

Coherent high-frequency modulation of Mössbauer γ radiation is a very effective method for studying time and frequency relations in the resonant interaction between radiation and solids. A doppler shift of the γ source (resonant absorber) at frequencies exceeding the width of the Mössbauer level Γ_0 leads to significant changes in the temporal and spectral composition of the radiation intensity traversing the resonant absorber. Splitting of a single line into a series of satellites,[1] dispersion of lines in time-resolved spectra[2,3] and quantum beat type oscillations[2,3] are examples of such changes arising from sinusoidal modulation.

There is considerable interest in elucidating interference effects which result from rapid (compared with the Mössbauer level lifetime τ_0) stepwise destruction of the conditions for resonance of the γ source with the absorber. The theory of coherent transient effects in Mössbauer spectroscopy has recently been developed and its main conclusions have been confirmed by experiments on the 93.3 keV γ resonance of the ^{67}Zn isotope ($\tau_0 = 13.2\,\mu$s).[4-6] We have demonstrated[7] the possibility of observing coherent transient effects in the 14.4 keV ^{57}Fe γ radiation. The high Debye-Waller factor ($f \approx 0.7$–0.8) and the short lifetime $\tau_0 = 14$ ns) determine the advantage of the isotope ^{57}Fe for these studies. We note that similar phenomena are known in NMR (Ref. 8) and in laser spectroscopy.[9]

Results of experimental investigations of coherent transient effects on stepwise modulation of ^{57}Fe Mössbauer radiation are given in the present work, and a theoretical analysis of the phenomenon is presented.

STEPWISE PHASE MODULATION

We will consider the following simplified model of the experiment: the γ source and the absorber with single resonance lines of Lorentzian shape are characterized by the corresponding central frequencies ω_s and ω_a, line half-widths Γ_s and Γ_a and Debye-Waller factors f_s and f_a. The value of the stationary resonance effect is given by ε_0. The stepwise phase shift is $a\theta\,(t - T)$, where $\theta\,(t - T)$ is the step function,

$a = \Delta x/\lambda$ and $2\pi\lambda$ is the γ-radiation wavelength which arises on an instantaneous mechanical displacement of the source by a distance Δx at the instant of time $t = T$. We consider that the conditions for resonance $\Delta\omega = \omega_s - \omega_a$ (isomeric shift of the lines $S \sim \Delta\omega$) before and after the stepwise phase shift are maintained ($\Delta\omega = $ const.).

The amplitude of the electric field of the resonant γ-radiation source is expressed in the form[5]

$$E_s(t) \sim E(t-t_0)\exp[ia\theta(t-T)],$$
$$E(t-t_0) = \exp[-i\omega_s(t-t_0)-\Gamma_s(t-t_0)/2]\theta(t-t_0), \tag{1}$$

where t_0 is the formation time of the Mössbauer level.

The field of Eq. (1) can be represented in the form of a sum:

$$E_s(t) \sim E_r(t-t_0)+E_i(t-T), \tag{2}$$

where the components $E_r(t - t_0)$ and $E_i(t - T)$ are the wave trains emitted respectively at an arbitrary instant t_0 and at the time of the phase step $t = T$:

$$E_r(t-t_0) = E(t-t_0)\exp[ia\theta(t_0-T)], \tag{3a}$$

$$E_i(t-T) = (e^{ia}-1)E(t-T)E(T-t_0). \tag{3b}$$

There is then no discontinuity in field intensity $E_s(t)$ at the moment $t = T$ as a result of the interference of the components $E_r(t - t_0)$ and $E_i(t - T)$.

The situation changes when the radiation traverses a resonant absorber. The field $E_s(t)$ is now modified and the amplitude of the absorber response in the frequency representation has the form[10]

$$R_a(\omega) = \exp\left(-\frac{iT_M\Gamma_0/4}{\omega-\omega_a+i\Gamma_a/2}\right), \tag{4}$$

where the Mössbauer thickness of the absorber is $T_M = f_a \sigma_0 n$ (σ_0 is the cross section of the resonant absorber and n is the surface density of resonant nuclei). As a result, the phase relations between the components $E_r(t - t_0)$ and $E_i(t - T)$ are destroyed and a transient pulse of the following form arises in the intensity traversing the absorber:

$$I(t-T, \Delta\omega) = \frac{N(t-T) - N(t<T)}{N_\infty}$$

$$= f_s \int\int_{-\infty}^{\infty} d\omega \, d\omega' R_a(\omega) R_a^*(\omega')$$

$$\times \exp[i(\omega'-\omega)t]\{E(\omega,T)E^*(\omega',T)$$

$$\times [2(1-\cos a)\langle |E(T-t_0)|^2\rangle_{t_0}$$
$$+(e^{ia}-1)\langle \exp[i\omega'(t_0-T)]E(T-t_0)\rangle_{t_0}$$
$$+(e^{-ia}-1)\langle \exp[i\omega(t_0-T)]E^*(T-t_0)\rangle_{t_0}]\},$$

(5)

$$E(\omega,t) = \frac{e^{i\omega T}}{\omega-\omega_s+i\Gamma_s/2},$$

$$E(T-t_0) = \exp\left[-i\omega_s(T-t_0) - \frac{1}{2}\Gamma_s(T-t_0)\right]\theta(T-t_0),$$

where N_∞, $N(t<T)$ and $N(t-T)$ are the counting rates of γ-quanta outside the resonance $\Delta\omega \to \infty$, up to and after the stepwise shift, respectively; $\langle \cdots \rangle_{t_0}$ indicates averaging over arbitrary times t_0 and the asterisks denote the complex conjugate.

Integrating over frequencies in Eq. (5) leads to the following dependence of intensity on time t and the isomeric shift of the lines, $\Delta\omega$:

$$I(t, \Delta\omega) = 2f_s e^{-\Gamma_a t}[(1-\cos a)\,\mathrm{Re}(AB) - \sin a\,\mathrm{Im}(AB)]\theta(t),$$

$$A = A(t, \Delta\omega) = \sum_{n=0}^{\infty} J_n((T_M\Gamma_a t)^{1/2})(-1)^n\left(\frac{4t}{T_M\Gamma_a}\right)^{n/2}$$

$$\times\left[i\Delta\omega + \frac{\Gamma_s-\Gamma_a}{2}\right]^n,$$

(6)

$$B = B(t, \Delta\omega) = \sum_{n=1}^{\infty} J_n((T_M\Gamma_a t)^{1/2})(-1)^{n+1}\left(\frac{T_M\Gamma_a}{4t}\right)^{n/2}$$

$$\times\left[i\Delta\omega + \frac{\Gamma_s+\Gamma_a}{2}\right]^{-n}.$$

J_n is the Bessel function of the first kind of order n, $T = 0$.

For exact resonance, i.e., if the isomeric shifts of the lines are absent ($\Delta\omega = 0$) and the widths are equal, $\Gamma_s = \Gamma_a = \Gamma$, Eq. (6) simplifies:

$$I(t) = \theta(t)2f_s(1-\cos a)e^{-\Gamma t}J_0((T_M\Gamma t)^{1/2})$$

$$\times \sum_{n=1}^{\infty} (-1)^{n+1}J_n((T_M\Gamma t)^{1/2})\left(\frac{T_M}{4\Gamma t}\right)^{n/2},$$

(7)

and for short times $t < (\Gamma/2)^{-1}$ this expression can be transformed to a form more convenient for calculations:

$$I(t) = \theta(t)2f_s(1-\cos a)J_0((T_M\Gamma t)^{1/2})$$

$$\times\left[e^{-\Gamma t}\sum_{n=0}^{\infty}J_n((T_M\Gamma t)^{1/2})\left(\frac{4\Gamma t}{T_M}\right)^{n/2} - e^{-T_M/4}\right].$$

(8)

Analysis of Eqs. (6)–(8) shows that a stepwise phase shift is accompanied by an appreciable jump in the γ-radiation intensity traversing the absorber. In fact, we obtain from Eq. (6), under the conditions $t = 0$ and $\Gamma_s = \Gamma_a = \Gamma$, for the peak intensity

$$I(0, \Delta\omega) = 2f_s\left\{(1-\cos a)\left[1 - \exp\left(-\frac{T_M/4}{1+(\Delta\omega/\Gamma)^2}\right)\right.\right.$$

$$\left.\times\cos\frac{T_M\Delta\omega/4\Gamma}{1+(\Delta\omega/\Gamma)^2}\right]$$

$$\left.+ \sin a \exp\left(-\frac{T_M/4}{1+(\Delta\omega/\Gamma)^2}\right)\sin\frac{T_M\Delta\omega/4\Gamma}{1+(\Delta\omega/\Gamma)^2}\right\}.$$

(9)

For zero isomeric shift $\Delta\omega = 0$, and for a thick absorber, $T_M \to \infty$ and $a = (2n+1)\pi$, it follows from Eq. (9) that $I(0,0) \to 4f_s$. The maximum intensity is thus four times greater than the stationary effect δ_0. Still larger intensities should be observed if the shifts differ from zero, $\Delta\omega > 0$ or $\Delta\omega < 0$, when $I(0,\Delta\omega)$ can in the limit reach $8f_s$. For example, for $a = \pi$, $T_M = 200$, we have $I(0,15,85\Gamma) = 7.28f_s$. The approximately exponential fall which follows the jump in intensity, can be characterized by an attenuation constant τ_n according to the empirical formula obtained by calculating the time dependences of Eqs. (7) and (8):

$$\tau_n \approx \tau_0(\Gamma/\Gamma_0 + T_M/8)^{-k},$$

where $k = 2.5$ for $0 < T_M < 8$ and $k = 2$ for $T_M > 8$. As T_M increases, the peak intensity of the jump consequently increases and the duration of the fall τ_n is strongly reduced ($\tau_n \approx 4$ ns for $T_M = 40$). The exponential fall ends in oscillations with amplitudes which also grow as T_M increases, and their maxima shift in the direction of $t \to 0$. It also emerges from Eq. (9) that for $a = 2n\pi$ the intensity satisfies $I(0,0) = 0$, i.e., a transient pulse is not observed.

The intensities of the transient pulses depend on the magnitude and sign of the phase step and on the isomeric shift of the lines. The relations $I_+(0,\Delta\omega) = I(0,\Delta\omega,a)$ and $I_-(0,\Delta\omega) = I(0,\Delta\omega - a)$ corresponding to the phase (mechanical) shifts of the source in the direction of the absorber, a, and away from the absorber, $-a$, are symmetrical relative to the resonance $\Delta\omega_+ 0$ and pass through a maximum in the regions $\Delta\omega < 0$ and $\Delta\omega > 0$ respectively. The sensitivity of $I_+(0,\Delta\omega)$ to the isomer shift is larger for smaller values of a. The ratio of I_+ to I_- can be used for accurate measurements of the shift of the lines within relatively narrow limits $|\Delta\omega| \leqslant (1-2)\Gamma$ (Ref. 7).

Equations (5) to (9) were obtained on the assumption of an ideal stepwise phase shift, which is not strictly achieved. The analysis of an actual experimental situation therefore

requires the following factors to be taken into account: 1) the phase jump takes place over a finite time Δt according to some law

$$a(t) = \frac{1}{\lambda} \int\limits_{-\infty}^{t} v(\tau) d\tau,$$

where $v(\tau)$ is the velocity of relative motion of source and absorber: 2) a spread of phase shifts exists within the body of the γ source; 3) the recording apparatus has a finite resolving time.

The finite duration of the phase jump indicates, essentially, the occurrence of frequency modulation. This type of modulation leads to substitution in Eq. (1) of the phase factor $\exp[ia\theta(t-T)]$ by

$$\Phi(t) = \exp\left[\frac{i}{\lambda} \int\limits_{-\infty}^{t} v(\tau) d\tau\right].$$

The intensity of the transient pulse is then

$$I(t, \Delta\omega) = 2f_s e^{-\Gamma_a t} \mathrm{Re}\left\{\int\limits_{-\infty}^{t} d\tau \int\limits_{-\infty}^{\tau} d\tau' \Phi^*(\tau) \dot{\Phi}(\tau') A(t-\tau, \Delta\omega)\right.$$

$$\times \exp\left[\left(\frac{\Gamma_a - \Gamma_s}{2} - i\Delta\omega\right)\tau\right] B(t-\tau', \Delta\omega)$$

$$\times \exp\left[\left(\frac{\Gamma_a - \Gamma_s}{2} + i\Delta\omega\right)\tau\right]$$

$$+ \int\limits_{-\infty}^{t} d\tau \dot{\Phi}^*(\tau) \Phi(\tau) e^{\Gamma_a \tau}[\,|A(t-\tau, \Delta\omega)|^2$$

$$\left. -A(t-\tau, \Delta\omega)B(t-\tau, \Delta\omega)]\right\}, \quad (10)$$

where $\dot{\Phi}(\tau) = d\Phi(\tau)/d\tau$.

In the case of a stepwise mechanical shift of the source at the instant $t = 0$ with a jump of duration $\Delta t \ll (T_M \Gamma)^{-1}$, Eq. (10) simplifies appreciably and takes the following form:

$$I(t, \Delta\omega) = 2f_s e^{-\Gamma_a t} \mathrm{Re}[(1-\Phi(t))A(t, \Delta\omega)B(t, \Delta\omega)]. \quad (11)$$

It follows from Eqs. (10) and (11) that taking account of the finite duration Δt (compared with the case $\Delta t = 0$) causes a reduction in peak intensities, to an increase in the duration of the transient pulses and of their fronts, and under certain conditions also to the appearance of additional oscillations.

A spread in the mechanical shifts $\Delta x(r)$ within the volume of the γ source caused by modulation amplifies these tendencies even further. We have shown[3] that in the case of in-phase modulation in the modulators we used with a thin source (absorber), the real shifts are satisfactorily described by a cosine radial distribution

$$\Delta x(r) = \Delta x_m \cos\frac{\pi}{2}\frac{r}{r_0}$$

over the source surface (r_0 is the substrate radius, Δx_m is the shift at the center $r = 0$). The spread indicated was calculated by averaging the phase factor $\Phi(\tau, r)$ over the area of the active spot of the γ source within the limits of the radius r_n.

In the present work we have neglected the resolving time of the apparatus ($\leqslant 5$ ns for 14.4 keV ^{57}Fe), since it is appreciably (three times) less than the rise time of the pulses excited.

EXPERIMENTAL INVESTIGATION OF TRANSIENT PULSES AND ANALYSIS OF THE RESULTS

Experiments on stepwise modulation of ^{57}Fe γ radiation were carried out with a ^{57}Co(Pd) source (10 mC) and a K_4^{57}Fe(CN)$_6 \cdot 3H_2O$ absorber ($T_M = 8 + 2$). The stationary value of the resonance effect was $\varepsilon_0 = 0.37 \pm 0.02$; the γ source (diameter 8×0.02 mm, $r_n/r_0 = 0.67$), stuck to a disk of x-cut quartz (diameter 25×0.25 mm), served as one of the electrodes of a piezoelectric transducer. The other electrode was a large aluminum plate. Stepwise modulation was acheived by feeding rectangular wave voltage pulses to the transducer (duration ~ 3000 ns) with a steep front and decay (~ 15 ns). The repetition frequency was 123.5 kHz and the amplitude U was varied from 0 to 40 V.

The amplitudes of the pulses in U were maintained constant during the course of the experiment (5–6 h) to an accuracy of ± 0.2 V. Different values of the isomer shift were
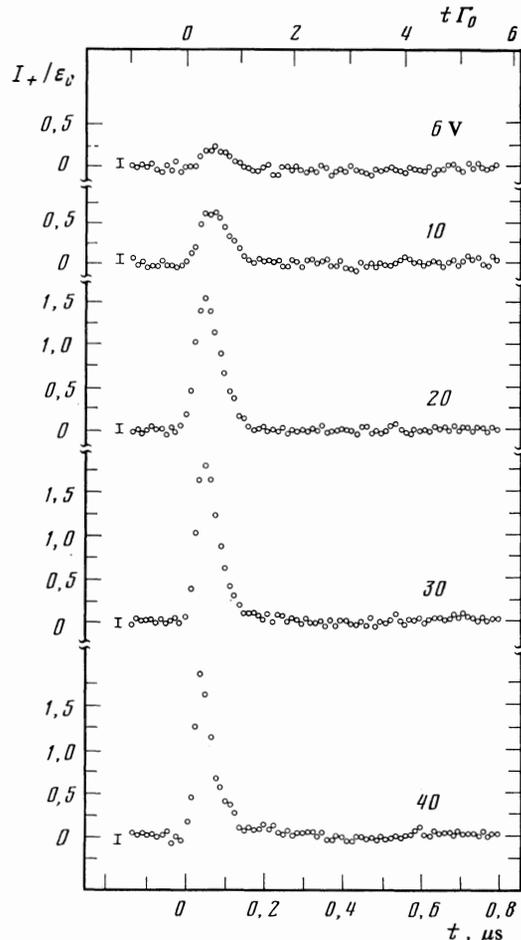

FIG. 1. Transient pulses of 14.4 keV ^{57}Fe resonance γ radiation (γ source ^{57}Co(Pd), absorber K_4^{57}Fe(CN)$_6 \cdot 3H_2O$, $T_m \approx 8$) for different pulse amplitudes U exciting the piezoelectric transducer and $s \sim \Delta\omega = 0$. The intensities $I_+(U)$ are expressed in units of ε_0.

given by a slow (10 Hz) doppler shift of the source, to an accuracy of ± 0.01 mm $\cdot$ s^{-1}, and were checked with the help of a laser velocity calibrator.

The time dependence of the intensity of the collimated γ-quanta beam traversing the absorber relative to the front of the excited voltage pulses was recorded with a time-differential Mössbauer spectrometer with a time-amplitude converter.[11] The resolving time of the apparatus was determined by the "transitory" curves of the 12–14 keV gamma–x-ray coincidences of the ^{88}Y isotope and was 4.8 ± 0.2 ns. A control experiment with sinusoidal excitation of the modulator at a frequency of 11 MHz showed that almost complete ($\geqslant 95\%$) in-phase modulation was achieved. Since the symmetry of the transient pulses relative to the sign of the phase jump, which comes out of the theory, really occurs, we only show the curves corresponding to positive phase shift ($+a$), a shift Δx of the source in the direction of the absorber.

A series of transient pulses of 14.4 keV ^{57}Fe γ-radiation is shown in Fig. 1, measured at several values of the ampli-
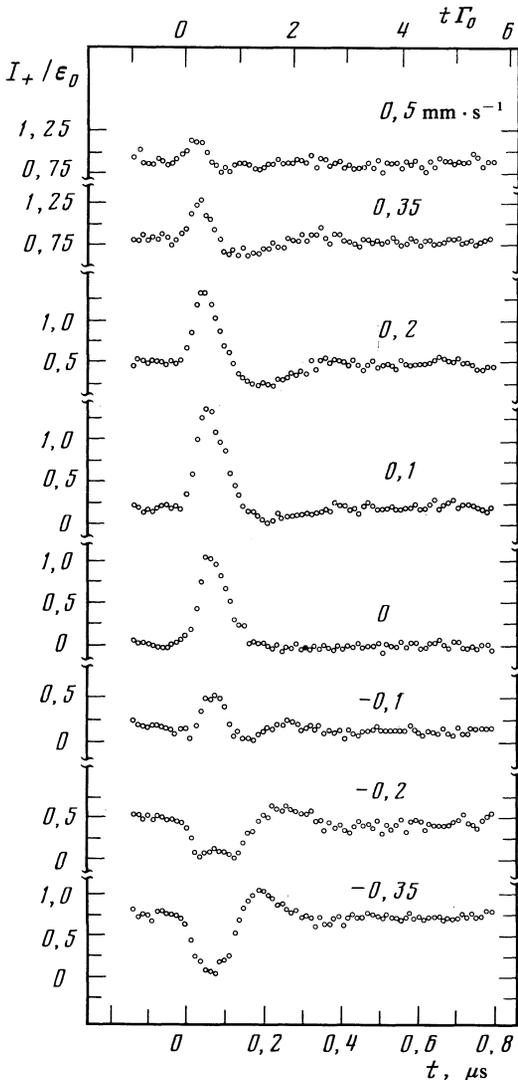
tudes U in the case of zero shift of the lines $S \sim \Delta\omega = 0$. A stepwise phase shift produces a jump in intensity which grows with increasing amplitude up to $U = 25$ V and then stays practically constant (relative to the background): $I_+(0) = 2.08$. Relative to the stationary effect this value becomes $1.84 \, \varepsilon_0$. An increase in U is also accompanied by a change in the temporal characteristics; the transient pulse becomes shorter and its maximum is shifted towards shorter times. The rise time decreases from 40 ± 2 ns for $U = 10$ V to 27.5 ± 2 ns for 40 V and the time to the half height from 78 ± 4 ns ($U = 15$ V) to 48 ± 4 ns ($U = 40$ V). The reduction in the duration of transient pulses below τ_0 and the increase of intensity above ε_0 found experimentally indicate that coherent transient effects are realized for the ^{57}Fe isotope.

The change in magnitude and sign of the isomer shift S of the lines has a strong influence on the shape and intensity of the transient pulses. The basic rules for such shifts for $U = 15$ V are shown in Fig. 2. In the region $S < 0$, the intensity has a maximum value $I_+ = 1.38 \, \varepsilon_0$ at $S = -0.1$ mm $\cdot$ s^{-1}, exceeding the corresponding value of 107 ε_0 for $S = 0$. In the case of $S > 0$ and increase in S leads to an approximately linear reduction in intensity; the transient pulse becomes negative relative to the intensity before the phase jump. Oscillations are observed with a contribution to the intensity which grows for large values of $|S|$.

In order to study the shape in more detail, transient pulses were plotted with improved reolution time (Fig. 3). It follows from the experimental results that the pulses ob-

FIG. 2. Transient pulses for different isomer shifts of the lines S and constant $U = 15$ V.
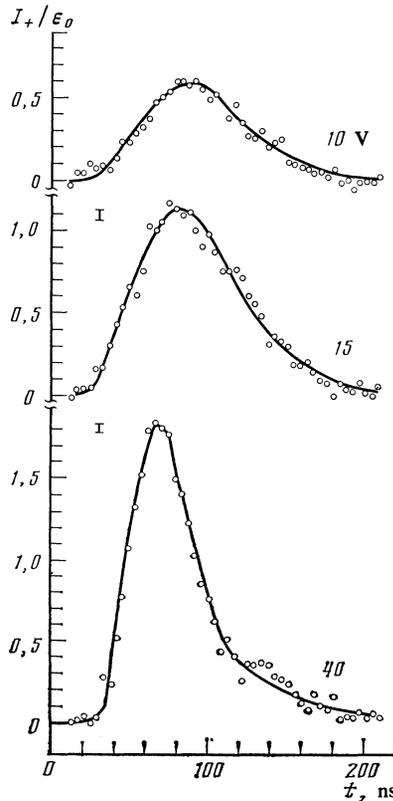

FIG. 3. Comparison between the calculated (full lines) and experimental (circles) shapes of the transient pulses for $U = 10$, 15 and 40 V and $S = 0$ (see text).
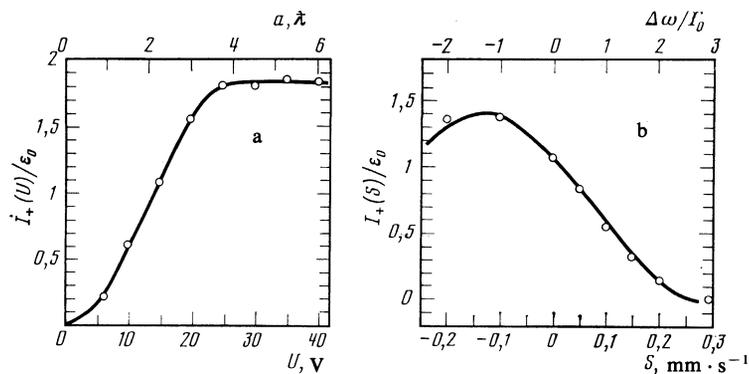
FIG. 4. The dependence of the intensity of the transient pulses a) on the amplitude of the exciting pulses $U \sim a_m$ ($\bar{\lambda} = 0.137$ Å) for $S = 0$; b) on the isomer shift of the lines $S \sim \Delta\omega$ for $U = 15$ V. The full lines correspond to the curves calculated according to Eq. (10) (see text).

tained and their characteristics cannot be explained quantitatively within the framework of pure phase modulation ($\Delta t = 0$). It is essential to take account of the finite time for the mechanical shift of the source and the spread in it. The absence of independent data on the real γ-source shift forces us to use the following simplified model of the stepwise modulation. A voltage pulse with a steep front produces a mechanical shift $\Delta x = d_{11} U$ ($d_{11} = 0.021$ Å/V is the piezoelectric modulus of the x-cut quartz). Due to the inertia of the γ-source and the loading of the modulator, the displacement of the source up to Δx_m takes place at constant velocity during an interval $\Delta t =$ const. for any values of U. In this case the phase discontinuity for the center of the γ source takes the form

$$a_m(t) = \begin{cases} 0, & t<0, \\ d_{11}Ut/\bar{\lambda}\Delta t, & 0 \leqslant t \leqslant \Delta t, \\ d_{11}U/\bar{\lambda}, & t>\Delta t. \end{cases}$$

The sinusoidal radial distribution $a(t,R)$ over the limits of the active spot of the thin γ source is evaluated by averaging

$$\langle \cos a(t) \rangle = \frac{2}{R_n^2} \int_0^{R_n} \cos\left[a_m(t) \cos\frac{\pi}{2}R \right] R\, dR,$$

$$\langle \sin a(t) \rangle = \frac{2}{R_n^2} \int_0^{R_n} \sin\left[a_m(t) \cos\frac{\pi}{2}R \right] R\, dR,$$

where $R = r/r_0$ and $R_n = r_n/r_0 = 0.67$ for the source which we used. In such a model, Δt remains the only parameter and to choose it we started from a calculation of the minimum build-up time $\Delta t \gtrsim 44$ ns of the modulator used under ideal conditions.

Good agreement between the shapes of the transient pulses calculated form Eq. (10) and experimental shapes was obtained for $\Delta t = 80$ ns (Fig. 3, solid lines). Overall agrement of the measured and calculated values was also observed for the dependence of the intensity of the transient pulses on the voltage $U \sim a_m$ (Fig. 4a) and on the isomeric shifts of the lines $S \sim \Delta\omega$ (Fig. 4b). In this way it can be concluded that the relatively crude model proposed for the modulator is adequate to describe the main relationships of the coherent transient effects and at the same time confirms the correctness of the theoretical discussion given.

As follows from Fig. 4, the $I_+(U,S = $ const.) relation can be used for an accurate measurement of the amplitude of short-lived (10^{-9}–10^{-7} s) periodical mechanical shifts with-

in the limits 0–1 Å, while the $I_+(U = $ const, S) relation contains the information on low-frequency shifts (or on slow relative velocities) within the limits of ± 150 m $\cdot$ s^{-1}. A certain inconvenience in the calibration, consisting in the nonlinearity of the relations mentioned, is compensated by the possibility of their theoretical derivation.

Periodic transient pulsing of the 14.4 keV ^{57}Fe resonance γ radiation, obtained from stepwise modulation, has appreciable advantages (high intensity and small duration) over γ choppers of the mechanical[12,13] and magnetic[14] types. At the same time, the characteristics of the transient pulses can be improved appreciably by refining the modulator design. It is convenient to apply coherent transient effects of ^{57}Fe resonance γ radiation for temporal studies of Mössbauer emission, scattering, diffraction, etc., where the use of the delayed γ–γ coincidence technique is made difficult by its characteristic limitations by the source activity.

[1] S. L. Ruby and D. I. Bolef, Phys. Rev. Lett. **5**, 5 (1960); A. R. Mkrtchyan, A. R. Arakelyan, G. A. Arutyunyan, and A. A. Kocharyan, Pis'ma Zh. Eksp. Teor. Fiz. **26**, 599 (1977) [JETP Lett. **26**, 449 (1977)].

[2] J. E. Monahan and G. J. Perlow, Phys. Rev. A**20**, 1499 (1979).

[3] R. Koch and E. Realo, Proc. Int. Conf. on Applications of the Mössbauer Effect, Alma-Ata, 1983 (in print).

[4] P. Helistö, T. Katila, W. Potzel, and K. Riski, Phys. Lett. A**85**, 177 (1981).

[5] P. Helistö, E. Ikonen, T. Katila, and K. Riski, Phys. Rev. Lett. **49**, 1209 (1982).

[6] P. Helistö, E. Ikonen, T. Katila, and K. Riski, Prelim. Rep. TKK-F-A551, Helsinki Univ. of Technology, Helsinki (1984).

[7] R. Koch, E. Realo, K. Rebane, and J. Jõgi, Proc. Int. Conf. on Applications of the Mössbauer Effect, Alma-Ata, 1983 (in print).

[8] N. Bloembergen, E. M. Purcell, and R. V. Pound, Phys. Rev. **73**, 679 (1948).

[9] A. Z. Genack, D. A. Weitz, R. M. Macfarlane, R. M. Shelby, and A. Schenzle, Phys. Rev. Lett. **45**, 438 (1980).

[10] F. J. Lynch, R. E. Holland, and M. Hammermesh, Phys. Rev. **120**, 513 (1960).

[11] R. Koch and E. Realo, Izv. Akad. Nauk. Est. SSR Ser. Fiz.-Mat. Nauk **28**, 374 (1979).

[12] S. L. Ruby, R. S. Preston, C. E. Skov, and B. J. Zabransky, Phys. Rev. A**8**, 59 (1973).

[13] A. N. Artem'ev, K. P. Aleshin, V. V. Sklyarevskiĭ, and E. P. Stepanov, Prib. Tekh. Eksp. No. 4, 82 (1978) [Instrum. Exp. Tech. **21**, 924 (1978)].

[14] G. V. Smirnov, Yu. V. Shvyd'ko, O. S. Kolotov, V. A. Pogozhev, M. Kotrbova, S. Kadechkova, and I. Novak, Zh. Eksp. Teor. Fiz. **86**, 1495 (1984) [Sov. Phys. JETP **59**, 875 (1984)].

Translated by R. Berman

Theory of the self-trapping rate

A. S. Ioselevich and É. I. Rashba

L. D. Landau Institute of Theoretical Physics, Academy of Sciences of the USSR
(Submitted 17 December 1984)
Zh. Eksp. Teor. Fiz. **88**, 1873–1897 (May 1985)

A theory is derived for the temperature dependence of the rate of self-trapping in the exponential approximation. This rate is determined by the process in which the composite system consisting of the lattice and the exciton (or electron) passes over a self-trapping barrier in configuration space. There are two distinct mechanisms for getting over the self-trapping barrier. At $T > T_c$ the process is purely one of activation, and its rate is described by $w \propto \exp(-W/T)$, where W is the height of the self-trapping barrier, and the critical temperature T_c is on the order of phonon frequencies. At $T < T_c$, in contrast, an instanton mechanism operates and describes thermally activated tunneling. At $T \approx T_c$ we should observe a slope change on the $w = w(T)$ curve, becoming more abrupt with increasing W/T_c. Solutions are constructed for three models: a small-radius exciton (or charge carrier) interacting with acoustic phonons, the same, interacting with optical phonons, and a Wannier-Mott exciton interacting with polar phonons. At low temperatures $(T < T_c)$ we have $\ln w \propto (b_1 + b_2 T^4)$ for acoustic phonons and $\ln w \propto (c_1 + c_2 \exp(-w_0/T))$ for optical phonons of frequency ω_0. The low-temperature behavior of the exponential factor in w is thus determined by acoustic phonons, regardless of which phonons are responsible for most of the self-trapping. At the lowest temperatures T the functional dependence $w(T)$ is determined by the coefficient of the exponential function. After passage over the barrier, the system collapses, and the energy which is released goes into the kinetic energy of several atoms. This process may terminate in defect formation.

1. STATEMENT OF THE PROBLEM AND BASIC RESULTS

In crystals exhibiting a strong interaction between quasiparticles (holes, electrons, or excitons) and phonons, self-trapping is possible. If the coupling with the phonons is a short-range phenomenon, a free state persists as a metastable state, in addition to the self-trapping state (the lowest-lying state). These two states are separated by a self-trapping barrier (see the review by Rashba[1]). Since excitons are electrically neutral, their coupling with phonons is always of short range, even for polar phonons. Since essentially all experiments on the coexistence of free and self-trapping states (separated by a barrier) have been carried out with excitons, we will for definiteness speak exclusively in terms of excitons below, but all the results apply equally well to holes and electrons which are interacting with nonpolar phonons.

Self-trapping occurs in two steps: passing over the barrier and subsequent relaxation of the highly nonequilibrium exciton-phonon system to the lowest-lying self-trapping state. In principle, either of these steps may be the "bottleneck," but we will assume everywhere below that the rate of the overall process is controlled by the first step.

The probability for tunneling through a barrier at a temperature $T = 0$ was calculated by Iordanskiĭ and Rashba[2] in the exponential approximation. In the present paper we determine the changes produced in the passage over the barrier by a nonzero temperature. For this purpose we use the mathematical apparatus developed by Ioselevich[3,4] in a theory for the long-wave fundamental absorption edge. The results of this analysis have been summarized elsewhere.[5] In the present paper we restrict the analysis to the exponential approximation in the calculation of the self-trapping rate w; we will analyze the coefficient of the exponential factor in w in a separate paper.

The self-trapping barrier is a potential barrier not only for excitons but also for the lattice, i.e., for a system having a large number of degrees of freedom. Such a system can cross a barrier either by tunneling or by an activation process. As it climbs over a barrier, an exciton adiabatically follows the lattice deformation, and the adiabatic potential of the lattice determines the dynamics of the exciton. A barrier can substantially limit the rate of the process only if its height W exceeds the characteristic frequency of the phonons, $\bar{\omega}$. In this case the motion of the lattice near the barrier can be described semiclassically, and the optimum path is determined from the condition for a steady state action functional S.

We can summarize the results of this study as follows: There exist three types of extremals of the functional S and, correspondingly, three competing paths in phase space. These paths can be sketched as in Fig. 1. In this figure, we have retained only two of the infinite number of degrees of freedom: R, the scale size of the deformation region, and $X = QR^{3/2}$, where Q is a characteristic displacement of the nuclei in a region $\sim R^3$. The total potential energy of the lattice, U, including the exciton-phonon interaction, is plotted as a function of R and X. This particular normalization of the coordinate X is convenient because it allows us to write an expression in the usual form for the truncated action of the entire system:

$$S_0 \sim \int d\mathbf{r} \int dQ(\mathbf{r}) \, [2U(\mathbf{r})]^{1/2} \sim \int dX [2U(X)]^{1/2},$$
$$U(X) \sim \int d\mathbf{r} \, U(\mathbf{r}). \qquad (1)$$

The line $X = 0$ corresponds to the undeformed lattice (on which we have $U = 0$). The self-trapping state lies in the "ravine" separated from the free states by a potential barrier.
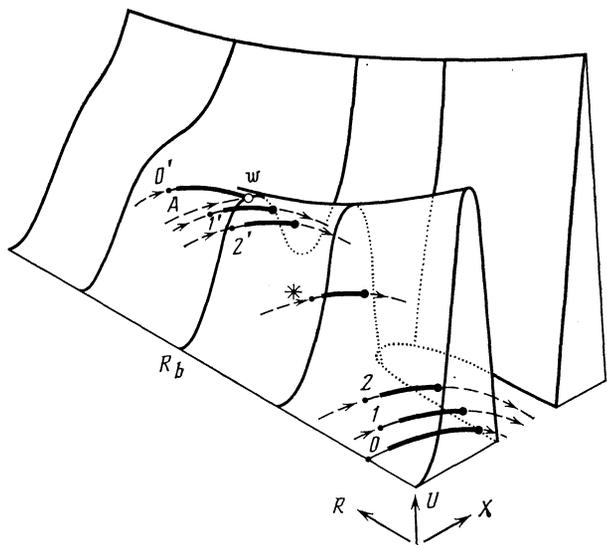
FIG. 1. Passage over the self-trapping barrier. U—Potential energy; X—deformation; R—size of the deformed region. Paths 0, 0') The temperature $T = 0$; 1, 1'), T_1; 2, 2') $T_2 > T_1$; *) the path corresponding to T*. Path A exists at all T. Dashed lines: Regions of the paths which lie in the classically accessible region; solid lines—regions of tunneling; light lines—free tunneling of the lattice; heavy lines—tunneling with a trapped exciton.

At the saddle point $\mathscr{W}$, the height of the potential barrier is at its minimum value W, which we will call the "height of the self-trapping barrier." The extremal path A, which corresponds to getting over the barrier by activation, passes through this point. This path does not depend on T. At $T = 0$, the tunneling occurs along path 0, which maximizes the tunneling transmission of the barrier, $\mathscr{D}$, when the total energy of the system is $\varepsilon = 0$. Because of the multidimensional nature of the configuration space, path 0 may lie at values of R quite different from the value (R_b) which corresponds to the point $\mathscr{W}$. For example, for acoustic phonons the barrier contracts markedly with decreasing R, so that the tunneling occurs at $R < R_b$ (see Ref. 2 and §6 of the present paper), despite the fact that the barrier is higher in this region. For optical phonons (§4), in contrast, the barrier broadens with decreasing R, and the tunneling occurs at $R \sim R_b$.

It is clear from the equations of Ref. 2 that there is no direct relationship between the barrier height W and the self-trapping rate w. This has apparently been established experimentally. According to results reported by Unuma et al.[6] for the two iodides KI and RbI, the barrier height W is lower in RbI, but the value of w at $T \approx 0$ is nevertheless lower.

At $T \neq 0$ tunneling occurs at energies $\varepsilon > 0$. With increasing ε, the transmission of the barrier increases exponentially, $\mathscr{D} = \exp(-2S_0(\varepsilon))$, but there is a simultaneous decrease in the Gibbs factor $e^{-\beta\varepsilon}$, where $\beta = T^{-1}$. We thus have

$$w \propto \exp(-S), \quad S = 2S_0(\varepsilon) + \beta\varepsilon. \quad (2)$$

The optimum value of ε is determined from the condition $\partial S/\partial\varepsilon = 0$ or $\partial S_0/\partial\varepsilon = -\beta/2$. The derivative $\partial S_0/\partial\varepsilon = -\tau$ is the imaginary time spent by the system below the barrier (because of the transformation to an imaginary time, this equation differs by a sign from the well-known

equation from analytic dynamics). We thus have $\tau = \beta/2$, and S has the meaning of the total action Hamiltonian for a back-and-forth crossing of the barrier in a time β. We will call $S = S(\beta)$ the "action."

In addition to path A, there are two other paths which are extremals of $S(\beta)$; in contrast with A, these two other paths are displaced in configuration space when β changes. On paths 1, 2, ..., the action is lower than on the corresponding paths 1', 2', ... Curve 0' is the limiting path of this latter family as $T \to 0$ (Fig. 1). The action on 0', like that on A, is infinite at $T = 0$. With increasing T, the two stationary paths move closer together. At $T^* \sim \bar{\omega}$ they merge and disappear. At $T > T^*$, we are left with the unique extremal path A.

Each path consists of two regions. In the first region, a freely moving lattice undergoes tunneling; the exciton remains in a free state and has essentially no effect on the motion of the lattice. In the second region, the exciton is in a local level and interacts strongly with the lattice. The time τ_0 of the motion in this second region is short for paths 0, 1, ...: $\tau_0 < \bar{\omega}^{-1}$ (a short instanton). On paths 0', 1', 2', ... we have $\tau_0 \approx \beta/2$ at $T \ll \bar{\omega}$; i.e., this time is essentially the same as the tunneling time (a long instanton).

Figure 2 shows the temperature dependence $S(\beta)$ for the three extremal paths. The overall temperature dependence seems to be generally the same, although the curves plotted here have been calculated for a specific problem, the interaction with nonpolar optical phonons[1] (§4). At temperatures $T > T_c$, path A wins the competition; on this path, the action reaches an absolute minimum W/T. According to (2), therefore, an Arrhenius activation law holds at $T > T_c$:

$$w \propto \exp(-W/T). \quad (3)$$

In the continuum approximation we have $W \approx 11/m^3v^2\omega_0^2\gamma^4$, for the nonpolar optical phonons and $W \approx 44\rho^2s^4/m^3C^4$ for acoustic phonons (Ref. 7). Here m is the effective mass of the exciton, v is the volume per atom, ω_0 is the frequency of the optical phonons, γ is the dimensionless coupling constant with these phonons, ρ is the density, s is the sound velocity, and C is the strain energy.
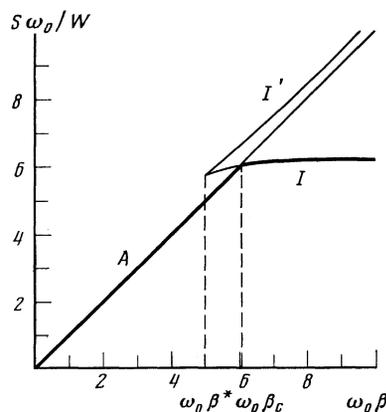

FIG. 2. Temperature dependence of the action for various solutions. I, I'—Short and long instantons; A—activation solution. $\beta^* = 1/T^*$, $\beta_c = 1/T_c$, $T^* > T_c$. The heavy lines correspond to the absolute minimum of the action.

A. S. Ioselevich and É. I. Rashba   1111

At $T < T_c$, the short instanton I, which describes thermally activated tunneling, wins the competition. In the low-temperature limit we have

$$S_I(T) \approx S_I(0)\{1 - b e^{-\omega_0/T}\},$$
$$S_I(0) \approx 6.2W/\omega_0, \quad b \approx 10.3, \tag{4}$$

for the nonpolar optical phonons or

$$S_I(T) \approx S_I(0)\{1 - b(T/\bar\omega)^4\},$$
$$S_I(0) \approx 2.8\rho s/m^2 C^2 \tag{5}$$

for the acoustic phonons. Here b is a dimensionless combination of parameters [see Eq. (68)].

Without reference to these models, under some extremely general assumptions, law (3) holds at high temperatures, while at low temperatures we have a T^4 law for S, since the low-temperature behavior is always determined by acoustic phonons. The only restriction on the applicability of expression (5) on the low-temperature side stems from the condition $S_I(0)b(T/\bar\omega)^4 > 1$; at lower temperatures T, the temperature dependence of the coefficient of the exponential function should dominate.

Expressions (3)–(5) for $S_I(T)$ hold for thermalized excitons, which have the temperature of the lattice. In contrast, in many experiments the excitons are produced with large initial energies, and in such a case the self-trapping rate may exceed the energy relaxation rate of the free excitons, and w will increase rapidly with the energy of the exciton. Evidence for this conclusion comes from, for example, the data of Roick et al.[8] on Xe and the data of Lushchik et al.[9] on CsBr. For hot excitons the self-trapping rate is not described by expressions (3)–(5); we will examine the case separately.

The problem of getting over the barrier is closely related to several other problems, primarily problems in quantum nucleation in phase-transition theory[10,11] and problems of the decay of a false vacuum in quantum field theory.[12] For this reason, the successful methods of the self-trapping theory based on a generalization of Refs. 2–4 to some extent border on the methods used by Langer[10] and Coleman.[12] In the self-trapping problem, however, there is a distinctive feature: The exciton remains free in the first stage of the tunneling of the lattice and is then trapped in the second state (thereby changing the adiabatic potential). This distinctive feature has several qualitative consequences.

Nasu and Toyozawa[13] describe the self-trapping rate on the basis of the theory of many-phonon transitions[14–16] originally worked out for radiationless transitions at impurity centers. That theory ignores the dependence of the electron wave function ψ on the lattice configuration, i.e., a dependence which plays a decisive role in the theory and which forms the main exponential factor in $\mathscr{D}$. In the absence of a correct procedure for determining the exponential factor, the expression derived in Ref. 13 for the coefficient of the exponential factor is obviously even less likely to inspire much confidence. In addition, the general approach of Ref. 13—an effort to derive a universal quantitative description of the processes which occur at the scale of the lattice constant—is clearly unrealistic. That such efforts are illusory is

demonstrated by, for example, the analysis in Ref. 6 of the experimental data.

2. GENERAL EXPRESSION FOR THE SELF-TRAPPING RATE

The number of transitions from the free state with momentum $\mathbf{k}$ to the self-trapping state (ST) during the short time interval $t_2 - t_1$ is equal to the square modulus of the transition amplitude averaged over the phonons;

$$w(\mathbf{k})(t_2 - t_1) = \langle |\mathscr{A}((ST)t_2, \mathbf{k}t_1)|^2 \rangle_{ph}$$
$$= \int d\mathbf{r}_1 \, d\mathbf{r}_2 \, d\mathbf{r}_1' \, d\mathbf{r}_2' \, \Psi_{ST}^*(\mathbf{r}_2 t_2)$$
$$\times \Psi_{ST}(\mathbf{r}_2' t_2) F(\mathbf{r}_2 t_2, \ \mathbf{r}_2' t_2; \ \mathbf{r}_1 t_1, \ \mathbf{r}_1' t_1) \Psi_{\mathbf{k}}(\mathbf{r}_1 t_1) \Psi_{\mathbf{k}}^*(\mathbf{r}_1' t_1). \tag{6}$$

Here $\psi_{\mathbf{k}}$ and ψ_{ST} are the exciton wave functions of the free and self-trapping states, and

$$F(\mathbf{r}_2 t_2, \mathbf{r}_2' t_2; \mathbf{r}_1 t_1, \mathbf{r}_1' t_1)$$
$$= \sum_{if} \exp\{\beta(\mu - \varepsilon_i)\} \langle i | \hat\psi(\mathbf{r}_1' t_1) \hat\psi^+(\mathbf{r}_2' t_2) | f \rangle$$
$$\times \langle f | \hat\psi(\mathbf{r}_2 t_2) \hat\psi^+(\mathbf{r}_1 t_1) | i \rangle. \tag{7}$$

State i and f correspond to purely phonon excitations. It is assumed that the phonons in the initial state, i, are in equilibrium; $e^{\beta\mu}$ is a normalization factor; and the $\hat\psi$ are the exciton annihilation operators.

The function F can be expressed in terms of the exciton-phonon Green's functions G and the free phonon propagators K:

$$F(\mathbf{r}_2 t_2, \mathbf{r}_2' t_2; \mathbf{r}_1 t_1, \mathbf{r}_1' t_1) = e^{\beta\mu} \text{Tr} \{ G^R(\mathbf{r}_2 t_2, \mathbf{r}_1 t_1)$$
$$\times K(t_1 - i\beta, t_1) G^A(\mathbf{r}_1' t_1, \mathbf{r}_2' t_2) \}. \tag{8}$$

The functions involved here can be written as functional integrals as follows[3]:

$$G^R(Q_2 \mathbf{r}_2 t_2, Q_1 \mathbf{r}_1 t_1)$$
$$= -i \int \mathscr{D}Q \mathscr{D}\Psi_R \ \Psi_R(\mathbf{r}_2 t_2) \Psi_R^*(\mathbf{r}_1 t_1) \exp\left\{ i \int_{t_1}^{t_2} dt L_{tot}(\Psi_R, Q) \right\}, \tag{9}$$
$$\{Q(t_2) = Q_2, \ Q(t_1) = Q_1\};$$
$$G^A(Q_1' \mathbf{r}_1' t_1', Q_2' \mathbf{r}_2' t_2') = i \int \mathscr{D}Q \mathscr{D}\Psi_A \ \Psi_A(\mathbf{r}_1' t_1') \Psi_A^*(\mathbf{r}_2' t_2')$$
$$\times \exp\left\{ i \int_{t_2'}^{t_1'} dt L_{tot}(\Psi_A, Q) \right\}, \tag{10}$$
$$\{Q(t_2') = Q_2', \ Q(t_1') = Q_1'\};$$
$$K(Q_2' t_2', Q_2 t_2) = \int \mathscr{D}Q \exp\left\{ i \int_{t_2}^{} dt L_{lat}(Q) \right\}; \tag{11}$$
$$\{Q(t_2') = Q_2', \ Q(t_2) = Q_2\};$$

Here L_{lat} and L_{tot} are the Lagrangian of the free lattice and the total Lagrangian of the exciton-phonon system, given by

$$L_{lat} = \frac{1}{2} \int \frac{d\mathbf{q}}{(2\pi)^3} \{ \partial_t Q_{\mathbf{q}} \partial_t Q_{-\mathbf{q}} - \omega_{\mathbf{q}}^2 Q_{\mathbf{q}} Q_{-\mathbf{q}} \}, \tag{12}$$

A. S. Ioselevich and É. I. Rashba   1112

$$L_{tot} = L_{lat} + \int d\mathbf{r} \left\{ i \Psi^* \partial_t \Psi + \frac{1}{2m} \Psi^* \nabla^2 \Psi \right.$$

$$\left. - \int \frac{d\mathbf{q}}{(2\pi)^3} \gamma_q Q_q e^{i\mathbf{q}\mathbf{r}} | \Psi(\mathbf{r}t) |^2 \right\}; \quad (13)$$

ω_q and γ_q are the phonon frequencies and the exciton-phonon coupling coefficients; m is the effective mass of an exciton; and we are setting $\hbar = 1$. A field integration over ψ is carried out in (9) and (10). The absence of a nontrivial (i.e., Q-dependent) normalization factor is a consequence of the circumstance that this is a one-particle problem with respect to the exciton, so that there is no contribution from vacuum loops. This assertion is proved in Ref. 17. An ordinary path integral is carried out over Q (Ref. 18). This mixed path-field representation of the functional integral in this problem is advantageous because the motion of the lattice differs in nature from that of the exciton. Specifically, the motion of the lattice is semiclassical, so that we can speak in terms of definite values of the displacements Q at each time. The motion of the exciton, in contrast, is quantized, and it is described by an adiabatic wave function; it is thus convenient to speak in terms of the values of ψ at each instant. Since we will be using the method of steepest descent below, these quantities are natural choices as the variables for the functional integration. We have omitted Q from the arguments of all the functions in (8); we take a convolution of the Green's functions with respect to this displacement, and we calculate the trace. Substituting (9)–(11) into (8), we find

$$F(\mathbf{r}_2 t_2, \mathbf{r}_2' t_2; \mathbf{r}_1 t_1, \mathbf{r}_1' t_1)$$

$$= \exp(\beta\mu) \int \mathscr{D}Q \mathscr{D}\Psi_A \mathscr{D}\Psi_R \Psi_R(\mathbf{r}_2 t_2) \Psi_R^*(\mathbf{r}_1 t_1)$$

$$\times \Psi_A(\mathbf{r}_1' t_1) \Psi_A^*(\mathbf{r}_2' t_2) \exp\left\{ i \int_\Gamma dt L(\Psi Q) \right\}, \quad (14)$$

$$\{ Q(t_1) = Q(t_1 - i\beta) \}.$$

The contour Γ is shown in Fig. 3a. The direction of the integration along Γ, which is the same as the time-ordering direction, is indicated by the arrows. The shape of the contour is the same as that used by Konstantinov and Perel'.[19] As the Lagrangian L we should use L_{tot} on the parts of the contour shown in solid line, while on the dashed parts we should use L_{lat}. The contour and the analogous expression for the four-time correlator describing the shape of the secondary-emission bands are given in Ref. 5.

The function F is a two-particle Green's function averaged over the phonons. This is to be expected, since this function arises when an average is taken of the square amplitude for the transition from one fixed state (a state of a free exciton with a momentum $\mathbf{k}$) to another fixed state of an exciton (a self-trapping state). As usual, the single-particle Green's function incorporates information only on the total probability for the decay of state $\mathbf{k}$, i.e., on the sum of the probabilities for transitions to all final states. As an exception, the single-particle Green's function can be used to calculate w at $T = 0$, since in this case the only possible final state is a self-trapping state. In particular, with respect to exciton spectroscopy the single-particle exciton Green's
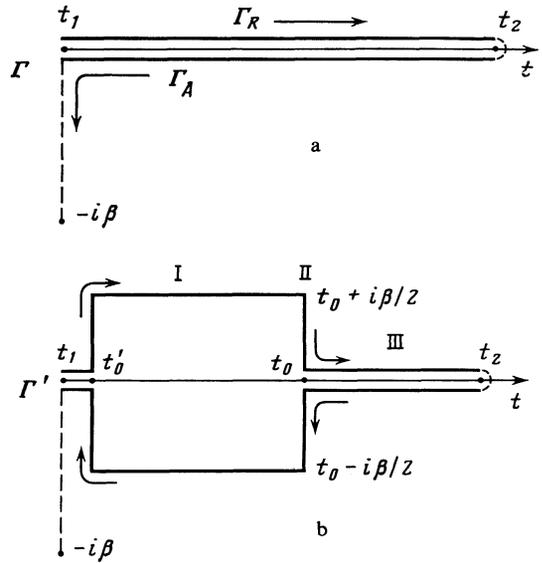

FIG. 3. Integration contours. a—Original contour Γ; b—deformed contour Γ'. The self-trapping time t_0 can be at an arbitrary point in the interval (t_1, t_2).

function gives a complete description of the absorption of light but does not describe the secondary emission (e.g., the luminescence spectrum). For this reason, w could not be derived at $T \neq 0$ by the approach of Ref. 4 in a calculation of the shape of the long-wave tail on the exciton absorption in a system with self-trapping.

Since the exciton-phonon interaction is linear in the displacements, we can carry out a Gaussian integration[18] over Q. For this purpose we must make a small displacement in (14) by the function $\overline{Q}$ defined by

$$\overline{Q}(\mathbf{q}t) = -\frac{1}{2} \gamma_q \int_\Gamma dt' \int d\mathbf{r} \overline{D}(\mathbf{q}, t-t') | \Psi(\mathbf{r}t') |^2 \exp(-i\mathbf{q}\mathbf{r}) \quad (15)$$

or

$$\overline{Q}_\alpha(\mathbf{q}t)$$

$$= \frac{1}{2} \gamma_q \sum_\beta \int_{\Gamma_\beta} dt' \int d\mathbf{r} D_{\alpha\beta}(\mathbf{q}, t-t') | \Psi_\beta(\mathbf{r}t') |^2 \exp(-i\mathbf{q}\mathbf{r}). \quad (16)$$

Here $\overline{Q} = \{ \overline{Q}_\alpha \}$ and $\Psi = \{ \Psi_\alpha \}$ are two-component vectors; we have $\alpha = R$ on Γ_R and $\alpha = A$ on Γ_A (Fig. 3a). On the vertical part of contour Γ_A we have $\Psi_A = 0$. The matrix Green's function $\widehat{D}$ is defined by the equation

$$\widehat{L}_q \widehat{D}(\mathbf{q}, t-t') = \delta(t-t'), \quad (17)$$

where

$$\widehat{L}_q = -\frac{1}{2} (\partial_t^2 + \omega_q^2) \hat{\sigma}_z. \quad (18)$$

Equation (17) is solved under boundary conditions which ensure a continuous joining of D and $\partial_t D$ at the ends of contour Γ at an arbitrary t':

$$D_{R\alpha}(t_1, t') = D_{A\alpha}(t_1 - i\beta, t'),$$

$$\partial_{t_1} D_{R\alpha}(t_1, t') = \partial_{t_1} D_{A\alpha}(t_1 - i\beta, t'). \quad (19)$$

The function D is

$$D_{\alpha\beta}(\mathbf{q}, t_\alpha, t_\beta') = -(i/\omega_\mathbf{q})\{(N_\mathbf{q}+1)\exp(-i\omega_\mathbf{q}[t_\alpha-t_\beta']_\Gamma)$$

$$+N_\mathbf{q}\exp(i\omega_\mathbf{q}[t_\alpha-t_\beta']_\Gamma)\}, \qquad (20)$$

where $[t_\alpha - t_\beta']_\Gamma$ is the difference between late and early times in the sense of the order on contour Γ, and the $N_\mathbf{q}$ are the phonon filling numbers. The function $D_{\alpha\beta}$ is analogous to that which figures in the Keldysh diagram technique.[20]

After the small displacement and the Gaussian integration over $Q - \hat{Q}$, expression (14) becomes

$$F(\mathbf{r}_2 t_2, \mathbf{r}_2't_2; \mathbf{r}_1 t_1, \mathbf{r}_1't_1)$$

$$= \int \mathscr{D}\Psi_A \mathscr{D}\Psi_R\, \Psi_R(\mathbf{r}_2 t_2)\Psi_R^*(\mathbf{r}_1 t_1)\Psi_A(\mathbf{r}_1't_1)$$

$$\times \Psi_A^*(\mathbf{r}_2't_2)\exp(i\mathscr{S}), \qquad (21)$$

where

$$\mathscr{S} = \int d\mathbf{r}\sum_\alpha \int_{\Gamma_\alpha} dt\, \Psi_\alpha^*\left(i\partial_t + \frac{1}{2m}\nabla^2\right)\Psi_\alpha$$

$$-\frac{1}{4}\sum_{\alpha\beta}\int_{\Gamma_\alpha} dt\int_{\Gamma_\beta} dt'\iint d\mathbf{r}\,d\mathbf{r}'$$

$$\times \int \frac{d\mathbf{q}}{(2\pi)^3}\gamma_\mathbf{q}^2 D_{\alpha\beta}(\mathbf{q}, t-t')\,|\Psi_\alpha(\mathbf{r}t)|^2|\Psi_\beta(\mathbf{r}'t')|^2$$

$$\times \exp[i\mathbf{q}(\mathbf{r}-\mathbf{r}')]. \qquad (22)$$

This expression is exact [for the model Lagrangian (12), (13)]. To derive physical results we need to evaluate the functional integral (21) approximately, making explicit use of the adiabatic parameter. Although the physical picture of self-trapping drawn in §1 was expressed in terms of the configuration coordinates Q, it is more convenient to eliminate Q from the calculations, as we did above, and to work with the action (22), which depends exclusively on Ψ. In this case the self-consistent displacements $\overline{Q}$ are determined in terms of Ψ from (15).

3. METHOD OF STEEPEST DESCENT

Since self-trapping involves getting over a high potential barrier, its probability is exponentially small. It is thus natural to turn to the method of steepest descent to evaluate integral (21).

Variation of the action S [given by (22)] with respect to Ψ^* leads to the nonlinear Schrödinger equation

$$\left\{i\partial_t + \frac{1}{2m}\nabla^2 - V(\mathbf{r}t)\right\}\Psi(\mathbf{r}t) = 0, \qquad (23)$$

where

$$V(\mathbf{r}t) = \int \frac{d\mathbf{q}}{(2\pi)^3}\gamma_\mathbf{q}\overline{Q}(\mathbf{q}t)e^{i\mathbf{q}\mathbf{r}}. \qquad (24)$$

Since V varies adiabatically slowly with t, the solution of the time-dependent equation (23) is

$$\Psi(\mathbf{r}t) = \psi(\mathbf{r}t)\exp\left\{-i\int^t E(t')\,dt'\right\}, \qquad (25)$$

where ψ is the solution of the nonlinear equation

$$\left\{E(t) + \frac{1}{2m}\nabla^2 - V(\mathbf{r}t)\right\}\psi(\mathbf{r}t) = 0; \qquad (26)$$

$E(t)$ and $\psi(\mathbf{r}t)$, which we will choose below to be real, represent an adiabatic energy and an adiabatic wave function of an exciton in the displacement field $\overline{Q}(\mathbf{r}t)$. It was shown in Ref. 3 that the extremals of $\psi(\mathbf{r}t)$ are normalized (to unity). Equations (16) and (26) make the displacements $\overline{Q}$ and the wave function ψ self-consistent. The motion as a function of $\overline{Q}$ is semiclassical (as is usual in an adiabatic theory).

To determine the extremal action, we must deform the contour Γ as shown in Fig. 3b. The physical meaning of the various elements of this contour is as follows: The vertical section at the right, II, which runs along the imaginary-time axis, corresponds to tunneling of the system. In Fig. 1, this section corresponds to the solid-line sections of the paths. Horizontal sections I and III correspond to motion in the classically accessible region from the sides of the free and self-trapping states, respectively.

The deformed contour Γ' is symmetric with respect to the t axis. We can therefore write $\psi_R(t) = \psi_A(t*)$ and $V_R(t) = V_A(t*)$, which are natural from the physical standpoint. Furthermore, at all points on this contour $\overline{Q}$ and V are real. Finally, the entire imaginary contribution to the action comes from the integration over the vertical portion of contour Γ' at the right, where the system is undergoing tunneling. That the horizontal sections do not contribute to Im$\{\widetilde{S}\}$ follows from the reality of V, while the contributions of the symmetric portions to Re$\{\widetilde{S}\}$ cancel out in pairs. The contribution to $\widetilde{S}$ from the (solid-line) vertical portion of the contour at the left is zero, since at early times t_0' the exciton remains free, so that we have $\psi = 0$.

We now use (20) to transform expression (16) for $\overline{Q}$ into forms appropriate for the various parts of contour Γ':

$$\overline{Q}_\mathrm{I}(\mathbf{q}u) = -\frac{\gamma_\mathbf{q}}{\omega_\mathbf{q}}\int d\mathbf{r}e^{-i\mathbf{q}\mathbf{r}}\left\{\int_u^0 du'\sin[\omega_\mathbf{q}(u'-u)]\psi^2(\mathbf{r}u')\right.$$

$$\left.+\frac{\cos(\omega_\mathbf{q}u)}{2\,\mathrm{sh}(\omega_\mathbf{q}\beta/2)}\int_{-\beta/2}^{\beta/2} d\tau'\,\mathrm{ch}(\omega_\mathbf{q}\tau')\psi^2(\mathbf{r}\tau')\right\}, \qquad (27)$$

$$\overline{Q}_\mathrm{II}(\mathbf{q}\tau) = -\frac{\gamma_\mathbf{q}}{2\omega_\mathbf{q}}\int d\mathbf{r}\exp(-i\mathbf{q}\mathbf{r})\int_{-\beta/2}^{\beta/2} d\tau'\{(N_\mathbf{q}+1)$$

$$\times \exp(-\omega_\mathbf{q}|\tau-\tau'|)$$

$$+N_\mathbf{q}\exp(\omega_\mathbf{q}|\tau-\tau'|)\}\psi^2(\mathbf{r}\tau'), \qquad (28)$$

$$\overline{Q}_\mathrm{III}(\mathbf{q}t) = -\frac{\gamma_\mathbf{q}}{\omega_\mathbf{q}}\int d\mathbf{r}\exp(-i\mathbf{q}\mathbf{r})\left\{\int_{t_0}^t dt'\sin[\omega_\mathbf{q}(t-t')]\psi^2(\mathbf{r}t')\right.$$

$$+\frac{1}{2}\cos(\omega_\mathbf{q}t)\int_{-\beta/2}^{\beta/2} d\tau'[(N_\mathbf{q}+1)\exp(-\omega_\mathbf{q}|\tau'|)$$

$$\left.+N_\mathbf{q}\exp(\omega_\mathbf{q}|\tau'|)]\psi^2(\mathbf{r}\tau')\right\}. \qquad (29)$$

In region I the time is $t = t_0 + i\beta/2 + u$ ($u < 0$), while in region II it is $t = t_0 + i\tau$.. It can be seen from (27)–(29) that the

A. S. Ioselevich and É. I. Rashba   1114

displacements in region II are determined entirely in terms of ψ^2 in the same region, while the displacements in each of regions I and III are determined in terms of ψ^2 in the same region and also in region II. When the system goes below the barrier, and also when it emerges from below the barrier, its kinetic energy must vanish. In complete agreement with this requirement we easily find from (27)–(29)

$$\partial_u \overline{Q}_I(qu)|_{u=0} = \partial_\tau \overline{Q}_{II}(q\tau)|_{\tau=\beta/2}$$

$$= \partial_\tau \overline{Q}_{II}(q\tau)|_{\tau=0} = \partial_t \overline{Q}_{III}(qt)|_{t=t_0} = 0.$$

Transforming (22), in strict analogy with (27)–(29), we find that the "imaginary" action

$$S = -i\left(\tilde{S} - \int_{r'} dt E(t)\right) \tag{30}$$

is given by

$$S[\psi] = \int_{-\beta/2}^{\beta/2} d\tau \int dr \frac{1}{2m} |\nabla\psi(r\tau)|^2$$

$$- \frac{1}{4} \iint_{-\beta/2}^{\beta/2} d\tau\, d\tau' \iint dr\, dr' \int \frac{dq}{(2\pi)^3}$$

$$\times \frac{\gamma_q^2}{\omega_q} \exp\{iq(r-r')\}\{(N_q+1)\exp(-\omega_q|\tau-\tau'|)$$

$$+ N_q \exp(\omega_q|\tau-\tau'|)\}\psi^2(r\tau)\psi^2(r'\tau'). \tag{31}$$

The second term in (30) results from the substitution into Eq. (21) of expression (25) for the four Ψ functions, which form the coefficient of the exponential function, and from the incorporation in the action of the exponential time factors in those functions.

The function $\psi(r\tau)$ in (31) is found from the condition that the action $S[\psi]$ be stationary with the usual normalization condition. At $T = 0$, expression (31) converts into the corresponding expressions of Ref. 2. We will now classify the solutions of Eq. (26), following Ref. 4.

A. Static solution. We can show that we have $\psi = \psi(r)$, i.e., that a function which is independent of t is a solution of Eq. (26). Substituting $\psi(r)$ into (27)–(29), and evaluating the integrals over the time, we find

$$\overline{Q}_I(q) = \overline{Q}_{II}(q) = Q_{III}(q) = -\frac{\gamma_q}{\omega_q^2} \int dr \psi^2(r) \exp(-iqr). \tag{32}$$

Substitution into (24) yields

$$V(r) = -\int \frac{dq}{(2\pi)^3} \frac{\gamma_q^2}{\omega_q^2} \int dr' \psi^2(r') \exp[iq(r-r')]. \tag{33}$$

The potential is thus independent of the time, and Eq. (26) does in fact have a static solution. Substituting $\psi(r\tau) = \psi(r)$ into (31), we find

$$S_A[\psi] = \beta W,$$

$$W = \int dr \frac{1}{2} |\nabla\psi(r)|^2$$

$$- \frac{1}{2} \iint dr\, dr' \int \frac{dq}{(2\pi)^3} \left(\frac{\gamma_q}{\omega_q}\right)^2 \psi^2(r)\psi^2(r') e^{iq(r-r')}. \tag{34}$$

The expression in braces in (31) is in fact equal to the height of the self-trapping barrier. To demonstrate this point, we construct a Hamiltonian corresponding to Lagrangian (12), (13). The generalized momenta are

$$P_q = \partial L_{tot}/\partial(\partial_t Q_q) = \partial_t Q_{-q}, \quad p_r = \partial L_{tot}/\partial(\partial_t \Psi(rt)) = i\Psi^*, \tag{35}$$

so that we have

$$H = \int \frac{dq}{(2\pi)^3} P_q \partial_t Q_q + \int dr i\Psi^*(rt)\, \partial_t \Psi(rt) - L_{tot}$$

$$= \frac{1}{2} \int \frac{dq}{(2\pi)^3} \{P_q P_{-q} + \omega_q^2 Q_q Q_{-q}\}$$

$$- \int dr\left\{\frac{1}{2m}\Psi^*\cdot\nabla^2\Psi - \int \frac{dq}{(2\pi)^3} \gamma_q Q_q |\Psi(rt)|^2 e^{iqr}\right\}. \tag{36}$$

The first term in (36) is the energy of the lattice, and the second is the energy of the exciton which is interacting with the lattice, as can be seen from (24) and (26). The sum of all the terms except the kinetic energy is the adiabatic potential U. The stationary values of the energy of the lattice at rest can be determined, as in polaron theory,[21] by setting $P_q = 0$, $\delta H/\delta Q_q = 0$, and also $\delta H/\delta\psi^* = 0$ under the normalization condition. As a result, we find expression (34) for H at the stationary saddle point (i.e., for the barrier height $W[1]$).

For the static solution (which describes path A in §1) we thus have $S = \beta W$, and the self-trapping rate is determined by (3).

B. Instanton solutions. The qualitative analysis in Ref. 4 showed that under the condition $T \ll \bar{\omega}$ there exist two instanton solutions. These solutions describe processes in which an exciton is trapped by a deformation well (i.e., $\Psi \neq 0$, $E < 0$) only over a certain region $\tau_0 < \beta/2$ of tunneling region II of contour Γ'. The lattice begins to tunnel while it is still free (Fig. 1).

For a short instanton we have $\tau_0 \lesssim \bar{\omega}^{-+}$. This result means that over nearly the entire tunneling time ($\beta/2 - \tau_0$) the lattice remains free, and it traps an exciton only on the short final section of the tunneling motion. The total energy ε of the tunneling system tends toward zero at $T \to 0$. This solution remains meaningful and leads to a finite action at $T = 0$, and in this limit it agrees with the solution derived in Ref. 2. For a short instanton we have $S \sim W/\bar{\omega}$.

For a long instanton we have $(\beta/2 - \tau_0) \sim \bar{\omega}^{-1}$, so that the region of free tunneling is short. As $T \to 0$, the total energy ε tends toward the barrier height W from below, so that the system emerging from under the barrier at the time t_0 (Fig. 3b) is in the immediate vicinity of the point $\mathscr{W}$. Its velocity is low here, so that the system spends nearly the entire tunneling time near $\mathscr{W}$. For a long instanton we have $(S - \beta W) \sim W/\bar{\omega}$, so that S is larger than for a short instanton, and a long instanton never contributes to the self-trapping rate.

At high temperatures, $T \gg \bar{\omega}$, there are no instanton solutions.[4] It seems natural that these two solutions would disappear, merging with each other. This picture is confirmed by the variational calculations in §4 and by the asymptotic solutions in §6.

In the three following sections we examine the characteristic features of self-trapping and the functional depen-

dence $w(T)$ for the basic models of the exciton-phonon interaction. We find confirmation that the picture drawn above is of general applicability. In these following sections we make explicit use of the continuum approximation for the lattice and of the effective-mass method for the exciton. These approaches are not suitable for the self-trapping state itself, but they may prove legitimate for describing the passage over the barrier if the scale size of the barrier satisfies $r_b \sim \Lambda a_0 \gg a_0$, where a_0 is the lattice constant. The condition $\Lambda \gg 1$ may hold if the exciton-phonon coupling is strong.[7,1] It is not necessary to use the macroscopic approximation, since a mathematical apparatus can be developed for more general models. However, the continuum approximation makes it possible, by examining specific models, to trace the specific features of self-trapping which stem from specific types of exciton-phonon interactions, and in most cases it is possible to completely resolve the problem. We will return to this question in §7.

4. NONPOLAR OPTICAL PHONONS

In this section we assume $\omega_q = \omega_0$ and $\gamma_q = \gamma_0$; the constant γ_0 is related to the analogous constant γ in Ref. 2 by $\gamma_0 = \gamma \omega_0 (2\omega_0 v)^{1/2}$. Since an exciton is electrically neutral, this model describes small-radius excitons which are interacting with polar or nonpolar optical phonons (it is assumed that the exciton has no constant dipole moment). This model also describes the interaction with nonpolar phonons of charge carriers and excitons of large radius if $R_{ex} \gg W$, where R_{ex} is the exciton rydberg. In this section we will express the time in units of ω_0^{-1}.

Transforming (31) with $\gamma_q = \gamma_0$ and $\omega_q = \omega_0$; carrying out the scale transformation

$$\mathbf{r} \to r_b \mathbf{r}, \quad \tau \to \omega_0^{-1}\tau, \quad \psi \to r_b^{-3/2}\psi, \quad r_b \equiv m\gamma_0^2/\omega_0^2;$$

and introducing $S = (\omega_0^3/m^3\gamma_0^4)\mathscr{S}$, we find

$$\mathscr{S}[\psi] = \int d\mathbf{r} \left\{ \int_{-\beta\omega_0/2}^{\beta\omega_0/2} d\tau \cdot \frac{1}{2} |\nabla \psi(\mathbf{r}\tau)|^2 \right.$$
$$- \frac{1}{4} \int\int_{-\beta\omega_0/2}^{\beta\omega_0/2} d\tau\, d\tau' [(N+1)e^{-|\tau-\tau'|}$$
$$\left. + N e^{|\tau-\tau'|}]\psi^2(\mathbf{r}\tau)\psi^2(\mathbf{r}\tau') \right\}. \quad (37)$$

According to (34), the height of the barrier is

$$W = \frac{W_0\omega_0^4}{m^3\gamma_0^4}, \quad W_0 = \int d\mathbf{r}\left\{\frac{1}{2}|\nabla\psi(\mathbf{r})|^2 - \frac{1}{2}\psi^4(\mathbf{r})\right\} \approx 44. \quad (38)$$

This numerical value of the functional W_0 at its lower saddle point is well known.

For instanton solutions $\beta\omega_0/2$, we can replace the integration limits in (37) by the length of an instanton, τ_0, because of the condition $\tau_0 < \beta\omega_0/2$.

At low temperatures the self-trapping rate determines a short instanton. For it, at $T = 0$ (i.e., $N = 0$), expression (37) does not contain any parameters, and the condtion $\delta S = 0$ determines the universal function $\psi_0(\mathbf{r}\tau)$. We can express the instanton action $S_I(T)$ in the low-temperature region in

terms of this function. For this purpose, treating $\mathscr{S}_I$ as a function of ψ and N, we calculate

$$\frac{d\mathscr{S}_I}{dN} = \frac{\partial\mathscr{S}_I}{\partial N} + \int\int d\tau\, d\mathbf{r}\, \frac{\delta\mathscr{S}_I[\psi N]}{\delta\psi(\mathbf{r}\tau)}\, \frac{d\psi(\mathbf{r}\tau)}{dN}. \quad (39)$$

From (37) and (26) we have

$$\delta\mathscr{S}_I[\psi N]/\delta\psi(\mathbf{r}\tau) = 2E(\tau)\psi(\mathbf{r}\tau). \quad (40)$$

It follows from (40) and the normalization condition for ψ that the second term in (39) is zero. Accordingly, the expansion of $\mathscr{S}_I$ at $N \ll 1$ is

$$S_I(T) = \frac{\omega_0^3}{m^3\gamma_0^4}\left\{\mathscr{S}_I(0) - \frac{1}{2}e^{-\beta\omega_0}\int d\mathbf{r}\left[\int_{-\tau_0}^{\tau_0}d\tau\, \mathrm{ch}\,\tau\psi_0^2(\mathbf{r}\tau)\right]^2\right\}. \quad (41)$$

The last term in (41) gives us an explicit expression for the coefficient b in (4). The integral in (41) cannot be evaluated exactly, since ψ_0 is not known, but an accurate estimate can be found for b. From the virial theorem[21,1] written for the functional (37) at $T = 0$ we find

$$4\mathscr{S}_I(0) = \frac{1}{2}\int\int d\tau\, d\tau' \int d\mathbf{r}\psi_0^2(\mathbf{r}\tau)\psi_0^2(\mathbf{r}\tau')\exp(-|\tau-\tau'|)$$
$$< \frac{1}{2}\int\int d\tau\, d\tau' \int d\mathbf{r}\psi_0^2(\mathbf{r}\tau)\psi_0^2(\mathbf{r}\tau')\mathrm{ch}(\tau-\tau')$$
$$= \frac{1}{2}\int d\mathbf{r}\left(\int d\tau\, \mathrm{ch}\,\tau\psi_0^2(\mathbf{r}\tau)\right)^2 \quad (42)$$

Comparing with (4) and (41), we find the inequality $b > 4$.

Since we are calculating only the exponential factor in w, the low-temperature correction to S_I determines the temperature dependence of w only if $S_I(0) - S_I(T) \gg 1$. For $S_I(0) - S_I(T) \lesssim 1$, i.e., at the lowest temperatures, the dependent $w(T)$ is determined by a competition between the temperature dependence of S_I and the coefficient of the exponential function. The latter contribution may dominate. A similar restriction applies to the results of §§6 and 5.

After the system emerges from below the barrier (region III), the self-trapping goes into a stage in which the ψ function contracts rapidly, with a deepening of the deformation well. We now show that after the contraction stage has been established it can be described by a self-similar solution. The system collapses over a finite time of order ω_0^{-1}.

In this stage the deformation is pronounced, so that we need retain only the growing first term in (29); we ignore the (oscillatory) second term. The potential in the Schrödinger equation (26) then becomes

$$V_{III}(\mathbf{r}t) \approx \int_{t_0}^{t} dt'\, \sin(t-t')\psi^2(\mathbf{r}t'). \quad (43)$$

We seek a solution of Eq. (26) for $(t_c - t) \ll 1$ in the form

$$\psi(\mathbf{r}t) = a^{-3/2}(t)\psi_c(\mathbf{r}/a(t)), \quad a(t) = (t_c-t)^\alpha, \quad (44)$$

where t_c is the time of the collapse. Introducing the new variables $\rho = \mathbf{r}/a(t)$ and $\xi = t - t'$ and expanding the sine in (43), we can rewrite (26) as

$$\left\{ E(t) + \frac{1}{2(t_c-t)^{2\alpha}}\nabla_\rho^2 \right.$$
$$\left. + \int_0^{t-t_0}\frac{\xi d\xi}{(t_c-t+\xi)^{3\alpha}}\psi_c^2\left[\rho\left(\frac{t_c-t}{t_c-t+\xi}\right)^\alpha\right]\right\}\psi_c(\rho) = 0. \quad (45)$$

A. S. Ioselevich and É. I. Rashba   1116

The integral converges at $\xi \sim (t_c - t) \ll t_c - t_0$, so that the integration in (45) can be extended to infinity. Using the substitution $x = \xi /(t_c - t)$, and imposing the condition that all the terms in (45) grow the same way in time, we find $\alpha = 2$ and the following equation for $\psi_c(\rho)$:

$$\left(E_c + \frac{1}{2} \nabla_\rho^2 + \int_0^\infty \frac{x dx}{(1+x)^6} \psi_c^2 \left[\frac{\rho}{(1+x)^2} \right] \right) \psi_c(\rho) = 0. \qquad (46)$$

The basic quantities increase in accordance with

$$E(t) = \frac{E_c}{(t_c - t)^4},$$

$$\int d\mathbf{r} Q^2(\mathbf{r}t) \propto (t_c - t)^{-2}, \quad \int d\mathbf{r} [\partial_t Q(\mathbf{r}t)]^2 \propto (t_c - t)^{-4}.$$

Nearly all the lattice energy is kinetic energy, indicating that defect formation may occur in the course of self-trapping.[2] An energy of the order of an electron volt which is liberated as a result of self-trapping is concentrated in the form of kinetic energy of the lattice in a small number of degrees of freedom, the nearest neighbors of the self-trapped exciton. In this sense, the self-trapping of an exciton is analogous to inelastic collisions in a gas. The excess energy of the electronic excitation is initially converted into the kinetic energy of several atoms, and it is not transferred to the lattice in small portions in the form of separate phonons. The energy relaxation of the fast atoms in the lattice can take various paths. The energy may be expended on the formation of lattice defects and/or dissipated among phonons. This picture of the degradation of electronic-excitation energy is considerably more reminiscent of inelastic collisions than was originally proposed by Frenkel[23] and Peierls.[24] The collapse in region III is completely unrelated to the tunneling in region II, so that it should occur even if W is small ($W \lesssim \bar{\omega}$) and also during barrier-free self-trapping in 2D systems (at a surface, for example). In these cases in which self-trapping does not occur in a defect-free sample, but is made possible by impurities, it may lead to the formation of defects near impurity centers.

Figure 4 sketches the time evolution of the displacements, $Q(t)$, and of the exciton binding energy $E(t)$. In the classically accessible region, $Q(t)$ oscillates, while the lattice energy has the behavior $\varepsilon \sim We^{-\beta\omega_0}$. After the stopping

point (the boundary between regions I and II) is crossed, the system undergoes tunneling, and in the region of free motion we have $Q(\tau) \propto \text{ch}(\beta\omega_0/2 - \tau)$. An exciton is trapped at the time $\tau = \tau_0$, and after the second stopping point ($\tau = 0$) is crossed the system emerges from below the barrier and collapses in a time $t_c - t_0 \sim 1$.

The results above exhaust the possibilities of an analytic study of this problem. Quantitative results require a numerical determination of the saddle extremal of functional (37). We use a variational method, choosing $\psi(\mathbf{r}\tau)$ in the square-wave approximation,

$$\psi(\mathbf{r}\tau) = \psi(\mathbf{r}) \theta(\tau_0 - |\tau|), \qquad (47)$$

and then determining ψ and τ_0 from the condition $\delta S = 0$. Substituting (47) into (37), we find

$$\mathscr{S} = \int d\mathbf{r} \left\{ 2\tau_0 \frac{1}{2} |\nabla\psi(\mathbf{r})|^2 - f(\tau_0) \cdot \frac{1}{2} \psi^4(\mathbf{r}) \right\} = W_0 \frac{(2\tau_0)^3}{f^2(\tau_0)}, \qquad (48)$$

where

$$f(\tau_0) = 2\tau_0 + Ne^{2\tau_0} + (N+1) e^{-2\tau_0} - (2N+1). \qquad (49)$$

We have used (38) in deriving the second equation in (48). The time τ_0 is found from the condition for an extremum of the right side of (48). A direct check shows that the values of the total energy at the points $\tau = \tau_0$ and $\tau = 0$ are indeed equal (generally speaking, this equality is nontrivial in a variational calculation). The presence of stopping at $\tau = 0$ follows from the symmetry of the solutions with respect to the sign of τ.

With $\psi_0(\mathbf{r}\tau)$ chosen in the form in (47), we evaluate the integral in (41), and for the coefficient b we find $b \approx 10.3$ from (4).

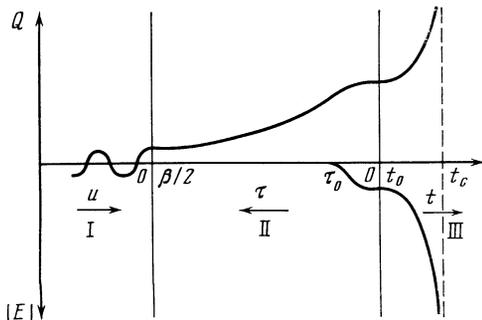Figure 5 shows $\mathscr{S}/W_0 = S\omega_0/W$ as a function of τ_0 for

FIG. 4. Time evolution $Q(t)$ and $E(t)$ for a short instanton (a nonpolar interaction with optical phonons). The upper half of contour Γ' is unfolded along the abscissa. The arrows show the direction in which the time (u, τ, or t) is measured in the corresponding region.
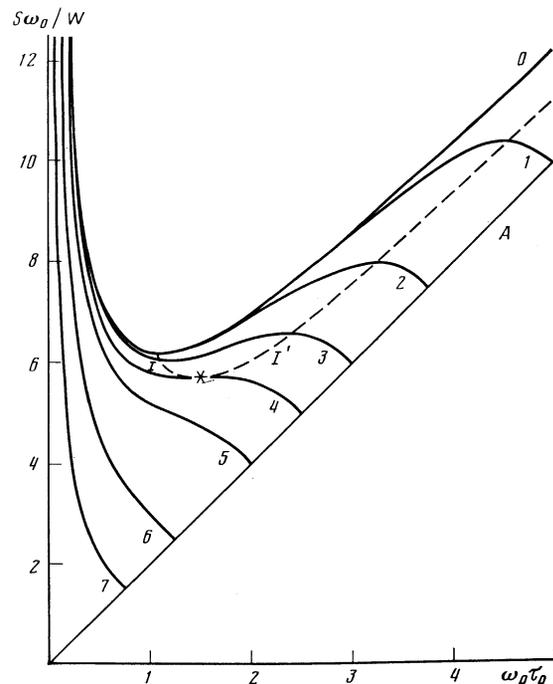

FIG. 5. The action as a function of the duration of an instanton (nonpolar interaction with optical phonons). The parameter of the curves is ω_0/T: 0—∞; 1—10; 2—7.5; 3—6; 4—5; 5—4; 6—2.5; 7—1.5. Line A shows the static solution; I and I' correspond to short and long instantons.

various values of $\beta\omega_0$. On the low-temperature curves there is a minimum corresponding to short instantons, and there is a maximum corresponding to long instantons. The end extrema, which lie on the diagonal $\tau_0 = \beta\omega_0/2$, correspond to the static solution.[3] The instanton extrema are connected by the dashed line. At the minimum on this line (the asterisk), the instantons merge and disappear. In Fig. 2, plotted on the basis of Fig. 5, this point corresponds to the "beak." The point at which curves A and I in Fig. 2 intersect, and at which the nature of the temperature dependence of w changes, corresponds to curve 3 in Fig. 5, for which the values of S at the instanton minimum and at the edge minimum agree. For numerical reasons, the switch in regime occurs early, at $T_c \sim \omega_0/6$, and the asymptotic behavior in (4) holds over the entire region $T < T_c$. The variational calculation thus confirms the general model described in §3 and illustrated in Fig. 2.

Although a short instanton gives rise to a minimum of the action in Fig. 5, and a long instanton gives rise to a maximum, in an infinite-dimensional path space these solutions correspond to more complicated stationary points; e.g., the point I is a saddle point.

5. WANNIER-MOTT EXCITON; POLAR OPTICAL PHONONS

The problem of a Wannier-Mott exciton is not described in the general Lagrangian (13), since an exciton of this sort consists of two particles and therefore has an internal degree of freedom. Kusmartsev and Meshkov[25] have shown, however, that the internal motion can be eliminated if the exciton binding energy satisfies $E_{ex} \gg W$, as it does when the masses of the electron and the hole are very different, $m_h/m_e \gg 1$. Kusmartsev and Meshkov[25] solved the problem at $T = 0$; here we generalize the solution to $T \neq 0$. After the change of variables

$$\mathbf{r} \to r_{ex}\mathbf{r}, \quad \tau \to \omega_0^{-1}\tau, \quad S \to [(\varepsilon_0 - \varepsilon_\infty)/\varepsilon_0 m_e r_{ex}^2 \omega_0]\mathscr{S},$$

where $r_{ex} = \varepsilon_\infty/m_e e^2$ is the exciton radius, and ε_0 and ε_∞ are the static and high-frequency dielectric permeabilities, we find expression (31) for the dimensionless action $\mathscr{S}$ with

$$m = M = \frac{m_h(\varepsilon_0 - \varepsilon_\infty)}{m_e \varepsilon_0},$$

$$\omega_{\mathbf{q}} = 1, \quad \gamma_{\mathbf{q}} = \frac{4\pi}{q}\left\{1 - \left[1 + \left(\frac{q}{2}\right)^2\right]^{-2}\right\}. \quad (50)$$

As Kusmartsev and Meshkov showed,[25] an exciton in a polarization well is in a shallow level ($|E| \ll |V|$ is a characteristic value of the potential energy). This is true both for the static solution, which determines the barrier height,

$$W = W_0(\varepsilon_0 - \varepsilon_\infty)/\varepsilon_0 m_e r_{ex}^2 M^2, \quad W_0 \approx 1.07,$$

and for an instanton. Introducing $\varkappa(\tau) = (2M|E(\tau)|)^{1/2}$, and following the arguments of Ref. 25, we find the following result for an instanton:

$$W_0\left(\frac{2}{E_0}\right)^{1/2} + \varkappa(\tau)$$

$$= \frac{W_0 M}{2E_0}\int_{-\tau_0}^{\tau_0}\{(N+1)e^{-|\tau-\tau'|} + Ne^{|\tau-\tau'|}\}\varkappa(\tau')d\tau', \quad (51)$$

where $E_0 \approx 3.1$. The solution of (5.1) for $M \gg 1$ is

$$\varkappa(\tau) = \frac{(2W_0/M)^{1/2}}{2N+1}\cos\left[\left(\frac{W_0 M}{E_0}\right)^{1/2}\tau\right],$$

$$\tau_0 = \frac{\pi}{2}\left(\frac{E_0}{W_0 M}\right)^{1/2}. \quad (52)$$

The action corresponding to this solution is

$$S_I(T) = 2W\,\mathrm{th}(\beta\omega_0/2). \quad (53)$$

At $T = 0$, this action is the same as that derived in Ref. 25 (aside from a factor of 2 resulting from a misprint in Ref. 25). For the static solution we have

$$S_A = W\beta. \quad (54)$$

We see that at all temperatures we have $S_I < S_A$, and the curves of $S(\beta)$ are tangent at $\beta = 0$. This case corresponds to the limiting case in Fig. 2 with $T_c = T^* = \infty$. We thus find a degenerate situation. At all values of T, the short instanton wins the competition with the Arrhenius solution, but at high values of T the two solutions are essentially indistinguishable. This degeneracy prevails only as $M \to \infty$; at finite values of M we would naturally expect that the usual picture in Fig. 2 would be restored. The low-temperature behavior is described by (4), as it is for nonpolar phonons. We show in §7 that this result is a general result for all optical phonons.

There is another upper limit on M because of the condition for the exciton state to be adiabatic. This condition, $E\tau_0 \gg 1$, can be written in the form $1 \ll M^{1/2} \ll S_I(T)$ with the help of (52) and (53). Since $S_I(T) \gg 1$, there must exist a region in which this inequality holds; this region will contract slightly as the temperature is raised.

6. ACOUSTIC PHONONS

In this section we assume $\omega_{\mathbf{q}} = sq$ and $\gamma_{\mathbf{q}} = \gamma q$. The latter dependence corresponds to the model of a strain energy. This model applies equally well to excitons and charge carriers. Using the change of variables

$$\mathbf{r} \to r_b\mathbf{r}, \quad \tau \to \omega_b^{-1}\tau, \quad \beta \to \omega_b^{-1}\beta,$$

$$S \to \mathscr{S}/mr_b^2\omega_b, \quad r_b = m\gamma^2/s^2, \quad \omega_b = s^3/m\gamma^2, \quad (55)$$

in (31) we find

$$\mathscr{S} = \int_{-\beta/2}^{\beta/2}d\tau\int d\mathbf{r}\frac{1}{2}|\nabla\psi|^2$$

$$- \frac{1}{4}\int\!\!\int_{-\beta/2}^{\beta/2}d\tau\,d\tau'\int\!\!\int d\mathbf{r}\,d\mathbf{r}'\int\frac{d\mathbf{q}}{(2\pi)^3}q$$

$$\times[(N_{\mathbf{q}}+1)e^{-q|\tau-\tau'|} + N_{\mathbf{q}}e^{q|\tau-\tau'|}]\psi^2(\mathbf{r}\tau)\psi^2(\mathbf{r}'\tau')\exp[i\mathbf{q}(\mathbf{r}-\mathbf{r}')]. \quad (56)$$

The barrier height found from (34) leads to the functional (38), as in §4. We thus have

$$W = W_0 s^4/m^3\gamma^4 \approx 44/mr_b^2, \quad S_A \approx 44/mr_b^2 T. \quad (57)$$

The spatial scale of the barrier, $r_b = m\gamma^2/s^2 \approx \Lambda a_0$, is of macroscopic size if $\Lambda \gg 1$. The satisfaction of this criterion justifies the use of the continuum approximation in (57). As was shown in Ref. 2, however, the spatial and temporal

scales of the instanton (a and τ_0) found from (56) at $T = 0$ are zero. This result follows from the virial theorem. At small values of a, however, deviations from the continuum approximation become important. When they are taken into account, a and τ_0 become nonzero, with $a \ll r_b$. Depending on the magnitude and sign of the various corrections, we can have two cases: $a \approx a_0$ or $a_0 \ll a \ll r_b$. In this section of the paper we will consider only the second of these cases. In this second case, we can retain the continuum form of the expression for $\mathscr{S}$ in (56) in a first approximation, assuming that a in this expression is given. In this approximation, the shape of the instanton and its length τ_0 are expressed in terms of a. In particular, we will show that the relation $\tau_0 \sim \omega_b^{-1}(a/r_b)^{3/2}$ holds. Since the characteristic momenta of the phonons which form the instanton are of order a^{-1}, their frequencies satisfy $\omega_q \sim \omega_b r_b /a$. We thus have $\omega_q \tau_0 \sim (a/r_b)^{1/2} \ll 1$, or, in dimensionless units,

$$q\tau_0 \sim a^{1/2} \ll 1. \tag{58}$$

This inequality shows that the potential well varies only slightly over the lifetime of an instanton. Consequently, the local level which arises is shallow over the entire time[4] τ_0.

The wave function of the shallow level is

$$\psi(\mathbf{r}\tau) = \frac{\varkappa^{1/2}(\tau)}{a} e^{-\varkappa(\tau)\tau} \chi\left(\frac{r}{a}\right),$$

$$\chi(\rho) \approx \frac{1}{(2\pi)^{1/2}} \frac{1}{\rho} \quad \text{as} \quad \rho \to \infty. \tag{59}$$

Here $\varkappa(\tau) = (2|E(\tau)|)^{1/2}$, and we have $a\varkappa(\tau) \ll 1$ for all τ.

Using (58) in (56), we can make the replacement

$$(N_q+1)e^{-q|\tau-\tau'|} + N_q e^{q|\tau-\tau'|} \approx (2N_q+1) - q|\tau-\tau'|. \tag{60}$$

For $T = 0$, substituting (59) and (60) into (56), we find, in the zeroth approximation in $q\tau_0$,

$$\mathscr{S} = A_0 y - \frac{1}{4} B_0 y^2, \quad y = \int_{-\tau_0}^{\tau_0} d\tau\, \varkappa(\tau)/a, \tag{61}$$

where

$$A_n = \int d\rho \frac{1}{2} |\nabla^{n+1}\chi(\rho)|^2,$$

$$B_n = \int \frac{d\mathbf{k}}{(2\pi)^3} k^{n+1} \left| \int d\rho\, e^{i\mathbf{k}\rho} \chi^2(\rho) \right|^2. \tag{62}$$

In (62) we have made use of the circumstance that the integrals converge at $\rho \sim 1$, so that we can omit $e^{-\varkappa r} \approx 1$ from (59). Varying (61) with respect to y, we find

$$y = \frac{2A_0}{B_0}, \quad \mathscr{S}_I^{(0)} = \frac{A_0^2}{B_0}, \quad S_I^{(0)} = \frac{A_0^2}{B_0}(mr_b^2\omega_b)^{-1}. \tag{63}$$

The function $\chi(\rho)$ is found from the condition that expression (63) for $\mathscr{S}$ be stationary in the class of functions χ with the asymptotic behavior in (59).

In order to derive an equation for $\varkappa(\tau)$ analogous to (51), we need to retain in (60) the term linear in q. Using Eq. (26), multiplying it by ψ, and integrating with the help of (62), we find, by analogy with (51),

$$\frac{A_0}{a}\varkappa(\tau) - \frac{1}{2}\frac{\varkappa(\tau)}{a^2} \int_{-\tau_0}^{\tau_0} d\tau'\, \varkappa(\tau')\left(B_0 - \frac{B_1}{a}|\tau-\tau'|\right)$$

$$= E(\tau) = -\frac{1}{2}\varkappa^2(\tau). \tag{64}$$

Its solution is

$$\varkappa(\tau) = \varkappa \cos(\pi\tau/2\tau_0)\theta(\tau_0-|\tau|), \tag{65}$$

where

$$\tau_0 = \pi a^{3/2}/2(2B_1)^{1/2}, \quad \varkappa = A_0(2B_1)^{1/2}/B_0 a^{1/2}. \tag{66}$$

It follows that the inequalities $\varkappa a \sim a^{1/2} \ll 1$ and $q\tau_0 \sim \tau_0/a \sim a^{1/2} \ll 1$ hold, confirming the assumptions made earlier. Incorporating the last term in (60), in the substitution into (56), we find a correction $Za^{1/2}$ to the action $\mathscr{S}_I$, where $Z = (\pi/8)(2B_1)^{1/2}A_0^2/B_0^2$.

We turn now to the calculation of $\Delta S_I(T)$—the low-temperature contribution to S_I. By analogy with §4 we write

$$\Delta S_I = -\frac{1}{2}\int\int_{-\tau_0}^{\tau_0} d\tau\, d\tau' \int \frac{d\mathbf{q}}{(2\pi)^3} q N_q \left| \int d\mathbf{r}\psi^2(\mathbf{r}\tau)e^{i\mathbf{q}\mathbf{r}} \right|^2, \tag{67}$$

where ψ is given by (59). Although the values of N_q are not small for $q \lesssim T$ in the limit $T \to 0$, in contrast with the case of optical phonons, the temperature contribution is, on the whole, small because the corresponding phase volume is small, so that expansion (67) is valid. In the integral (67), the characteristic values of q are $\sim \beta^{-1}$. Consequently, the Fourier component of the density ψ^2 in (67) takes different forms for $\beta > \varkappa^{-1}$ and $a < \beta < \varkappa^{-1}$. Its behavior at $\beta < a$ is irrelevant, since there are no instantons in this region, as we will see below. In the first case [$T < \omega_b(r_b/a)^{1/2}$], the integral over $\mathbf{r}$ reduces to a normalization integral, and we can write

$$\Delta\mathscr{S}_I(T) \approx -2\tau_0^2 \int \frac{d\mathbf{q}}{(2\pi)^3} q N_q$$

$$= -\frac{\pi^2}{15}\frac{\tau_0^2}{\beta^4} \equiv -U_1\left(\frac{a}{r_b}\right)^3\left(\frac{T}{\omega_b}\right)^4, \tag{68}$$

$$U_1 = \pi^4/120B_1.$$

Since $r_b > a$, and the transition to the Arrhenius limit occurs at $T_c \sim \omega_b$ [as follows from a comparison of (57) and (63)], expression (68) is valid everywhere at $T < T_c$ (in the limit $\Lambda \gg 1$). Because of the large numerical factor of 44 in (57), however, we actually have $T_c > \omega_b$, and we will therefore also consider the second case: $\omega_b(r_b/a)^{1/2} < T < \omega_b r_b/a$. Under these conditions the Fourier component of ψ^2 is $\pi\varkappa(\tau)/q$ and we have

$$\Delta\mathscr{S}_I(T) \approx -\frac{\pi^2}{6}\left(\frac{A_0}{B_0}\right)^2\left(\frac{a}{\beta}\right)^2 = -U_2\left(\frac{a}{r_b}\right)^2\left(\frac{T}{\omega_b}\right)^2,$$

$$U_2 = \frac{\pi^2}{6}\frac{A_0^2}{B_0^2}. \tag{69}$$

This case can occur only at extremely large values $\gtrsim 10^5$ of Λ.

All the numerical coefficients were determined with the trial function $\chi(\rho) = 2\pi(1+\rho^2)]^{-1/2}$. We then have

A. S. Ioselevich and É. I. Rashba

$$A_0 = \frac{3\pi}{16}, \qquad A_1 = \frac{45\pi}{256}, \qquad B_n = \frac{(n+1)!}{2^{(n+3)}}$$

and

$$\mathscr{S}_I^{(0)} = \frac{9\pi^2}{32} \approx 2.8, \qquad Z = \frac{9\pi^3}{64} \approx 4.4,$$

$$U_1 = \frac{\pi^4}{15} \approx 6.5, \qquad U_2 = \frac{3\pi^4}{8} \approx 36.5. \tag{70}$$

According to (2), the value of $\mathscr{S}_I^{(0)}$ determines the self-trapping rate at $T = 0$. Its value here, 2.8, is half the value of 5.6 which follows from expression (48) in Ref. 2. The reason for the difference is the better choice of the form of the instanton and of the tail function $\chi(\rho)$, which has the correct asymptotic behavior as $\rho \to \infty$, as described in (59).

We have been working in the continuum approximation, expressing all quantities in terms of a. In order to determine a we need to consider the corrections to the action (56). These corrections arise because of the nonparabolic exciton dispersion law and also becuase of the spatial dispersion of the strain energy and of the velocity of sound:

$$\frac{1}{2}|\nabla\psi|^2 \to \frac{1}{2}|\nabla\psi|^2 + \frac{1}{9}\lambda_0 a_0^2|\nabla^2\psi|^2,$$
$$\tag{71}$$
$$\gamma_q \to \gamma q[1 + \lambda_1(a_0 q)^2], \qquad \omega_q \to sq[1 + \lambda_2(a_0 q)^2].$$

As a result, a correction $(a_0/a)^2 Y$ is made to (56), where

$$Y = 2A_0 A_1 \lambda_0/B_0 - A_0^2 B_2(2\lambda_1 - \lambda_2)/B_0^2. \tag{72}$$

In most cases we apparently have $\lambda_0, \lambda_1, \lambda_2 < 0$, and since these quantities appear in Y with opposite signs, the sign of Y is not determined. Collecting all the corrections to the action, we find

$$\Delta\mathscr{S}_I = Za^{1/2} + Y(a_0/a)^2 - U_1 a^3/\beta^4, \quad a^{1/2} \ll \beta,$$
$$\tag{73}$$
$$\Delta\mathscr{S}_I = Za^{1/2} + Y(a_0/a)^2 - U_2 a^2/\beta^2, \quad a \ll \beta \ll a^{1/2}.$$

Here $\mathscr{S}_I^{(0)}$ does not depend on a, we determine a from the condition for $\Delta\mathscr{S}_I$ to have an extremum. In the first region ($\beta \gg a^{1/2}$), the quantity $\Delta\mathscr{S}_I$ has a unique extremum, a minimum if $Y > 0$:

$$a_I = a_0(4Y/Z)^{2/5}(r_b/a_0)^{1/5} \sim \Lambda^{1/5}a_0. \tag{74}$$

For $\Lambda \gg 1$, we can use the continuum approximation, i.e., $a_I \gg a_0$. If $Y < 0$, then $\Delta\mathscr{S}_I$ falls off monotonically with decreasing a. In this case we cannot carry out an expansion in $a_0 q$ as in (71), and we have $a_I \approx a_0$; i.e., the continuum approximation is clearly incorrect. In the second region ($a \ll \beta \ll a^{1/2}$) the quantity $\Delta\mathscr{S}_I$ has, for $Y > 0$, a minimum at $a = a_I(\beta)$, which corresponds to a short instanton, and it has a maximum at $a = a_{I'}(\beta)$ corresponding to a long instanton. If $a_I^{3/4} \ll \beta \ll a_I^{1/2}$, then

$$a_I(\beta) \approx a_I, \quad a_{I'}(\beta) = r_b(Z/4U_2)^{2/3}(\omega_b/T)^{4/3}.$$

For $\beta \sim a_I^{3/4}$, the instantons merge and disappear. At $Y < 0$, there is no minimum, and we have $a_I \approx a_0$. Consequently, for acoustic phonons with $\Lambda \gg 1$ and $Y > 0$ it is possible to analytically follow the picture of the coalescence and disappearance of instantons which was predicted qualitatively in Ref.

4 and which is confirmed by the variational calculations in §4 of the present paper (see Fig. 5).

As in §5, the small depth of the exciton level, $E(\tau)$, and the short duration of the instanton, τ_0, make it difficult to satisfy the adiabatic condition $\tau_0 E(\tau) \gg 1$, and they impose an upper limit on Λ. From (66), (63), and (74) we find

$$(r_b/a_I)^{1/2} \sim \Lambda^{2/5} \ll S_I^{(0)}. \tag{75}$$

If the opposite condition holds, the scheme outlined above for determining a is incorrect. The finite value of a [which is greater than the value given for a_I by (74)] results not from corrections to the continuum model (71), but from corrections to the adiabatic approximation (i.e., the coefficient of the exponential function). In this case the continuum model is logically closed, and we have $a_I \sim r_b(S_I^{(0)})^{-2}$. We will not treat this situation in detail in the present paper, since it requires an analysis of the coefficient of the exponential function. Furthermore, condition (75) apparently does hold in the experiments of which we are aware.

Returning to Fig. 1, we note that there is a difference in the spatial scales for a short instanton and for the A path according to (75): $a_I/r_b \sim \Lambda^{-4/5} \ll 1$. For optical phonons, this ratio is of order unit (§4).

We have yet to discuss the spatial behavior of the potentials $V(\mathbf{r}t)$ [see (26)]. For optical phonons, V is a monotonic function of r, and the width of the well retains a scale size r_b all the way to the point t_0 (Fig. 3b), while the depth of the well increases monotonically. For acoustic phonons the pattern is considerably more complicated. The potential is nonmonotonic, and this behavior is capable in principle of causing qualitative changes in the self-trapping picture. Using (27) and (28) in region I (Fig. 3b; i.e., in the classically accessible region), we find

$$V_I(ru) = -\frac{\pi\tau_0}{2\beta^3 r}\left\{\frac{\text{sh}[\pi(r-u)/\beta]}{\text{ch}^3[\pi(r-u)/\beta]} + \frac{\text{sh}[\pi(r+u)/\beta]}{\text{ch}^3[\pi(r+u)/\beta]}\right\},$$

$$u < 0, \tag{76}$$

while in the region of free tunneling, at $\tau > a^{1/2} > \tau_0$, we find

$$V_{II}(r\tau) =$$

$$\frac{2\pi\tau_0}{\beta^3 r}\text{sh}\left(\frac{2\pi r}{\beta}\right)\frac{\cos^2(2\pi\tau/\beta) - 2 + \cos(2\pi\tau/\beta)\text{ch}(2\pi r/\beta)}{[\text{ch}(2\pi r/\beta) - \cos(2\pi\tau/\beta)]^3}.$$

$$\tag{77}$$

Figure 6 shows r profiles of V_I and V_{II} for various values of u and τ. In contrast with the optical phonons, in addition to the potential well there is a potential barrier here: a barrier for an exciton (not for the lattice, as the self-trapping barrier is). At large negative values of u (region I), i.e., long before the lattice goes under the barrier, the oscillating potential V_I corresponds to a spherical wave converging on a center. By the time $u = 0$, the potential V_I has become a potential of constant sign (attractive; Fig. 6a). In region II, the potential barrier for the exciton reappears, but now surrounded by a potential well near $\mathbf{r} = 0$, whose depth (like the barrier height) increases with decreasing τ, while its width decreases (Fig. 6b). In the later stage of the tunneling, at $\tau < a^{1/2}$, the barrier stabilizes, while the well continues to become deeper.
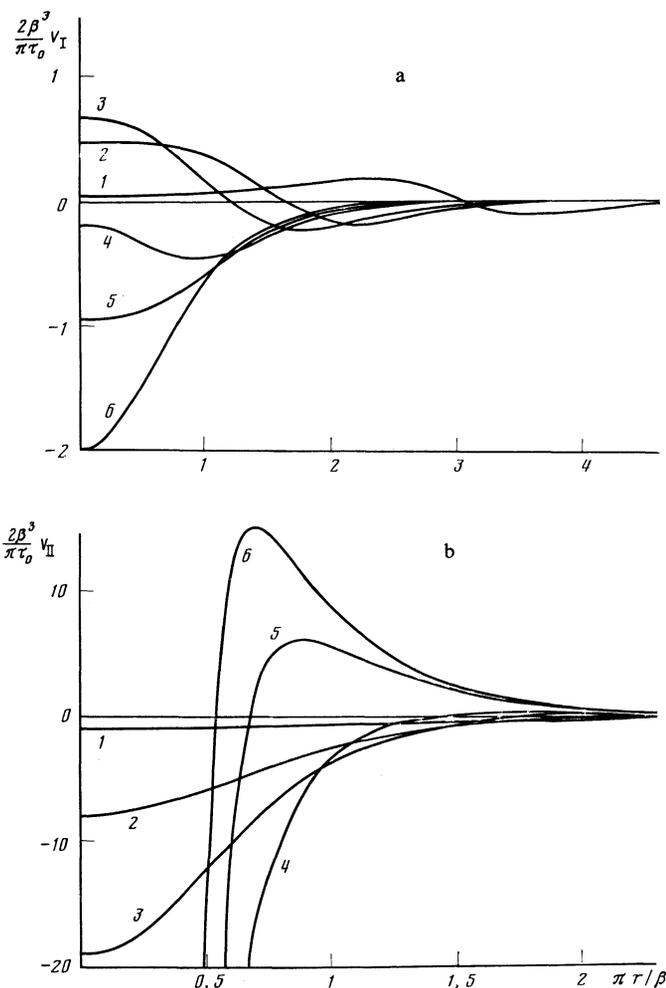
FIG. 6. Spatial distribution of the potentials $V_I(ru)$ and $V_{II}(r\tau)$ (acoustic phonons) for various times u and τ: a: Curve 1—$(-\pi u/\beta) = 3$; 2—1.6; 3—1.2; 4—0.6; 5—0.4; 6—0. b: Curve 1—$\pi\tau\beta = \pi/2$; 2—$\pi/4$; 3—$\pi/5$; 4—$\pi/8$; 5—$\pi/16$; 6—$\pi/20$.

At $\tau \sim a$ the depth reaches $\sim a^{-2}$, and the further changes in the depth are slower. At $\tau = \tau_0$, a local level appears in the well. In the continuum model, the transmission of the exciton barrier is $\sim \exp(-\text{const}\cdot a^{1/4})$ and approximately empty since $a \ll 1$. For realistic values of a, however, the barrier can dominate in the trapping of an exciton.

In region III, the level ceases to be shallow, and the system collapses, by analogy with the case of optical phonons (§4).

Strictly speaking, all the results found here are valid in the limit $\Lambda \gg 1$ and for $Y > 0$. In particular, equating S_A and $S_I^{(0)}$ from (57) and (70), we find

$$T_c \approx 16\omega_b \approx 16\Lambda^{-1}\omega_D, \tag{78}$$

where ω_D is the Debye frequency. For all of the results to be quantitatively correct, we would actually need unrealistically high values $\Lambda \gtrsim 10^2$. From the experimental standpoint, an important parameter is the ratio T_c/ω_D, which determines the point of the change in regime. For optical phonons the corresponding ratio is $T_c/\omega_0 = 1/6$ (§4), while for the acoustic phonons it depends strongly on Λ. At the boundary of the region within which the continuum approximation is applicable, to a short instanton ($\Lambda \approx 10^2$), expression (78) gives us $T_c/\omega_D \approx 0.2$. In the opposite limit, $\Lambda \approx 1$, with $a_I \approx a_0$, we should not expect any significant difference

between T_c/ω_D and T_c/ω_0, since at short wavelengths there is no significant difference at all between acoustic and optical phonons. We should thus not place much importance on the fact that at the actual values $\Lambda \sim 3$–5 expression (78) predicts a value of T_c/ω_D significantly greater than 1. Apparently, T_c/ω_D remains on the order of a few tenths over the entire realistic range of Λ for both acoustic and optical phonons.

To conclude this section, we summarize the conditions for the applicability of the continuum approximation for the interaction of an exciton with acoustic phonons. At the moderate values $\Lambda \sim 5$ we can expect satisfactory results for W and also for $\mathscr{S}_I(T = 0)$ (if $Y > 0$). The expressions for the short instanton, however, especially the analytic expression for the coefficient b in (5) and expression (78) for the ratio T_c/ω_D, are valid only at the unrealistically large values $\Lambda \gtrsim 10^2$ (since fractional powers of Λ up to $\Lambda^{1/5}$ figure in the criteria).

7. DISCUSSION OF RESULTS AND OF EXPERIMENTAL DATA

The basic results derived in §§3–6 hold under extremely general assumptions. The most important point we wish to make is that these results are unrelated to the continuum approximation. The continuum approximation used here has made it possible to pursue the solution of the problem

1121   Sov. Phys. JETP **61** (5), May 1985

A. S. Ioselevich and É. I. Rashba   1121

essentially to its end. As a result (first), we find the picture of the sequence of temperature regimes shown in Fig. 2. There is no reason to doubt that this picture is quite general, but to derive it outside the continuum approximation would require laborious numerical calculations. Second, the continuum approximation has made it possible to derive some simple analytic expressions for W and $w(T)$, describing self-trapping in terms of the macroscopic parameters (the effective mass, the stain energy, the velocity of sound, etc.). Even if these expressions do not give us the accuracy we need (this situation is not surprising, since Λ is not very large), they do give a correct description of the qualitative behavior. Third, the continuum approximation clearly reveals the characteristic features which stem from the interaction with the various branches of the phonon spectrum.

Foremost among the general results is the Arrhenius law (3), with an activation energy equal to the barrier height. It can be shown that this law is valid for an arbitrary structure of the exciton band, for an arbitrary phonon spectrum, and for an arbitrary exciton-phonon interaction. The existence of this dependence, with the barrier height as an activation energy, is not a trivial result. For example, Sumi[26] asserts that law (3), being "classical," should not apply in the quantum region, $T < \bar{\omega}$. Actually, the classical behavior is a consequence of the adiabatic nature of the situation, i.e., the criterion $W/\bar{\omega} \gg 1$, not a consequence of the condition $T > \bar{\omega}$. For this reason, passage over the barrier through an activated process is described by the classical Arrhenius law regardless of the value of T, and the region in which this law describes self-trapping is determined exclusively by the competition between the purely activation process and tunneling. In particular, for nonpolar optical phonons the Arrhenius law describes self-trapping for all T down to $T_c = \omega_0/6 < \bar{\omega}$.

Furthermore, the functional dependence $S(T)$ is general in nature at low temperatures. It is determined exclusively by the type of phonons and by the particular type of exciton-phonon interaction. This fact becomes particularly clear when we calculate the truncated action S_0 through the direct use of the Maupertuis principle. Since the arguments for optical and acoustic phonons are slightly different, we will discuss these two cases separately.

We begin with dispersion-free optical phonons. In this case the temperature dependence of S is determined by the region of free tunneling, in which we have, according to (36),

$$U[Q(\mathbf{r}\tau)] = \frac{1}{2} \omega_0^2 \int d\mathbf{r}\, Q^2(\mathbf{r}\tau). \tag{79}$$

Since there is no dispersion, phonons with all $\mathbf{q}$ oscillate in synchronism, and since they have a common stopping point at the time $\tau = \beta/2$ their phases are also the same. We can thus factorize $Q(\mathbf{r}\tau)$ as $Q(\mathbf{r}\tau) = \Phi(\mathbf{r})\theta(\tau)$ and we find that the free tunneling occurs as it does in a single-mode system with the coordinate $Q(\tau)$. If we choose $\Phi(\mathbf{r})$ to be normalized, we find from (79)

$$U(Q) = \omega_0^2 Q^2/2, \quad S_0(\varepsilon) = \int_{(2\varepsilon)^{1/2}/\omega_0}^{Q_0} (\omega_0^2 Q^2 - 2\varepsilon)^{1/2}\, dQ. \tag{80}$$

Although Q_0 refers to the time of the emergence from under the barrier ($\tau = 0$), where expression (79) has become applicable, and the motion is no longer a single-mode motion, there is no effect on the results, since at small values of ε the behavior $S_0(\varepsilon)$ is determined by the region $Q \ll Q_0$:

$$S_0(\varepsilon) \approx \frac{\omega_0}{2} Q_0^2 - \frac{\varepsilon}{\omega_0} \ln\left(\frac{\omega_0 Q_0}{(2\varepsilon)^{1/2}} \right). \tag{81}$$

Calculating the tunneling time $(-\partial S_0/\partial \varepsilon)$ and equating it to $\beta/2$, we find

$$\varepsilon(\beta) = {}^1\!/_2\omega_0^2 Q_0^2 \exp\{-(1+\beta\omega_0)\}. \tag{82}$$

Substituting (82) into (81) and (2), we find

$$S(\beta) = S(\beta = \infty) - \varepsilon(\beta)/\omega_0. \tag{83}$$

This expression leads to the exponential behavior (4) for $S(T)$. The second term in (41), derived previously, is also equal to $\varepsilon(\beta)/\omega_0$, as can be seen with the help of (28). Since the temperature dependence in (83) is determined exclusively by the region of free motion, the exciton-phonon interaction does not appear explicitly in (83), and this expression is equally valid for excitons, charge carriers, and polar and nonpolar phonons.

For acoustic phonons, the problem is essentially a multimode problem, but the low-temperature behavior of S can nevertheless be found since it is determined by long-wave phonons. Such phonons have essentially no effect on the shape of the exciton wave function ψ, so that the potential energy associated with the mode $\mathbf{q}$ is

$$U_\mathbf{q} = \frac{1}{2} \omega_\mathbf{q}^2 Q_\mathbf{q} Q_{-\mathbf{q}} + \gamma_\mathbf{q} Q_\mathbf{q} \int d\mathbf{r}\, \psi^2(\mathbf{r}) \exp(i\mathbf{q}\mathbf{r}). \tag{84}$$

In the long-wave approximation, the latter integral reduces to a normalization integral at $\tau < \tau_0$ and vanishes at $\tau > \tau_0$. Since $\tau_0 \ll \beta/2$, the quantity $Q_\mathbf{q}$ manages to change very little in the region $\tau < \tau_0$, and we have

$$S_0(\varepsilon_\mathbf{q}, Q_\mathbf{q}^{(0)}) = \int_{(2\varepsilon_\mathbf{q})^{1/2}/\omega\mathbf{q}}^{Q_\mathbf{q}^{(0)}} dQ_\mathbf{q}\, (\omega_\mathbf{q}^2 Q_\mathbf{q}^2 - 2\varepsilon_\mathbf{q})^{1/2} - \tau_0 \gamma_\mathbf{q} Q_\mathbf{q}^{(0)}. \tag{85}$$

The two parameters here, $Q_\mathbf{q}^{(0)}$ and $\varepsilon_\mathbf{q}$, must be found from a condition on the duration of the tunneling, $\partial S_0/\partial \varepsilon_\mathbf{q} = -\beta/2$, and from the presence of a stopping point at $\tau = 0$:

$$P_\mathbf{q} = \partial S/\partial Q_\mathbf{q} = 0,$$

$$\mathrm{Arch}[\omega_\mathbf{q} Q_\mathbf{q}^{(0)}/(2\varepsilon_\mathbf{q})^{1/2}] = \beta\omega_\mathbf{q}/2, \tag{86}$$

$$[\omega_\mathbf{q}^2(Q_\mathbf{q}^{(0)})^2 - 2\varepsilon_\mathbf{q}]^{1/2} - \tau_0\gamma_\mathbf{q} = 0. \tag{87}$$

The first term in (87) is the momentum $P_\mathbf{q}$ at the exit from the region of free tunneling, and the second term is the momentum of the "impact" of the exciton on the mode $\mathbf{q}$ during the brief existence of a local level. Solving (86) and (87), and substituting into (85) and (2), we find

$$S_\mathbf{q}(\beta) = -(\tau_0^2\gamma_\mathbf{q}^2/\omega_\mathbf{q})(1+2N_\mathbf{q}). \tag{88}$$

Integration of the temperature contribution to $\mathbf{q}$ finally yields

$$S(\beta) = S(\beta = \infty) - \tau_0^2 \int \frac{d\mathbf{q}}{(2\pi)^3} \frac{\gamma_\mathbf{q}^2}{\omega_\mathbf{q}} N_\mathbf{q}. \tag{89}$$

A. S. Ioselevich and É. I. Rashba   1122

For the deformation interaction the second term in (89) agrees with (68) and gives us a T^4 law.

Only a single parameter of the instanton—its duration τ_0—figures in expression (89). Consequently, regardless of which phonons form the instanton, the temperature contribution to the acoustic phonons is always described by expression (89). In particular, in those piezoelectric materials in which the self-trapping of charge carriers results from interaction with other groups of phonons, the low-temperature behavior of S is determined by the interaction with piezoelectric-effect acoustic phonons. Since $\gamma_{\mathbf{q}}$ remains finite as $\mathbf{q} \to 0$, the latter phonons give rise to a T^2 law.

It follows in particular that calculations based on the introduction of an "interaction mode"[27,13] are incapable in principle of describing the functional dependence $w = w(T)$. The dependence $S_I(T)$ is determined by phonons with thermal wavelengths and is totally independent of those (shorterwave) phonons which are responsible for the self-trapping at $T = 0$. The process is thus essentially a multimode process, and even if we do speak in terms of an "interaction mode" this mode must have at least two spatial scales, which must be redetermined at each temperature. In contrast, it is not possible to introduce a universal interaction mode, independent of the temperature. An analogous situation arises in a description of long-wave absorption. As was pointed out in Ref. 4, the description of the many-phonon absorption of light by means of a universal interaction mode which was proposed by Sumi and Toyozawa[28] leads to an incorrect exponential dependence of the absorption coefficient on the frequency of the light. It can thus be concluded that the apparent simplification which results from the introduction of an interaction mode is actually illusory, since in order to obtain correct results it is necessary to assume that the structure of this mode is unknown at the outset and to determine it in a self-consistent way.

Since we still lack a complete theory for the temperature dependence of the coefficient of the exponential function, we will simply discuss the experimental data, and this only very briefly. Toyozawa[29] lists the crystals in which self-trapping has been observed and classifies them. In all the experimental studies the region in which w grow rapidly with the temperature is described by an exponential function $w(T) \propto \exp(-\Delta E /T)$. According to recent data, for example, ΔE is 30 and 18 meV in KI and RbI, respectively.[6] If we assume that the interaction with acoustic phonons is dominant near the maximum of their density (~ 7.5 meV), and if we identify ΔE with W, we find $W/\bar\omega \approx 3$, so that the adiabatic approximation is near the limit of its applicability. The temperature at which the exponential region begins corresponds to ≈ 30 K ≈ 2.5 meV $\approx \bar\omega/3$. The estimate $T_c /\bar\omega \approx 1/3$ does not contradict the results of §§4 and 6. In these crystals the half-width of the exciton band is $E_B \approx 0.3$–0.35 eV, so that we have $W/E_B < 0.1$. The small value of this ratio may be taken as an indication that the continuum approximation is valid in the barrier region, since $W/E_B \approx \Lambda^{-2}$. We then find $\Lambda \approx 3$–4. Analysis of data on inert gas crystals leads to similar conclusions. For Xe the activation energy is[30] $W \approx 60$ meV, so that for $\bar\omega \approx 0.5\omega_D \approx 10$ meV we would have

$W/\bar\omega \approx 6$. On the other hand, we have $T_c \approx 50$ K and thus $T_c /\omega_D \approx 0.25$, while we have $W/E_B \approx 60/500 \approx 0.1$. Fugol' and Tarasova's estimate[31] of the parameters on the basis of experimental data showed that Λ increases progressively (from about 2 to 7) as we go from Xe to Ne. We can therefore expect the continuum approximation to yield reasonable estimates for W and $\mathscr{S}_I(T = 0)$. A systematic calculation of W and S_I, however, in the continuum apppproximation, would require a significantly more comprehensive set of crystal parameters than is presently available. The reason is the degeneracy of the bands, which causes an increase in the total number of parameters (effective masses, strain energies, etc.) and a spontaneous lowering of the symmetry of the barrier.[32] The same effect results from the dependence of the coupling constants $\gamma_{\mathbf{q}}$ on the exciton momentum $\mathbf{k}$ (Ref. 26). The existence of self-trapping quasimolecule states is testimony in favor of this dependence.

Roick et al.[8] have studied the low-temperature ($T < 40$ K) behavior $w(T)$ for Xe. They found that at these temperatures the behavior $w(T)$ is not exponential. What evidence is available at this point suggests that this behavior is determined by the coefficient of the exponential function.

We conclude with a look at data on defect formation. That defect formation is possible during the self-trapping stage was first pointed out in Ref. 22, where a specific mechanism for the defect formation was considered: the formation and growth of an exciton cavity under the influence of the quantum pressure of the electron wave function ψ on the lattice. The appearance of fast atoms in the course of the self-trapping (§4) suggests a mechanism in addition to the conventional mechanisms for defect formation[33] and the mechanism of Ref. 22. This other mechanism would operate in an early stage of the relaxation of the self-trapped exciton, specifically, at times on the order of $\bar\omega^{-1}$. This mechanism may be seen both in substances in which cavities do not form and in the initial stage of the formation of a "large" exciton cavity,[22,34] i.e., a cavity from which several atoms have been removed. Defect formation in the course of exciton self-trapping in Ar has recently been detected from measurements of exciton-stimulated desorption by Coletti et al.[35]

[1)]Other extremal paths may also exist, but path A or I would apparently always correspond to the least action.
[2)]This possibility was established in Ref. 22 on the basis of completely different considerations.
[3)]The finite slope of the curves at the end extrema in Fig. 5 would seem at first glance to contradict the assertion, proved in Subsection 3A, that the action is stationary for the static solution. The contradiction is removed by a suitable change in the parametrization near $\tau_0 = \beta\omega_0/2$.
[4)]This circumstance has also been pointed out by F. V.Kusmartsev.

---

[1]E. I. Rashba, in: Excitons (ed. E. I. Rashba and M. D. Sturge), North-Holland, Amsterdam, 1982, p. 543.
[2]S. V. Iordanskiĭ and É. I. Rashba, Zh. Eksp. Teor. Fiz. **74**, 1872 (1978) [Sov. Phys. JETP **47**, 975 (1978)].
[3]A. S. Ioselevich, Zh. Eksp. Teor. Fiz. **81**, 1508 (1982) [Sov. Phys. JETP **54**, 800 (1981)].
[4]A. S. Ioselevich, Zh. Eksp. Teor. Fiz. **83**, 743 (1982) [Sov. Phys. JETP **56**, 415 (1982)].
[5]A. S. Ioselevich and É.I. Rashba, Pis'ma Zh. Eksp. Teor. Fiz. **40**, 348 (1984) [JETP Lett. **40**, 1151 (1984)].
[6]Y. Unuma, Y. Masumoto, S. Shionoya, and H. Nishimura, J. Phys. Soc.

Jpn. **52**, 4277 (1983).

[7]É. I. Rashba, Fiz. Nizk. Temp. **3**, 524 (1977) [Sov. J. Low Temp. Phys. **3**, 254 (1977)].

[8]E. Roick, R. Gaethke, and G. Zimmerer, Solid State Commun. **47**, 333 (1983).

[9]Ch. B. Lushchik, E. A. Vasil'chenko, P. Kh. Liblik, A. Ch. Lushchuk, N. E. Lushchik, Kh. A. Soovik, and M. M. Taĭirov, Trudy inst. fiz. AN Est. SSR (Tartu) **52**, 7 (1981).

[10]J. S. Langer, Ann. Phys. **41**, 108 (1967).

[11]I. M. Lifshits and Yu. Kagan, Zh. Eksp. Teor. Fiz. **62**, 385 (1972) [Sov. Phys. JETP **35**, 206 (1972)].

[12]S. Coleman, Phys. Rev. **D15**, 2929 (1977).

[13]T. Nasu and Y. Toyozawa, J. Phys. Soc. Jpn. **50**, 235 (1981).

[14]K. Huang and A. Rhys, Proc. R. Soc. London **204**, 406 (1950).

[15]M. A. Krivoglaz, Zh. Eksp. Teor. Fiz. **25**, 191 (1953).

[16]R. Kubo and Y. Toyozawa, Prog. Theor. Phys. **13**, 160 (1955).

[17]A. G. Aronov and A. S. Ioselevich, Pis'ma Zh. Eksp. Teor. Fiz. **41**, 71 (1985) [JETP Lett. **41**, 84 (1985)].

[18]R. P. Feynman and A. R. Hibbs, Quantum Mechanics and Path Integrals, McGraw-Hill, New York (1965) (Russ. transl. Mir, Moscow, 1968).

[19]O. V. Konstantinov and V. I. Perel', Zh. Eksp. Teor. Fiz. **39**, 197 (1960) [Sov. Phys. JETP **12**, 142 (1960)].

[20]L. V. Keldysh, Zh. Eksp. Teor. Fiz. **47**, 1515 (1964) [Sov. Phys. JETP **20** 1018 (1964)].

[21]S. I. Pekar, Issledovaniya po élektronnoĭ teorii kristallov (Research in the Electronic Theory of Crystals), Gostekhizdat, Moscow, Leningrad, 1951.

[22]F. V. Kusmartsev and É. I. Rashba, Proceedings of the International Conference on Radiation Physics of Semiconductors and Related Materials (Tbilisi, 1979), (ed. G. P. Kekelidze and V. I. Shakhovtsov), Tbilisi University Press, 1980, p. 448; Czech. J. Phys. **B32**, 54 (1982).

[23]J. Frenkel, Phys. Rev. **37**, 17, 1276 (1931).

[24]R. Peierls, Ann. Phys. (Paris) **13**, 59 (1932).

[25]F. V. Kusmartsev and S. V. Meshkov, Zh. Ekps. Teor. Fiz. **85**, 1500 (1983) [Sov. Phys. JETP **58**, 870 (1983)].

[26]H. Sumi, J. Phys. Soc. Jpn. **53**, 3498, 3511 (1984).

[27]Y. Toyozawa and M. Inoue, J. Phys. Soc. Jpn. **21**, 1663 (1966).

[28]H. Sumi and Y. Toyozawa, J. Phys. Soc. Jpn. **31**, 342 (1971).

[29]Y. Toyozawa, Physica (Utrecht) **117/118B**, 23 (1983).

[30]R. A. Kink, L. É. L. Lykhmus, and M. V. Sel'g, Pis'ma Zh. Eksp. Teor. Fiz. **28**, 505 (1978) [JETP Lett. **28**, 469 (1978)].

[31]I. Ya. Fugol' and E. I. Tarasova, Fiz. Nizk. Temp. **3**, 335 (1977) [Sov. J. Low Temp. Phys. **3**, 160 (1977)].

[32]F. V. Kusmartsev and É.I. Rashba, Pis'ma Zh. Eksp. Teor. Fiz. **33**, 164 (1981) [JETP Lett. **33**, 155 (1981)]; Zh. Eksp. Teor. Fiz. **86**, 1142 (1984) [Sov. Phys. JETP **59**, 668 (1984)].

[33]N. Itoh, Adv. Phys. **31**, 491 (1982).

[34]T. Suemoto and H. Kanzaki, J. Phys. Soc. Jpn. **49**, 1039 (1980).

[35]F. Coletti, J. M. Debever, and G. Zimmerer, J. Phys. Lett. **45**, L-467 (1984).

Translated by Dave Parsons